UNIVERSITY OF LEICESTER

Bayesian methods in health technology assessment: application to overactive bladder syndrome

by

Rhiannon Kate Owen

Thesis submitted for the degree of
Doctor of Philosophy

in the
College of Medicine, Biological Sciences and Psychology
Department of Health Sciences

June 2017

Bayesian methods in health technology assessment: application to overactive bladder syndrome

Rhiannon Kate Owen

This thesis addresses itself to the challenges of health technology assessment (HTA) to inform healthcare policy decision making in the presence of missing or sparse data. HTA advocates the use of evidence synthesis methods. Such an approach involves the analysis of clinical trial data through to the collation, and synthesis, of all relevant information pertaining to the decision question. Motivated by an example in overactive bladder syndrome (OAB), methodological developments together with practical applications are explored throughout the evidence synthesis process. This thesis begins with a novel application of Bayesian methodology to evaluate a large randomised controlled trial (RCT) of repeat treatment in patients with interval-censored data. Performance of Bayesian prediction models were assessed for varying proportions of missing data, and misspecification of distributional form, through a series of simulation studies. Following this, all RCTs evaluating interventions for OAB were identified in a systematic review, and critically appraised. In the current literature, all cross-modality treatment comparisons were performed using pairwise meta-analyses. In this thesis, a cross-modality treatment comparisons was performed using network meta-analysis (NMA) methods in order to obtain treatment effect estimates in terms of efficacy, safety, and tolerability. Network meta-regression techniques were employed to investigate the impact of potential treatment effect modifiers including baseline effects. Building on the general NMA framework, this model was extended to account for similarities between the same interventions with different methods of administration, making use of a natural treatment hierarchy, and where appropriate, incorporating dose-response constraints. Use of hierarchical NMA models increased the precision of treatment effect estimates used for decision-making. The hierarchical NMA model was further extended to incorporate a multivariate approach. This approach borrowed information across outcomes, increasing the precision in the treatment effect estimates. Multivariate hierarchical NMA allowed for the comparison of all interventions across all outcome measures, ameliorating the impact of outcome reporting bias, and thus increasing the ability to make decisions for healthcare policy. In doing so, sacral nerve stimulation (SNS) appeared to be the most promising intervention for the management of OAB.

Acknowledgements

I would like to express great appreciation to my supervisors, Professor Keith Abrams, Professor Douglas Tincello and Dr Sylwia Bujkiewicz for their advice, encouragement, and confidence throughout. Particular thanks to Professor Keith Abrams for giving me an initial insight in to the sphere of medical decision making with an enthusiasm I only wish to emulate.

My personal thanks are extended to the National Institute for Health Research (NIHR) for the many opportunities they have offered throughout my academic career, most noteably, the opportunity to carry out this research. I would also like to thank the staff and students of the University of Leicester's College of Medicine, Biological Sciences and Psychology, with specific thanks to the Biostatistics Research Group. My personal thanks are extended to Professor Alex Sutton and Professor Nicola Cooper for their understanding, continuous support, and contagious enthusiasm. Particular thanks to Dr Mark Rutherford, Dr Pascale Dequen, Dr Eleni Elia, Mr Virag Patel and Dr Jane Bethea who are, and have been, truly great friends.

Last but certainly not least, I would like to offer special thanks to my family and friends for their wholehearted love and support. Particular thanks goes to my parents, Kim and Garry. To my Dad for providing me with a solid mathematical grounding - teaching me the area of a circle, age 2, and proof, age 6 - I'm sure, from which, much of this work stems...! To my Mum, for providing my most valuable lesson in uncertainty, with strength, resilience and incredible work ethic. Special thanks to Nanny O for the incessant laughter, and for always being '*my* rock!' - your wish to attend graduation has been the greatest motivation. To Wendy, for your love and warm welcome was always worth the many miles back home. Many thanks goes to my sister, Carrie, and brother-in-law, Matt - thank you for your constant support and for being my second home! To my niece, Blake, for you are my best form of dissemination, and of course, for your many 'pick me up' smiles. To 'Oldrush, friends from Canada, Leicester, Loughborough, South and West Wales, thank you for the many adventures!

And finally, to the late Rob George, for reminding me always, to '*be happy.*'

"I enjoy communicating science.
It is important that the public understands basic science,
if they are not to leave vital decisions to others."

-Professor Stephen Hawking

Contents

Abstract i

Acknowledgements ii

List of Figures xii

List of Tables xv

Abbreviations xxi

1 Introduction 1
- 1.1 Health technology assessment and health policy decision making . . . 1
- 1.2 Bayesian methodology . . . 2
- 1.3 Detrusor overactivity and overactive bladder: a growing concern for healthcare provision . . . 3
- 1.4 Assessment of interventions for overactive bladder and detrusor overactivity . . . 4
- 1.5 Methodological issues concerning health technology assessment of interventions for overactive bladder . . . 5
- 1.6 Aims of the thesis . . . 7
- 1.7 Structure of the thesis . . . 7

2 Clinical Background 9
- 2.1 Chapter overview . . . 9
- 2.2 Detrusor overactivity (DO) . . . 9
- 2.3 Overactive bladder syndrome (OAB) . . . 10
  - 2.3.1 Urgency . . . 10
  - 2.3.2 Daytime frequency/voiding . . . 10
  - 2.3.3 Nocturia . . . 11
  - 2.3.4 Urinary incontinence . . . 11
  - 2.3.5 Relationship between detrusor overactivity and overactive bladder . . . 11
- 2.4 Prevalence and assessment . . . 12
  - 2.4.1 Bladder diaries . . . 13

2.4.2 Quality of life questionnaires . . . 13
2.4.3 Urodynamic investigation . . . 14
2.5 Key interventions for the management of overactive bladder and detrusor overactivity . . . 14
2.5.1 Lifestyle interventions . . . 15
2.5.2 Physical therapies . . . 15
2.5.2.1 Pelvic floor muscle training (PFMT)/Physiotherapy 15
2.5.2.2 Biofeedback . . . 16
2.5.2.3 Electrical stimulation . . . 16
2.5.2.4 Magnetic stimulation . . . 16
2.5.3 Behavioural therapies . . . 17
2.5.4 Bladder training (BT) . . . 17
2.5.5 Oral drug therapies . . . 17
2.5.5.1 Antimuscarinic drugs . . . 18
2.5.5.2 β_3-adrenoceptor agonists . . . 18
2.5.5.3 Oestrogens . . . 19
2.5.5.4 Other drugs . . . 19
2.5.6 Surgical management . . . 19
2.5.6.1 Sacral nerve stimulation (SNS) . . . 20
2.5.6.2 Botulinum toxin . . . 20
2.5.6.3 Major surgeries . . . 20
2.5.7 Multicomponent therapy . . . 21
2.6 Systematic reviews, meta-analysis and clinical guidelines . . . 21
2.7 Implications for healthcare provision . . . 23
2.8 Chapter summary . . . 23

3 Methodological Background 24
3.1 Chapter overview . . . 24
3.2 Bayesian methodology . . . 24
3.2.1 Markov Chain Monte Chain (MCMC) and Gibbs Sampling . 29
3.2.2 Convergence diagnostics . . . 30
3.2.2.1 Brooks-Gelman-Rubin plots . . . 31
3.2.2.2 Autocorrelation plots . . . 31
3.2.2.3 History/trace plots . . . 32
3.2.2.4 Density plots . . . 32
3.2.3 Goodness of fit and model selection . . . 32
3.2.4 Sampling and prior distributions . . . 33
3.2.4.1 Binomial distribution . . . 34
3.2.4.2 Normal distribution . . . 34
3.2.4.3 Half-normal distribution . . . 35
3.2.4.4 Exponential distribution . . . 35
3.2.4.5 Gamma distribution . . . 36
3.2.4.6 Weibull distribution . . . 36

3.2.4.7 Log-normal distribution . . . 37
3.2.4.8 Generalised gamma distribution . . . 37
3.2.4.9 Uniform distribution . . . 38
3.2.4.10 Multivariate normal distribution . . . 38
3.3 Health technology assessment (HTA) . . . 39
3.4 Evidence-synthesis . . . 39
3.4.1 Fixed effect (FE) model . . . 40
3.4.2 Random effects (RE) model . . . 41
3.4.3 Heterogeneity . . . 42
3.4.3.1 Meta-regression . . . 43
3.4.3.2 Adjusting for baseline risk . . . 44
3.4.4 Pairwise meta-analysis . . . 45
3.4.5 Indirect treatment comparisons and network meta-analysis (NMA) . . . 45
3.4.5.1 Network meta-regression . . . 49
3.4.5.2 Ranking interventions . . . 50
3.4.5.3 Assessing inconsistencies between direct and indirect information . . . 52
3.4.6 Advantages of a Bayesian approach to evidence synthesis . . . 52
3.5 Chapter summary . . . 53

4 Bayesian Methods for Clinical Trial Evaluation 54
4.1 Chapter overview . . . 54
4.2 Introduction . . . 55
4.2.1 The RELAX trial . . . 56
4.3 Methods . . . 58
4.3.1 Model 0: Bayesian parametric frailty model . . . 60
4.3.2 Model 1: Bayesian flexible parametric frailty model . . . 61
4.3.3 Model 2: Bayesian flexible parametric frailty model incorporating covariates . . . 62
4.3.4 Restricted mean survival time (RMST) . . . 64
4.3.5 Model comparison and computation . . . 65
4.4 Results . . . 66
4.4.1 Convergence diagnostics . . . 74
4.5 Simulation study . . . 77
4.5.1 Introduction . . . 77
4.5.2 Methods . . . 77
4.5.2.1 Simulation procedures . . . 77
4.5.2.2 Data generating mechanism . . . 79
4.5.2.3 Performance measures . . . 80
4.5.3 Results . . . 82
4.6 Discussion . . . 84
4.7 Chapter summary . . . 88

5 Systematic Review of Clinical Trials in Overactive Bladder 89
5.1 Chapter overview . . . 89
5.2 Searching and identification methods . . . 90
5.3 Data extraction and estimation methods . . . 93
5.3.1 Continuous data . . . 93
5.3.2 Binary data . . . 95
5.3.3 Predictors of response . . . 95
5.4 Study characteristics . . . 96
5.4.1 Trials not included in evidence synthesis methods . . . 102
5.5 Efficacy data . . . 108
5.6 Safety and tolerability data . . . 108
5.7 Publication bias . . . 109
5.7.1 Efficacy data . . . 110
5.7.2 Safety and tolerability data . . . 111
5.8 Quality assessment and risk of bias . . . 111
5.9 Networks of evidence . . . 115
5.9.1 Efficacy data . . . 115
5.9.2 Safety and tolerability data . . . 117
5.9.3 Interventions not included in the evidence synthesis . . . 118
5.10 Discussion . . . 126
5.11 Chapter summary . . . 128

6 Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder 129
6.1 Chapter overview . . . 129
6.2 Introduction . . . 130
6.3 Methods . . . 131
6.3.1 Network meta-analysis . . . 131
6.3.2 Missing data framework . . . 134
6.3.3 Network meta-regression and adjustment for baseline risk . . . 137
6.3.4 Model computation and convergence diagnostics . . . 138
6.3.5 Assessing inconsistencies between direct and indirect information . . . 138
6.3.6 Goodness of fit and model selection . . . 139
6.3.7 Sensitivity analysis . . . 139
6.4 Results . . . 140
6.4.1 Network meta-analysis of efficacy outcomes . . . 142
6.4.1.1 Incontinence . . . 142
6.4.1.2 Voiding . . . 146
6.4.1.3 Urgency . . . 150
6.4.1.4 Nocturia . . . 152
6.4.2 Network meta-analysis of safety and tolerability outcomes . . . 154

6.4.2.1 Number of patients experiencing adverse events . . 154
6.4.2.2 Discontinuation due to adverse events 155
6.4.2.3 Discontinuation due to a lack of efficacy 155
6.4.3 Treatment profiles 156
6.4.4 Model assessment 159
6.4.4.1 Assessing inconsistency between direct and indirect information 159
6.4.4.2 Convergence diagnostics 160
6.4.4.3 Model fit 160
6.4.5 Sensitivity analysis 161
6.5 Discussion 162
6.6 Chapter summary 168

7 Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder 170
7.1 Chapter overview 170
7.2 Introduction 171
7.3 Methods 174
7.3.1 Class-based network meta-analysis 178
7.3.2 Hierarchical network meta-analysis 180
7.3.2.1 Continuous outcomes 181
7.3.2.2 Binary outcomes 182
7.3.3 Incorporating dose constraints 184
7.3.4 Model computation and convergence diagnostics 185
7.3.5 Assessing inconsistencies between direct and indirect information 185
7.3.6 Goodness of fit and model selection 185
7.3.7 Sensitivity analysis 186
7.4 Results 187
7.4.1 Efficacy data 188
7.4.1.1 Class-based network meta-analysis 189
7.4.1.2 Hierarchical network meta-analysis 192
7.4.1.3 Hierarchical network meta-analysis incorporating dose constraints 193
7.4.2 Safety and tolerability data 200
7.4.2.1 Class-based network meta-analysis 200
7.4.2.2 Hierarchical network meta-analysis 203
7.4.2.3 Hierarchical network meta-analysis incorporating dose constraints 203
7.4.3 Efficacy, safety and tolerability treatment profiles 210
7.4.4 Model assessment 212
7.4.4.1 Assessing inconsistency between direct and indirect information 212
7.4.4.2 Convergence diagnostics 212
7.4.4.3 Sensitivity analysis 213

7.5 Discussion . . . 214
7.6 Chapter summary . . . 218

8 Multivariate Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder 220
8.1 Chapter overview . . . 220
8.2 Introduction . . . 221
8.3 Methods . . . 224
8.3.1 Multivariate network meta-analysis . . . 224
8.3.2 Multivariate hierarchical network meta-analysis . . . 228
8.3.3 Estimating between-study and within-study correlations . . . 231
8.3.4 Model computation and convergence diagnostics . . . 233
8.3.5 Assessing inconsistencies between direct and indirect information . . . 234
8.3.6 Goodness of fit and model selection . . . 234
8.3.7 Sensitivity analysis . . . 235
8.4 Results . . . 236
8.4.1 Multivariate hierarchical network meta-analysis . . . 237
8.4.2 Treatment profiles . . . 246
8.4.3 Model assessment . . . 248
8.4.3.1 Assessing inconsistency between direct and indirect information . . . 248
8.4.3.2 Convergence diagnostics . . . 248
8.4.3.3 Sensitivity analysis . . . 249
8.5 Discussion . . . 249
8.6 Chapter summary . . . 254

9 Discussion and Conclusions 256
9.1 Summary . . . 256
9.1.1 Methodological summary . . . 257
9.1.2 Clinical summary . . . 258
9.2 Discussion and limitations . . . 260
9.2.1 Methodological discussion . . . 260
9.2.2 Clinical discussion . . . 264
9.3 Further work . . . 266
9.4 Conclusions . . . 270

Bibliography 272

A Bayesian Methods for Clinical Trial Evaluation 312
A.1 Visual inspection of basic parametric distributions . . . 312
A.2 Example WinBUGS code - Model 0 . . . 316

A.3 Example WinBUGS code - Model 1 . . . 319
A.4 Example WinBUGS code - Model 2 . . . 322
A.5 Convergence diagnostics . . . 325

B Systematic Review of Clinical Trials in Overactive Bladder 327
B.1 Search Strategy . . . 328
B.2 Study references . . . 329
B.3 Efficacy data . . . 339
B.4 Safety and tolerability data . . . 344

C Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder 348
C.1 Example WinBUGS code for fixed effect NMA incorporating a missing data framework for continuous outcomes . . . 348
C.2 Example WinBUGS code for random effects NMA incorporating a missing data framework for continuous outcomes . . . 351
C.3 Example WinBUGS code for age adjusted random effects NMA incorporating a missing data framework for continuous outcomes . . . 353
C.4 Example WinBUGS code for baseline risk NMA incorporating a missing data framework for continuous outcomes . . . 355
C.5 Example WinBUGS code for fixed effect NMA of binary outcomes . 357
C.6 Example WinBUGS code for random effects NMA of binary outcomes 358
C.7 Results for number of patients experiencing adverse events . . . 360
C.8 Results for discontinuations due to adverse events . . . 362
C.9 Results for discontinuations due to a lack of efficacy . . . 364
C.10 Convergence diagnostics . . . 366
C.11 Assessing inconsistencies . . . 371
C.12 Sensitivity to variance priors . . . 376
C.13 Sensitivity to potentially biased studies . . . 383
C.14 Sensitivity to the inclusion of sacral nerve stimulation . . . 385

D Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder 387
D.1 Example WinBUGS code for hierarchical NMA of continuous outcomes . . . 387
D.2 Example WinBUGS code for hierarchical NMA of continuous outcomes incorporating dose constraints . . . 391
D.3 Example WinBUGS code for hierarchical NMA of binary outcomes 397
D.4 Example WinBUGS code for hierarchical NMA of binary outcomes incorporating dose constraints . . . 399
D.5 Results for efficacy outcomes . . . 405
D.6 Assessing inconsistencies . . . 409
D.7 Convergence diagnostics . . . 411
D.8 Sensitivity analysis . . . 420

E Multivariate Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder **428**
E.1 Example WinBUGS code for multivariate hierarchical NMA 428
E.2 Multivariate NMA results 434
E.3 Convergence diagnostics 436
E.4 Assessing inconsistencies between direct and indirect information . . 442
E.5 Sensitivity analysis 444

F Published Papers **450**
F.1 Comparison of the effectiveness of repeated injections of onabotulinum toxin A for refractory idiopathic detrusor overactivity . . . 451
F.2 Patient factors associated with onabotulinum toxin A treatment outcome in women with detrusor overactivity 459
F.3 Hierarchical network meta-analysis: development of a three-level hierarchical modelling approach incorporating dose-related constraints 466

List of Figures

2.1 Schematic relationship between detrusor overactivity and symptoms of overactive bladder 12

3.1 Impact of an informative prior on the posterior distribution 27
3.2 Impact of a non-informative prior on the posterior distribution . . . 27
3.3 Example of connected and disconnected networks 46

4.1 Example of patient intervals 58
4.2 Repeat treatment timeline 66
4.3 Kaplan-Meir curve for patient reported time to recurrence of symptoms 67
4.4 Predicted survival of the 2-component poly-Weibull model without covariates stratified by active injection 69
4.5 Time to patient reported symptom recurrence for active injection 1 stratified by randomised treatment allocation 75
4.6 Time to patient reported symptom recurrence for active injection 2 stratified by randomised treatment allocation 76
4.7 Flow diagram illustrating the data generating mechanism 80

5.1 Flow diagram of studies identified in the systematic review 98
5.2 Assessment of publication bias: efficacy data 110
5.3 Assessment of publication bias: safety and tolerability data 111
5.4 Between-study quality assessment 112
5.5 Network of evidence (Incontinence) 119
5.6 Network of evidence (Voiding) 120
5.7 Network of evidence (Urgency) 121
5.8 Network of evidence (Nocturia) 122
5.9 Network of evidence (Number of patients with adverse events) . . . 123
5.10 Network of evidence (Discontinuation due to adverse events) 124
5.11 Network of evidence (Discontinuation due to lack of efficacy) 125

6.1 Example network of treatment comparisons for trials reporting treatment effects and measures of uncertainty for urinary incontinence episodes 136
6.2 Heatmap of treatment profiles across all outcomes 158

7.1 Example of intervention hierarchies for the management of OAB . . 173

7.2 Classification of interventions 175
7.3 Network of evidence for classes of interventions (Incontinence) . . . 176
7.4 Network of evidence for classes of interventions (Adverse events) . . 177
7.5 Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in incontinence episodes relative to placebo between individual-intervention, hierarchical, and dose constraint models for the top 10 interventions 199
7.6 Comparison of the estimated odds ratios (and 95% credible intervals) in change from baseline in incontinence episodes relative to placebo between individual-intervention, hierarchical, and dose constraint models 209
7.7 Heatmap of treatment profiles across all outcomes 211

8.1 Network of evidence for multivariate network meta-analysis 237
8.2 Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in incontinence episodes relative to placebo between individual-intervention, hierarchical, and multivariate hierarchical NMA models for the top 10 interventions 239
8.3 Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in voiding episodes relative to placebo between individual-intervention, hierarchical, and multivariate hierarchical NMA models for the top 10 interventions . 240
8.4 Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in urgency episodes relative to placebo between individual-intervention, hierarchical, and multivariate hierarchical NMA models for the top 10 interventions . 241
8.5 Heatmap of intervention profiles for the cardinal symptoms of OAB 247

A.1 Visual inspection of the exponential model 313
A.2 Visual inspection of the generalised gamma model 313
A.3 Visual inspection of the log-normal model 314
A.4 Visual inspection of the Weibull model 314
A.5 Log-log plot of survival for the exponential and generalised gamma models 315

C.1 Brooks-Gelman-Rubin plots for urinary incontinence episodes . . . 366
C.2 Autocorrelation plots for urinary incontinence episodes 367
C.3 History and trace plots for urinary incontinence episodes 368
C.4 Density plots for urinary incontinence episodes 369

D.1 Brooks-Gelman-Rubin plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes 411
D.2 Brooks-Gelman-Rubin plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events . . . 412

D.3 Autocorrelation plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes 413
D.4 Autocorrelation plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events 414
D.5 History and trace plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes 415
D.6 History and trace plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events . . . 416
D.7 Density plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes 418
D.8 Density plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events 419

E.1 Brooks-Gelman-Rubin plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes 436
E.2 Autocorrelation plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes 437
E.3 History and trace plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes 438
E.4 Density plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes 441

List of Tables

4.1 Model fit statistics . . . 68
4.2 Table of results for models without covariate effects . . . 71
4.3 Average patient symptoms at 6 weeks following each active onaBoNT-A injection . . . 73
4.4 Simulation results . . . 83

5.1 Study characteristics . . . 99
5.1 Study characteristics (cont.) . . . 100
5.1 Study characteristics (cont.) . . . 101
5.2 Patient characteristics . . . 105
5.2 Patient characteristics (cont.) . . . 106
5.3 Baseline characteristics of patients receiving first, second, and third-line therapies as recommended by the NICE in England and Wales 107
5.4 Within-study quality assessment . . . 113
5.4 Within study quality assessment (cont.) . . . 114
5.5 Intervention codes . . . 117

6.1 Scenarios for trial reporting . . . 135
6.2 Model fit statistics . . . 141
6.3 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from unadjusted random effects network meta-analysis . . . 144
6.3 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from unadjusted random effects network meta-analysis (cont.) . . . 145
6.4 Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age adjusted random effects network meta-analysis . . . 148
6.4 Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age adjusted random effects network meta-analysis (cont.) 149
6.5 Estimated posterior median difference (and 95% credible interval) in change from baseline for urgency episodes relative to placebo obtained from random effects network meta-analysis adjusted for baseline risk . . . 151

6.6 Estimated posterior median difference (and 95% credible interval) in change from baseline for nocturia episodes relative to placebo obtained from unadjusted random effects network meta-analysis . . 153

7.1 Model fit statistics 188
7.2 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to a control group obtained from a class-based network meta-analysis 191
7.3 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis 194
7.3 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis (cont.) 195
7.4 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis incorporating dose constraints 196
7.4 Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis incorporating dose constraints (cont.) 197
7.5 Estimated posterior odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from a class-based network meta-analysis 202
7.6 Estimated posterior odds ratios (95% credible intervals) for the proportion of patients experiencing adverse events obtained from hierarchical network meta-analysis 204
7.6 Estimated posterior odds ratios (95% credible intervals) for the proportion of patients experiencing adverse events obtained from hierarchical network meta-analysis (cont.) 205
7.7 Estimated posterior odds ratios (95% credible intervals) for the proportion of patients experiencing adverse events obtained from hierarchical network meta-analysis incorporating dose constraints 206
7.7 Estimated posterior odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from hierarchical network meta-analysis incorporating dose constraints (cont.) 207

8.1 Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis 242
8.1 Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis (cont.) 243

8.2 Estimated intervention rank (95% credible intervals) and probability that each intervention is the best for the management of urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis . . . 244
8.2 Estimated intervention rank (95% credible intervals) and probability that each intervention is the best for the management of urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis (cont.) . . . 245

B.1 Study references . . . 330
B.2 Efficacy data . . . 339
B.2 Efficacy data (cont.) . . . 340
B.2 Efficacy data (cont.) . . . 341
B.2 Efficacy data (cont.) . . . 342
B.3 Safety and tolerability data . . . 344
B.3 Safety and tolerability data (cont.) . . . 345
B.3 Safety and tolerability data (cont.) . . . 346
B.3 Safety and tolerability data (cont.) . . . 347

C.1 Estimated odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from random effects network meta-analysis . . . 360
C.1 Estimated odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from random effects network meta-analysis (cont.) . . . 361
C.2 Estimated odds ratios (95% credible intervals) for discontinuations due to adverse events obtained from random effects network meta-analysis . . . 362
C.2 Estimated odds ratios (95% credible intervals) for discontinuations due to adverse events obtained from random effects network meta-analysis (cont.) . . . 363
C.3 Estimated odds ratios (95% credible intervals) for discontinuations due to a lack of efficacy obtained from random effects network meta-analysis . . . 364
C.3 Estimated odds ratios (95% credible intervals) for discontinuations due to a lack of efficacy obtained from random effects network meta-analysis (cont.) . . . 365
C.4 Inconsistency between direct and indirect information for urinary incontinence episodes . . . 371
C.5 Inconsistency between direct and indirect information for voiding episodes . . . 372
C.6 Inconsistency between direct and indirect information for urgency episodes . . . 373
C.7 Inconsistency between direct and indirect information for nocturia episodes . . . 374

C.8 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters . . . 376
C.8 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters (cont.) . . . 377
C.9 Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters . . . 378
C.9 Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters (cont.) . . . 379
C.10 Sensitivity analyses assessing the impact in change from baseline in urgency episodes for different choices of prior distribution on variance parameters . . . 380
C.11 Sensitivity analyses assessing the impact in change from baseline in nocturia episodes for different choices of prior distribution on variance parameters . . . 381
C.12 Sensitivity analysis assessing the impact in change from baseline in urgency episodes having excluded studies A157 and A158 . . . 383
C.13 Sensitivity analysis assessing the impact in change from baseline in incontinence episodes having excluded sacral nerve stimulation from the treatment network . . . 385
C.13 Sensitivity analysis assessing the impact in change from baseline in incontinence episodes having excluded sacral nerve stimulation from the treatment network (cont.) . . . 386

D.1 Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age-adjusted hierarchical network meta-analysis assuming a common regression coefficient . . . 405
D.1 Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age-adjusted hierarchical network meta-analysis assuming a common beta coefficient (cont.) . . . 406
D.2 Estimated posterior median difference (and 95% credible interval) in change from baseline for urgency episodes relative to placebo obtained from a hierarchical network meta-analysis . . . 407
D.3 Estimated posterior median difference (and 95% credible interval) in change from baseline for nocturia episodes relative to placebo obtained from a hierarchical network meta-analysis . . . 408
D.4 Inconsistency between direct and indirect information obtained from hierarchical NMA for urinary incontinence episodes . . . 409
D.5 Inconsistency between direct and indirect information obtained from hierarchical NMA for number of patients experiencing adverse events 410

D.6 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for hierarchical network meta-analysis 420
D.6 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for hierarchical network meta-analysis (cont.) . 421
D.7 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on class-specific variance parameters for hierarchical network meta-analysis 422
D.7 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on class-specific variance parameters for hierarchical network meta-analysis (cont.) 423
D.8 Sensitivity analyses assessing the impact of different choices of prior distributions on variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events 424
D.8 Sensitivity analyses assessing the impact of different choices of prior distributions on variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events (cont.) 425
D.9 Sensitivity analyses assessing the impact of different choices of prior distributions on class-specific variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events 426
D.9 Sensitivity analyses assessing the impact of different choices of prior distributions on class-specific variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events (cont.) 427

E.1 Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate network meta-analysis . 434
E.1 Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate network meta-analysis (cont.) 435
E.2 Sensitivity analysis assessing the impact of studies A157 and A158 on the estimated posterior median differences (and 95% credible intervals) obtained from multivariate hierarchical network meta-analyses 442
E.2 Sensitivity analysis assessing the impact of studies A157 and A158 on the estimated posterior median differences (and 95% credible intervals) obtained from multivariate hierarchical network meta-analyses (cont.) 443

E.3 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis . . . 444
E.3 Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis (cont.) . . . 445
E.4 Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis 446
E.4 Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis (cont.) . . . 447
E.5 Sensitivity analyses assessing the impact in change from baseline in urgency episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis 448
E.5 Sensitivity analyses assessing the impact in change from baseline in urgency episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis (cont.) . . . 449

Abbreviations

BT	**B**ladder **T**raining
CI	**C**onfidence **I**nterval
COMET	**C**ore **O**utcome **M**easures in **E**ffectiveness **T**rials
COS	**C**ore **O**utcome **S**ets
CrI	**Cr**edible **I**nterval
DIC	**D**eviance **I**nformation **C**riterion
DO	**D**etrusor **O**veractivity
ER	**E**xtended **R**elease
FE	**F**ixed **E**ffects
HPD	**H**ighest **P**osterior **D**ensity
HR	**H**azard **R**atio
HTA	**H**ealth **T**echnology **A**ssessment
ICIQ	**I**nternational **C**onsultation on **I**ncontinence Modular **Q**uestionnaire
IDO	**I**diopathic **D**etrusor **O**veractivity
IPD	**I**ndividual **P**articipant **D**ata
I-QOL	**I**ncontinence **Q**uality **o**f **L**ife instrument
IQR	**I**nter**q**uartile **R**ange
IR	**I**mmediate **R**elease
IUSS	**I**ndevus **U**rgency **S**everity **S**cale
MAR	**M**issing **A**t **R**andom
MCDA	**M**ulti-**C**riteria **D**ecision **A**nalysis
MCMC	**M**arkov **C**hain **M**onte **C**arlo
MSE	**M**ean **S**quare **E**rror
MVNMA	**M**ultivariate **N**etwork **M**eta-**A**nalysis

NHS	**N**ational **H**ealth **S**ervice
NICE	**N**ational **I**nstitute for Health and **C**are **E**xcellence
NMA	**N**etwork **M**eta-**A**nalysis
OAB	**O**vera**B**ladder
onaBoNT-A	**Onabo**tuli**n**um **T**oxin Type-**A**
OR	**O**dds **R**atio
PFMT	**P**elvic **F**loor **M**uscle **T**raining
PGI-I	**P**atient **G**lobal **I**mpression on **I**mprovement
RCT	**R**andomised **C**ontrolled **T**rial
RE	**R**andom **E**ffects
RMST	**R**estricted **M**ean **S**urvival **T**ime
SD	**S**tandard **D**eviation
SE	**S**tandard **E**rror
SNS	**S**acral **N**erve **S**timulation
SUI	**S**tress **U**rinary **I**ncontinence
UK	**U**nited **K**ingdom
UUI	**U**rge **U**rinary **I**ncontinence

Chapter 1

Introduction

1.1 Health technology assessment and health policy decision making

Health technology assessment (HTA) is a multidisciplinary process that aims to evaluate the suitability, effectiveness, and cost of emerging healthcare interventions. The main purpose of HTA is to inform health policy decision-making for the individual patient, as well as the healthcare provider (Goodman, 2004). As such, HTA acts as a bridge between medical research and health policy decision making, where scientific findings are conceptualised in terms of the needs and resources of healthcare systems (Battista and Hodge, 1999). HTA has become a valuable tool to examine health technologies in terms of clinical effectiveness, quality of care, patient outcomes, and cost-effectiveness, and is currently used to support reimbursement decisions at regional, national and international levels worldwide (Perry et al., 1997). In England and Wales, the National Institute for Health and Care Excellence (NICE) is tasked with the responsibility of deciding whether new health technologies are reimbursed by the National Health Service (NHS). In doing so, NICE ensure that new interventions are appropriately appraised, NHS expenditure is minimised, and the availability and quality of NHS care is consistent across all NHS trusts (Rawlins and Culyer, 2004). In recent

years, HTA and health policy decision making have given rise to the practice of evidence-based medicine. Evidence-based medicine is an approach to medical practice that ensures that clinical recommendations are established from empirical evidence that is systematically collated from all pertinent trials relating to the decision problem. This approach advocates the use of the highest quality of evidence, which is usually obtained from meta-analyses, systematic reviews, and randomised controlled trials (RCTs).

1.2 Bayesian methodology

One of the defining features of Bayesian statistics is the focus on parameter uncertainty (Spiegelhalter et al., 1999a). Since health policy decision making is often required under a level of uncertainty, Bayesian methods have been widely recognised as a natural approach for performing both clinical and cost-effectiveness analyses (Spiegelhalter et al., 2004; Briggs et al., 1999). Another defining feature of a Bayesian approach is the ability to update prior information in the light of new evidence to form posterior distributions summarising the current state of knowledge regarding a parameter or set of parameters. This creates a flexible framework which allows the incorporation of all available evidence. Previously, complex Bayesian models such as these presented intractable difficulties but with the recent development of Markov Chain Monte Carlo (MCMC) software such as WinBUGS (Spiegelhalter et al., 2003), increasingly complex Bayesian models can be easily estimated in a flexible, simulation based framework (Welton et al., 2012). In a decision making context, incorporating all available evidence pertaining to the decision question is often highly desirable, and this approach is considered more ethical than that of frequentist alternatives as data from all previous participants can be utilised in the decision making framework (Kadane, 1995). Furthermore, a Bayesian approach allows conclusions to be presented in a form that is more accessible for health policy decision makers (such as posterior probability statements), which in the presence of inconclusive evidence more accurately reflects

the degree of belief in the intervention effects of emerging treatments (Lilford and Braunholtz, 1996).

1.3 Detrusor overactivity and overactive bladder: a growing concern for healthcare provision

Urinary incontinence is a common complaint in males and females, generating a need for healthcare intervention in over a third of adults over the age of 40 (McGrother et al., 2004). In the UK, the annual cost burden for treating urinary incontinence is currently in excess of £500 million (Turner et al., 2004). One of the common causes of urinary incontinence is the spontaneous contraction of the detrusor muscle during bladder filling, a condition known as detrusor overactivity (DO). The International Continence Society recognises overactive bladder (OAB) syndrome as the symptomatic manifestation of DO (Abrams, 2003a), where the symptoms of OAB are those of urgency, with or without incontinence, urinary frequency, and nocturia (Diamond et al., 2012). Detrusor overactivity and OAB both have a damaging effect on patients' quality of life across physical, social, cognitive, and financial domains (Girman et al., 1998; Abrams et al., 2000; Stewart et al., 2003; Coyne et al., 2003; O'Donnell et al., 2005; Coyne et al., 2007). Abrams et al. (2000) described the quality of life disturbance in social and functional domains of patients with OAB to be lower than that of patients with diabetes. Certainly, many patients become effectively house-bound and suffer loss of self-esteem and depression.

Treatment of OAB and DO is time-consuming and disappointing from both a patient's and clinician's perspective. For the conservative and minimally invasive

management of OAB symptoms, the National Institute for Health and Care Excellence (2013b) (NICE) in the UK currently recommends a selection of lifestyle interventions, physical therapies, behavioural therapies, oral drug therapies, surgical management, and more recently, β_3-adrenoceptor agonists, sacral nerve stimulation (SNS), and botulinum toxin type A (onaBoNT-A). In the current literature, there is little evidence of the most effective intervention for the management of OAB syndrome, and clinical practice is somewhat variable.

The James Lind Alliance Priority Setting Partnership produced a set of treatment uncertainties relating to urinary incontinence, which was devised from a unique collaboration of patients and clinical experts from several disciplines. Among the top five research priorities was the identification of the most effective interventions for reducing urinary frequency, and urgency (Buckley et al., 2010). It is well recognised by clinicians and patients that the efficacy of antimuscarinic agents is only moderate, and that fewer than 25% of patients prescribed antimuscarinic therapies continue to use them in the long term (Kelleher et al., 1997b). Given the availability of numerous drugs and emerging alternative treatments such as β_3-adrenoceptor agonists, SNS, and onaBoNT-A - which appear more effective but with different side effect profiles and costs - there is an increasing need to identify the most beneficial intervention for the management of OAB and DO. In December 2015, the Chief Medical Officer in the UK - Professor Dame Sally Davies - echoed this concern and called for action on Women's health with specific mention to the management of urinary incontinence (Davies, 2015).

1.4 Assessment of interventions for overactive bladder and detrusor overactivity

There is little evidence of the most effective intervention for the management of OAB and DO. This is largely because many of the published trials report data on interventions versus placebo, or with interventions of a similar nature (i.e. oral

drug therapy versus oral drug therapy). Consequently, there is little information regarding the most effective intervention across different treatment modalities, which makes comparison between active interventions difficult. This is particularly evident for trials evaluating oral drug therapies and onaBoNT-A. In the current literature, a number of published meta-analyses have synthesised data evaluating interventions on a head-to-head basis (Chapple et al., 2008a; Cui et al., 2013, 2014, 2015; Nitti et al., 2013; Sun et al., 2015; Reynolds et al., 2015; Buser et al., 2012; Novara et al., 2008a; Luo et al., 2012; Sweeney et al., 2016; Cao et al., 2016; Huang et al., 2015; Davis et al., 2015; Liu et al., 2014; Wu et al., 2014). Freemantle et al. (2016) used indirect treatment comparisons to compare the efficacy of onaBoNT-A versus mirabegron, but there have been no meta-analyses which have evaluated the comparative efficacy between the diverse range of interventions with different treatment regimes. A multi-modality treatment comparison has the ability to inform patients and physicians of the expected patient benefit and the risk of increased complications from increasing levels of interventions invasiveness. In a HTA context, this can be particularly useful for integrating patients' perspectives, establishing evidence-based treatment pathways, and ultimately informing cost-effectiveness analyses.

1.5 Methodological issues concerning health technology assessment of interventions for overactive bladder

Patients of chronic conditions such as OAB will often require repeat intervention and lifelong follow up. Consequently, in a clinical trial evaluating duration of treatment effect, participant follow-up may continue over several months or years. During this time, patient follow-up is often intermittent and information regarding treatment effect is often unreported or poorly reported. However, patients frequently return for repeat treatment and thus complete information is available for the interval in which the treatment had failed. This is often referred to as

interval censored data and arises when a failure time can not be observed or is missing but is known to fall in an interval obtained from a series of examinations (Peto, 1973). In these situations, a Bayesian approach has the advantage of providing a flexible framework with which to model complex data structures as well as the ability to obtain posterior predictive distributions for missing data.

Network meta-analyses (NMA) are widely used in a HTA setting due to the attractive nature of utilising all relevant information from both direct and indirect evidence (Spiegelhalter et al., 2004; Caldwell et al., 2005; Jansen et al., 2008). In many medical scenarios, there is often a wide array of treatment options, leading to a large network of evidence, and an abundant number of treatment comparisons. This is especially true for the management of OAB, for which there are over 140 different interventions, and thus, there are over 9730 potential pairwise treatment comparisons. For this reason, there is often little or no evidence of the relationship between many of the interventions of interest, both in terms of the number of studies and the number of treatment comparisons. This lack of original evidence makes it difficult to identify the most effective treatment(s) for the management of OAB (Lumley, 2002), and as a result, clinical practice can often be disappointing, and a burden to NHS resources. With the aim to increase precision in treatment effect estimates, similarities between interventions can be accounted for using hierarchical NMAs, incorporating dose-response constraints on increasing doses of interventions (Owen et al., 2015).

Furthermore, for clinical conditions which lack hard outcome measures, such as death, it is common for treatment efficacy to be based on a multitude of outcome measures. This poses two limitations; the first is that the original trials use different outcome measures for the reporting of primary outcomes and therefore different interventions are evaluated for different outcomes (Kirkham et al., 2012); and secondly, for each outcome measure the most effective intervention may differ and consequently it can be difficult to determine which single treatment is the most effective across the entire symptom syndrome (Bower et al., 2011). To overcome

these limitations, hierarchical NMAs can be extended to incorporate the correlation between outcomes in order to predict a distribution for missing outcome data using a multivariate approach (Achana et al., 2014).

1.6 Aims of the thesis

This thesis aims to systematically review and critically appraise all interventions for the management of OAB and DO. To facilitate this goal, a number of methodological developments are described and demonstrated in the context of HTA. Health technology assessment can be seen as a continuum of analyses from which current best evidence is summarised to form population-based clinical conclusions. At the foundation of this continuum are RCTs, and at the apogee are clinical- and cost-effectiveness analyses. This thesis aims to develop and apply methodologies for clinical trial evaluation and evidence synthesis in order to maximise current information with the view to aid decision making from patients', physicians' and health policy decision makers' perspective. By synthesizing existing data on clinical benefits, harms, and tolerance of individual interventions, the empirical findings presented in this thesis aim to provide the current best evidence to decision makers who are tasked with deciding which interventions to recommend for clinical practice.

1.7 Structure of the thesis

The remainder of this thesis is organised as follows. Chapter 2 begins with a more detailed introduction to the motivating clinical example in OAB syndrome. In this chapter, clinical definitions and terminologies used throughout this thesis are defined, existing interventions are introduced, and a brief overview of the current literature, together with the implications for healthcare decision making, are outlined. Chapter 3 introduces the fundamental statistical methodologies that

are implemented and developed throughout the succeeding chapters of this thesis. In this chapter, key assumptions and limitations of current methodologies are highlighted and discussed. One of the common limitations of many statistical methodologies is that of missing data. Chapter 4 aims to address this issue using novel Bayesian methods for repeated event data in RCTs of chronic conditions. In this chapter, Bayesian methods are applied to time-to-event data from a large RCT in OAB - the RELAX trial (Tincello et al., 2012) - and simulation studies are undertaken to assess the performance of the proposed Bayesian models. Evidence-based medicine is embedded in the HTA process and goes beyond that of assessing individual RCTs, whereby all relevant evidence for newly emerging and existing interventions are acquired and summarised. Chapter 5 collates data from all pertinent RCTs in OAB, critically appraises the evidence, and highlights some of the limitations of the published literature. Chapter 6 quantifies this information for both efficacy and safety outcomes using various network meta-analysis (NMA) techniques. In this chapter, many of the data limitations identified in Chapter 5 are explored and addressed using a series of NMA methods. Chapter 7 builds on this methodology with the aim to increase precision in the treatment effect estimates for decision making. To do this, similarities between the same interventions with different methods of administration are accounted for, making use of a natural treatment hierarchy, and constraints are placed on increasing doses of interventions. Chapter 8 extends the hierarchical NMA described in Chapter 7 to incorporate a multivariate approach. This approach utilises the correlation between outcomes and jointly synthesises data from multiple outcome measures in a single coherent analysis. The primary reason for developing multivariate analyses were largely motivated by the findings of Chapter 5, and aims to ameloriate potential outcome reporting bias in the original trials. Chapter 9 concludes the thesis with a summary of the important clinical and methodological findings, together with a discussion of the limitations, and the opportunities for further work.

Chapter 2

Clinical Background

2.1 Chapter overview

This chapter provides a clinical background for detrusor overactivity (DO) and overactive bladder (OAB) syndrome. Clinical definitions and terminology used throughout this thesis are explained, and existing interventions are introduced. A brief overview of the current literature, clinical consequences and future concerns for healthcare provision are described.

2.2 Detrusor overactivity (DO)

Detrusor overactivity is characterized by the spontaneous contraction of the detrusor muscle during bladder filling, with no defined cause. Individuals are diagnosed with DO if there is an apparent involuntary contraction of the detrusor muscle upon urodynamic testing. Detrusor overactivity can occur in both men and women alike, though it is more common in women (Abrams, 2003b). The known causes for DO are multifaceted and can be summarised as one of two mechanisms: neurogenic or idiopathic. Neurogenic DO occurs as a result of damage, or injury, to an individual's neural network that impairs the signal between the central nervous system, and the bladder (e.g, patients with spinal cord injury) (Haylen et al.,

2010). Idiopathic DO is defined as an increased excitability and pressure in the bladder with no defined cause (Abrams et al., 2002). Both forms of DO often cause individuals to experience symptoms of OAB syndrome.

2.3 Overactive bladder syndrome (OAB)

In comparison to DO, OAB is a symptom syndrome without definitive diagnosis. The symptoms of OAB are increased urgency, daytime frequency, nocturia, and urinary incontinence, though the cardinal symptoms of OAB are said to be urgency, daytime frequency, and urinary incontinence (Abrams, 2005). In 2002, the International Continence Society (ICS) defined standardised terminology for symptoms of overactive bladder (Abrams et al., 2002):

2.3.1 Urgency

Urgency is defined as 'the sudden compelling desire to pass urine which is difficult to defer'. In the previous literature there has been much controversy over the definition of urgency, and although it is considered the key symptom of overactive bladder syndrome, confusion in terminology and characterization has made it difficult to measure in an evaluative setting (Gormley et al., 2012).

2.3.2 Daytime frequency/voiding

Increased daytime frequency, also known as pollakisuria or increased voiding, is defined as 'the complaint by the patient who considers that he/she voids too often by day.' Healthy individuals are considered to experience up to 8 voids per 24 hours and thus any value above this range indicates OAB (Kreder and Dmochowski, 2007).

2.3.3 Nocturia

Nocturia is defined as 'the complaint that the individual has to wake at night one or more times to void.' It is thought that individuals with nocturia will return to full sleep following a voiding episode. Without returning to full sleep individuals are likely to suffer from a sleeping disorder (van Kerrebroeck et al., 2002).

2.3.4 Urinary incontinence

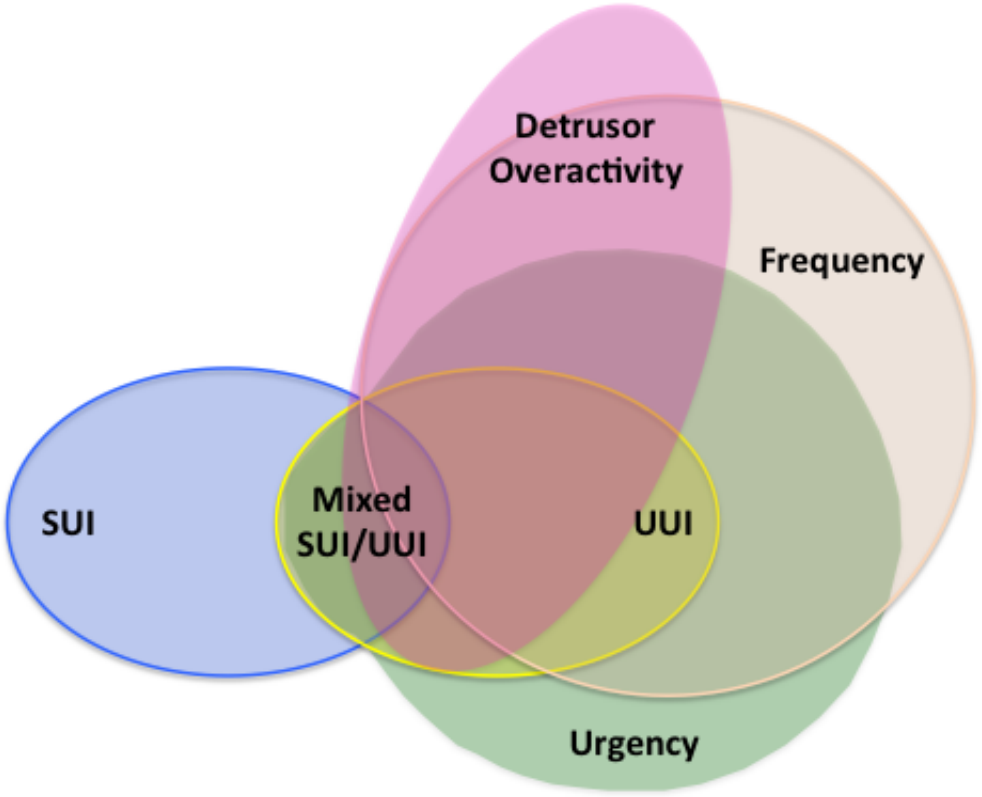
Urinary incontinence is defined as 'the complaint of any involuntary leakage of urine.' There are three common types of urinary incontinence, urge (UUI), stress (SUI) and mixed incontinence, each with a different aetiology. Urge urinary incontinence is associated with involuntary leakage of urine due to a sudden and forceful desire to void. Stress incontinence occurs as a result of physical exertion such as sneezing, coughing, laughing, or exercising. Mixed urinary incontinence is a combination of both urge and stress-related urinary incontinence episodes. It can be difficult to differentiate between subtypes of urinary incontinence, and thus trials in overactive bladder syndrome will often measure the total number of urinary incontinence episodes (Gormley et al., 2012). Individuals with symptoms of OAB including urinary incontinence are considered to have OAB wet, and those without symptoms of urinary incontinence are considered to have OAB dry (Tubaro, 2004).

2.3.5 Relationship between detrusor overactivity and overactive bladder

There is a reasonable degree of overlap between the conditions of OAB and DO, which has led to confusion in terminology. Figure 2.1 illustrates the schematic relationship between the symptomatic diagnosis of OAB syndrome and the urodynamic diagnosis of DO. Whilst these conditions co-exist in the majority of individuals, it is possible to experience asymptomatic DO as well as a combination,

or isolated symptoms of OAB syndrome. Throughout this thesis, I will refer to OAB as a collective term encompassing both patients with symptoms of OAB and patients diagnosed with symptomatic DO.

FIGURE 2.1: Schematic relationship between detrusor overactivity and symptoms of overactive bladder



SUI denotes stress urinary incontinence; UUI denotes urge urinary incontinence. Adapted from a figure by (Abrams, 2003b).

2.4 Prevalence and assessment

Overactive bladder is most common in individuals over 40 years of age (Milsom et al., 2001). It is possible for children and young adults to present with symptoms of overactive bladder, but the aetiology of the symptom syndromes are considered to be different for different age groups. Epidemiological studies estimate the prevalence of overactive bladder to range between 7-27% in men, and 9-43% in women

(Choo et al., 2007; Corcos and Schick, 2004; Coyne et al., 2011; Irwin et al., 2006; Milsom et al., 2001; Stewart et al., 2003; Tikkinen et al., 2008; Herschorn et al., 2008). Approximately 400 million of the 2008 worldwide population experiences OAB, with that number predicted to rise to 546 million by 2018 (Irwin et al., 2011). Due to the embarrassing nature of the condition, prevalence of OAB is thought to be under-reported (MacDiarmid, 2008) and thus these estimates remain conservative.

Overactive bladder syndrome is most commonly assessed though patient-reported bladder diaries, quality of life questionnaires and urodynamic investigation.

2.4.1 Bladder diaries

Individuals with OAB are regularly asked to record the frequency and severity of symptoms in patient-reported bladder diaries. A bladder diary records symptoms of urinary incontinence, voiding, urgency and nocturia on a daily basis. The National Institute for Health and Care Excellence (2015) recommend the use of urinary bladder diaries for a minimum of 3 days (Tincello et al., 2007). Typically, average symptom profiles are calculated per 24 hours. Patient reported bladder diaries are considered a gold standard tool for assessing the quantitative measure of OAB symptoms (Homma et al., 2011).

2.4.2 Quality of life questionnaires

Overactive bladder syndrome has a detrimental impact on all aspects of life including social, work-related, sexual, psychological and emotional aspects (Coyne et al., 2003). To capture this diverse burden, a number of quality of life questionnaires are used. These include general quality of life questionnaires such as EQ-5D (Gusi et al., 2010), as well as urinary incontinence-specific questionnaires such as the Incontinence Quality of Life instrument (I-QOL) (Wagner et al., 1996), International Consultation on Incontinence Modular Questionnaire (ICIQ) (Abrams

et al., 2006), and OAB-specific questionnaires such as King's Health Questionnaire (KHQ) (Kelleher et al., 1997c), Overactive Bladder questionnaire (OAB-q) (Coyne et al., 2002), and the Patient Perception of Bladder Condition (PPBC) (Coyne et al., 2006).

2.4.3 Urodynamic investigation

Urodynamic testing involves the insertion of a small bladder and rectal catheter to fill the bladder in order to investigate bladder condition. Volume of urine voided, rate of emptying, and pressure in the bladder, are often recorded. Urodynamic testing is not as common as other forms of OAB assessment as it is costly and highly invasive (Digesu et al., 2003), though it is regularly used to definitively diagnose DO.

2.5 Key interventions for the management of overactive bladder and detrusor overactivity

There is a large and diverse range of interventions for the management of OAB and DO. In 2013, the National Institute for Health and Care Excellence (2013b) (NICE) in England and Wales formulated a set of clinical guidelines. Key interventions recommended by NICE can be categorized into several broad classes including lifestyle interventions, physical therapies, behavioural therapies, oral drug therapies, and surgical management. Traditionally, DO is managed by multicomponent therapy - that is, the combination of drug therapies with other conservative methods for treating symptoms of OAB such as physical and behavioural therapies. In more extreme cases, surgery is considered as a second and third line intervention. More recently, emerging interventions such as injections of botulinum toxins and sacral nerve stimulation (SNS) have shown promising results and are less invasive than other forms of surgery.

2.5.1 Lifestyle interventions

Lifestyle interventions are procedures aimed at modifying lifestyle factors associated with OAB and DO; including, but not limited to, modifying dietary habits, caffeine intake, fluid consumption, smoking habits, weight loss, and physical exercise (Wyman et al., 2009).

2.5.2 Physical therapies

Physical therapies are used for both prevention and management of OAB and include pelvic floor muscle training (PFMT), biofeedback, electrical stimulation, and magnetic stimulation (National Institute for Health and Care Excellence, 2013b). The most commonly used physical therapy is PFMT (Dumoulin and Hay-Smith, 2008).

2.5.2.1 Pelvic floor muscle training (PFMT)/Physiotherapy

The pelvic floor is made up of 3 layers of striated muscles (Bø, 2004). The main role of the pelvic floor is to provide support for the core pelvic structures. It is these muscles that simultaneously contract to move the pelvic girdle in one direction. PFMT involves the repetitive contraction of the pelvic floor muscle, which aims to strengthen and build muscle tone, and provide perineal support (Price et al., 2010). Individuals are required to produce a voluntary inward and upward contraction, and hold, of the pelvic floor for a minimum of 8 contractions, 3 times a day (National Institute for Health and Care Excellence, 2013b). Pelvic floor muscle training, more broadly referred to as physiotherapy, is recommended for a minimum of 3 months as first-line therapy for OAB (National Institute for Health and Care Excellence, 2013b).

2.5.2.2 Biofeedback

Biofeedback, or applied psychophysiological feedback, teaches individuals to intrinsically control muscle tension, pain, and neural fluctuations through guided cognitive techniques, such as visualization and relaxation (Abrams et al., 2002). Biofeedback techniques are often used to facilitate PFMT (Herderschee et al., 2011).

2.5.2.3 Electrical stimulation

Electrical stimulation incites muscle contraction by sending an electric current to the pelvic floor muscles or nerves in the lower back. It is thought that electrical stimulation strengthens muscle tone and stimulates growth of the nerve cells responsible for bladder contraction (Abrams et al., 2002). There are several methods of electrical stimulation including pelvic floor electrical stimulation, posterior tibial nerve stimulation and transcutaneous electrical nerve stimulation. Pelvic floor electric stimulation directly stimulates the pelvic floor muscles by inserting a small vaginal electrode (Siegel et al., 1997). This form of electrical stimulation is often used to aid PFMT (Dumoulin et al., 2010). Posterior tibial nerve stimulation directly targets stimulation of the tibial nerve by inserting a small electrode in the lower leg (Vandoninck et al., 2003). For the treatment of urinary incontinence, transcutaneous electrical nerve stimulation specifically targets the S2 to S4 dermatomes (Hagstroem et al., 2009). For all methods of electrical stimulation, patients require electrical current over long periods of time (National Institute for Health and Care Excellence, 2013b).

2.5.2.4 Magnetic stimulation

Similar to electrical stimulation, magnetic stimulation aims to excite the pelvic floor muscles and sacral nerve root. However, magnetic stimulation induces an electric current by rapidly changing magnetic fields (But, 2003).

2.5.3 Behavioural therapies

Behavioural therapy involves the adaptation and rehabilitation of learnt behaviours. There are many types of behavioural therapies including bladder training, prompted voiding, and timed voiding (Abrams et al., 2002). For overactive bladder syndrome, it is common for behavioural therapies to be combined with alternative therapies to form a multicomponent treatment plan (Burgio et al., 2000; Wyman et al., 1998).

2.5.4 Bladder training (BT)

Bladder training is a multifaceted programme that can be made up of one or more techniques including dietary modification, delayed and timed voids, PFMT and biofeedback (Marinkovic et al., 2012). Dietary management usually includes a conscious reduction in fluid and caffeine intake (Gormley et al., 2012). Delayed and timed voids involve a disregard for the normal desire to void, and voiding only at timed, and pre-set intervals (Eustice et al., 2000). To begin with, voiding intervals tend to be as short as every 30 minutes. These intervals are slowly increased until the individual can maintain control over several hours (Burgio et al., 2000). This continuous routine will eventually increase bladder capacity and reduce urgency. NICE recommend a minimum duration of 6 weeks of bladder training as first-line therapy (National Institute for Health and Care Excellence, 2013b).

2.5.5 Oral drug therapies

Oral drug therapy, also called pharmacotherapy, includes oral medications that usually interact with receptors or enzymes in cells to reduce symptoms. Oral drug therapies for the management of OAB include antimuscarinic drugs, β_3-adrenoceptor agonists, oestrogens and other drugs. However, it is known that long term compliance with oral drug therapy is generally low (Kelleher et al., 1997a).

2.5.5.1 Antimuscarinic drugs

Antimuscarinic drugs, also known as anticholinergics, work in one of two ways. The first mechanism inhibits the binding of acetylcholine to receptors in the detrusor muscle, and the second inhibits urothelial sensory receptors which reduces afferent nerve activity. Previous studies (Robinson and Cardozo, 2012; Madhuvrata et al., 2012; Chapple et al., 2008b; Novara et al., 2008b) have shown that antimuscarinics are safe and effective for treating symptoms of OAB. However, many patients fail to tolerate antimuscarinics in the first three months after prescription due to widely documented side effects such as dry mouth, blurred vision and constipation (Cardozo et al., 2004; Chapple et al., 2008b). In order to address these side effects whilst maintaining efficacy, different formulations of the same antimuscarinic drug have been evaluated, such as transdermal (Dmochowski et al., 2003), and intravesical formulations (Kim et al., 2005). The most commonly used antimuscarinics in current clinical practice are oxybutynin, tolterodine, solifenacin, fesoterodine, imidafenacin, darifenacin, propiverine and trospium. NICE recommends immediate release formulation of non-proprietary oxybutynin as second line treatment for OAB, provided that bladder training and other physical and behavioural therapies have failed in the first instance (National Institute for Health and Care Excellence, 2013b).

2.5.5.2 β_3-adrenoceptor agonists

β_3-adrenoceptor agonists are pharmacological agents that stimulate the β_3 receptor predominantly found in the urinary bladder. Sympathetic stimulation of the β_3-receptor, the most commonly found β receptor subtype in the bladder, is thought to result in increased relaxation and increased bladder compliance (Thiagamoorthy et al., 2015).

2.5.5.3 Oestrogens

Oestrogens play an important role in sustaining the health of essential tissues involved in the maintenance of normal pressure transmission in the urethra of females (Hextall, 2000). Oestrogen deficient women tend to experience increased sensation and increased urgency (Robinson and Cardozo, 2003). For this reason, oestrogen deficiencies may be an aetiological factor associated with OAB in women. Oestrogen replacement is thought to improve and sustain the health of these essential tissues, and retain bladder function (Robinson and Cardozo, 2003). NICE currently recommends oestrogens for the treatment of OAB in post-menopausal women (National Institute for Health and Care Excellence, 2013b).

2.5.5.4 Other drugs

There are several existing and newly emerging drugs that have been evaluated for OAB and which work in different ways to those outlined above. Two of the most commonly used drugs are Duloxetine and Desmopressin. Duloxetine is a serotonin and noradrenaline reuptake inhibitor that targets the sacral spinal cord. Duloxetine is currently licensed specifically for SUI (Basu and Duckett, 2009), but has recently been evaluated for treating symptoms of OAB (Steers et al., 2007). Desmopressin is an antidiuretic hormone that inhibits diuresis, and is specifically recommended for treating nocturia (Lose et al., 2004).

2.5.6 Surgical management

Surgical management is currently offered when conservative medical measures have failed. Surgical management options currently include sacral nerve stimulation, botulinum toxin and major surgeries.

2.5.6.1 Sacral nerve stimulation (SNS)

Sacral nerve stimulation (SNS), or sacral neuromodulation, is a type of electrostimulation of the sacral reflex pathway with the aim to inhibit bladder reflex (Chancellor and Chartier-Kastler, 2000). Different to electrical stimulation, SNS is considered to be a more invasive intervention that involves the subcutaneous implantation of a programmable stimulator - similar to a pacemaker - which delivers low amplitude electrical current via a lead to the sacral nerve roots S2 and S3 (Leng and Chancellor, 2005).

2.5.6.2 Botulinum toxin

There are 7 types of botulinum toxins, A, B, C1, D, E, F and G. Types A and B are used to treat DO (Orasanu and Mahajan, 2013), though most studies explore the use of onabotulinum toxin type A (onaBoNT-A) (Lie et al., 2010). NICE currently recommend the use of onaBoNT-A but do not recommend the use of botulinum toxin type B for the treatment of overactive bladder (National Institute for Health and Care Excellence, 2013b). Botulinum toxins work by preventing the release of acetylcholine from cholinergic nerve terminals. This reduces the contractility of the muscle at the site of injection (Apostolidis et al., 2006). Usually the toxin is injected in to the bladder wall sparing the trigone. More recent studies have investigated injection into the bladder wall, base, and trigone (Kuo, 2011). Though there is a risk of urinary retention in patients receiving botulinum toxin, and intermittent self-catheterization is often required (Shaban and Drake, 2008).

2.5.6.3 Major surgeries

There are a number of alternative surgical interventions that require more invasive procedures involving the removal or transposal of the bladder wall, muscle and urethra. These include augmentation cystoplasty (Flood et al., 1995), detrusor myectomy(Swami et al., 1998), and urinary diversion (Singh et al., 1997). Surgery of this type often require life-long follow-up (National Institute for Health and

Care Excellence, 2013b) and can be largely detrimental to quality of life (Awad et al., 1998).

2.5.7 Multicomponent therapy

More recently, multicomponent or combination therapies have become increasingly popular for the management of OAB. The most common type of multicomponent therapies are oral drugs combined with behavioural or physical interventions (Burgio et al., 2008, 2010; Chancellor et al., 2008; Fitzgerald et al., 2008; Herschorn et al., 2004; Kafri et al., 2013; Kaya et al., 2011; Millard, 2004; Ozdedeli et al., 2010; Zimmern et al., 2010). In the emerging literature, the focus has turned towards combinations of drug therapies such as β_3-adrenoceptor agonists combined with antimuscarinic drug therapies (Abrams et al., 2015; Kosilov et al., 2015), or multicomponent antimuscarinic therapies (Kosilov et al., 2014).

2.6 Systematic reviews, meta-analysis and clinical guidelines

There is little evidence of the most clinically effective intervention for OAB, and consequently clinical practice is somewhat variable. Almost all published papers reporting data on interventions for OAB compare the intervention solely with a placebo or with interventions of the same class (i.e. oral drug therapy vs. oral drug therapy); and so there are a limited number of trials that compare interventions of different classes and different treatment modalities (e.g. oral drug therapy versus more invasive treatments), which makes comparison across the diverse range of interventions difficult. In a similar fashion, a number of published meta-analyses have evaluated interventions against placebo (Chapple et al., 2008a; Cui et al., 2013, 2014, 2015; Nitti et al., 2013; Sun et al., 2015; Reynolds et al., 2015), or more commonly against an intervention within the same class of therapy (Buser et al., 2012; Novara et al., 2008a; Luo et al., 2012; Sweeney et al., 2016; Cao et al.,

2016; Huang et al., 2015; Davis et al., 2015; Liu et al., 2014; Wu et al., 2014). One study by Freemantle et al. (2016) compared onaBoNT-A with mirabegron, but there have been no meta-analyses to evaluate comparative efficacy between diverse classes of interventions with different approaches, and consequently a superior intervention has not been identified. A cross-modality treatment comparison would be useful in the context of healthcare decision making, even in the presence of established treatment pathways, as it allows patients and physicians to make informed decisions with regard to the trade off between expected patient benefit, and increased levels of invasiveness.

The National Institute for Health and Care Excellence (2013b) in England and Wales currently recommend a 6-week course of BT and PFMT, together with a reduction in caffeine intake, as first line therapy for the conservative management of OAB. Electrical stimulation and/or biofeedback is recommended as first-line therapy for women who cannot actively contract pelvic floor muscles in order to aid motivation and adherence to treatment. If voiding frequency remains bothersome, NICE recommend that antimuscarinic drugs in combination with BT should be considered as second-line therapy. In situations where BT is ineffective, non-proprietary oxybutynin should be prescribed. Alternative antimuscarinic drug therapies such as darifenacin, solifenacin, tolterodine, and trospium are recommended as second-line therapies if oxybutynin can not be tolerated. Alternative second-line therapies such as Mirabegron - a β_3-adrenoceptor agonist, and percutaneous posterior tibial nerve stimulation are recommended if alternative conservative therapies are contraindicated, clinically ineffective, or present unacceptable side effects. Patients who are refractory to conservative measures are recommended minimally invasive interventions such as onaBonT-A, and SNS as third-line therapy. For the administration of onaBoNT-A, patients must be willing and able to self-catheterise. More invasive surgery such as augmentation cystoplasty and urinary diversion are recommended failing all other measures.

2.7 Implications for healthcare provision

Overactive bladder has a negative impact on all aspects of life leading to a significant burden not only to the sufferer but to society in general. With prevalence increasing with age and demographic trends indicating a larger elderly population, OAB is becoming a rapidly growing concern for healthcare provision worldwide. Cost estimates for OAB are in the order of €3.57 billion (Klotz et al., 2007) and $6 billion per annum, accounting for approximately 10% of total nursing home costs (Digesu et al., 2003). These costs are comparable worldwide for diabetes, dementia, and breast cancer (Hu and Wagner, 2000). It is therefore crucial to identify the most effective intervention for the management of OAB to minimise NHS time, expenditure and resource waste.

2.8 Chapter summary

Overactive bladder adversely affects a large proportion of the population. There is a plethora of interventions that are currently available for the management of OAB including lifestyle interventions, physical therapies, behavioural therapies, bladder training, oral drug therapies, and surgical management. In the current literature there is no coherent comparison between the diverse range of interventions and therefore identification of the most effective intervention is unknown. Overactive bladder has a detrimental impact on physical, mental and social well-being, costing the NHS approximately £500 million per year (Turner et al., 2004). Identification of the most effective intervention is essential to maximise patient benefit and ensure efficient use of NHS resources.

Chapter 3

Methodological Background

3.1 Chapter overview

This chapter introduces many of the statistical concepts and methodologies used and developed throughout this thesis. Fundamental statistical approaches are described and terminologies defined, highlighting where appropriate, key assumptions and limitations.

3.2 Bayesian methodology

A key characteristic of Bayesian methodology is the incorporation of full parameter uncertainty. In a frequentist framework, we make statements about the likelihood of the data at specific parameter values. A Bayesian approach assumes that every parameter in the model is unknown and should be estimated from a probability distribution; and thus, Bayesian methods differ to that of frequentist methods in that both the data and the model are assumed to be random elements (Spiegelhalter et al., 1999b). Bayesian methodology arose from the illustrious paper posthumously published by Thomas Bayes, where it states that:

> "*Given* the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named" (Bayes and Price, 1763, p.376).

In this extract, Bayes and Price (1763) highlight the use of probability statements as a means of expressing uncertainty in an unknown quantity, as well as the property of conditional probability i.e., what is the probability of failure in a single trial, conditional on data documenting the previous number of failures. To describe this, Bayes and Price (1763) proposed a theorem to relate conditional and marginal probabilities for observable events, which was subsequently termed Bayes' theorem.

Bayes' theorem can be used to make inferences about parameter estimates where instead of expressing parameters as observable event probabilities, the quantities of interest can be described as probability distributions (Lunn et al., 2012). Considering this from a health technology assessment perspective, assuming that θ is some unknown quantity of interest, such as a treatment effect, and y is the observed evidence, such as data from a clinical trial, Bayes' theorem can be used to calculate the probability distribution of θ conditional on the observed data y:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \tag{3.1}$$

where $p(\theta)$ denotes the prior distribution for θ, and is used to express beliefs about the uncertainty before taking in to account the observed data. The conditional probability, $p(\theta|y)$, is the posterior distribution of θ, and is used to express uncertainty about θ given the observed data. Similarly, the conditional probability, $p(y|\theta)$, denotes the dependence of the data, y, on values of the parameter θ. This is often expressed as a function of θ for fixed y, which is used as a basis for all standard likelihood-based models, and is termed the likelihood function. The normalising constant, $p(y)$, ensures that the posterior distribution integrates to 1. In most situations, it is not necessary to calculate the normalising constant in order

to assess properties of the posterior distribution (Lunn et al., 2012) and so, Bayes' theorem is often reduced to:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{3.2}$$

$$\text{i.e. Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Prior distributions can be thought of as expressing external information which could come from a variety of sources such as previous RCTs or meta-analyses (Spiegelhalter et al., 1999b). In the absence of external information, prior beliefs can be elicited from expert opinion; however, this approach can often result in much controversy (Spiegelhalter et al., 1999a). In order to avoid subjective prior beliefs about the parameter estimate, a 'non-informative' or 'vague' prior distribution - sometimes referred to as a 'flat' prior distribution - can be chosen to represent a feasible range of values that a parameter is expected to take, with large amounts of prior uncertainty (Carlin and Louis, 1997). The purpose of this approach is that the likelihood of the observed data is expected to drive posterior inference, and prior distributions are expected to have little impact in the overall parameter estimation (Spiegelhalter et al., 2000). This scenario is further illustrated through Figures 3.1 and 3.2. Figure 3.1 demonstrates the impact of informative prior distributions, selected from external evidence or expert opinion, on posterior estimation, whilst Figure 3.2 illustrates the impact of non-informative or vague prior distributions on posterior inference.

In most cases, estimation of the posterior distribution is of most interest, and these can be reported in the form of direct probability statements (Freedman, 1996). In this thesis, interest lies in assessing clinical effectiveness of interventions and as such posterior distributions for treatment effect estimates are summarised using the posterior median. The posterior median is of most use in situations where posterior distributions are skewed but should be similar to the posterior mean in situations where distributions are symmetrical. Typically, posterior distributions are presented with Bayesian 95% credible intervals (CrI). Credible intervals can be

FIGURE 3.1: Impact of an informative prior on the posterior distribution

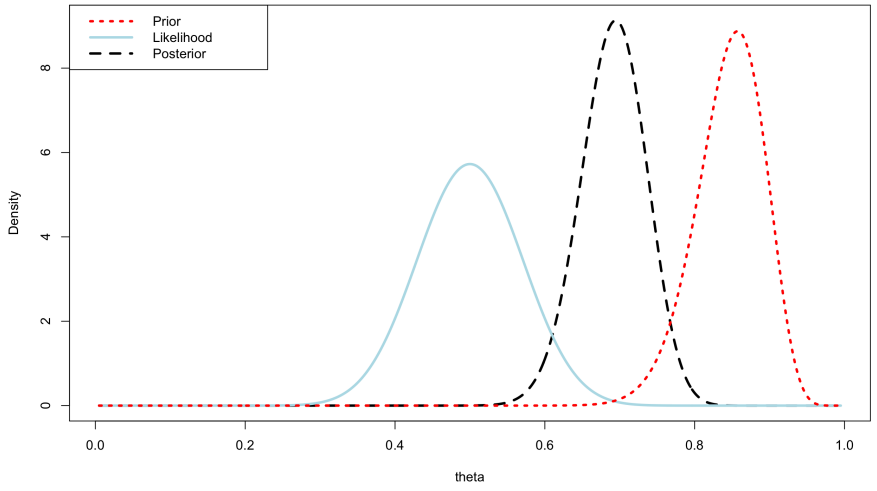
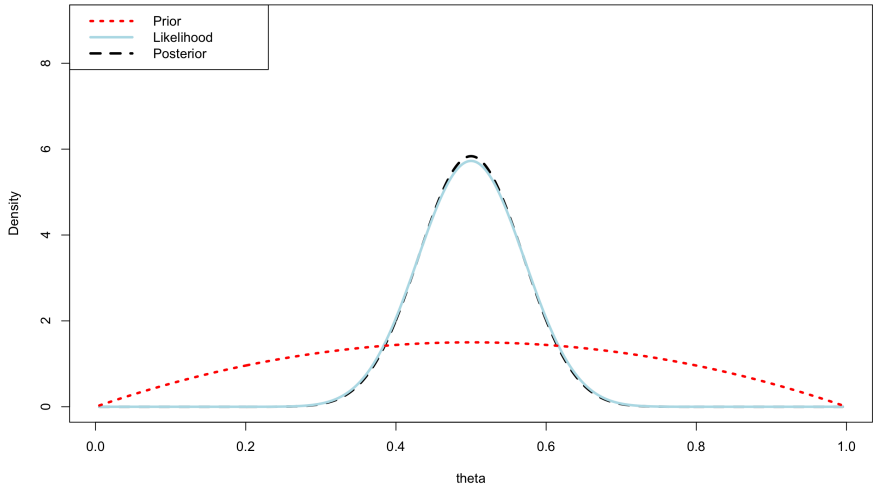


FIGURE 3.2: Impact of a non-informative prior on the posterior distribution



thought of as a Bayesian equivalent to 95% confidence intervals (CI). However, the width of credible intervals are determined by the standard deviation of the posterior distribution, whereas the width of a confidence interval is determined by the standard error of the estimate. Therefore, the interpretation of credible intervals are slightly different to that of confidence intervals - the correct interpretation of a credible interval is that there is a 95% possibility that the true parameter

value is contained within the 95% credible interval. A confidence interval can be interpreted as, in a series of confidence intervals constructed from replicated experiments, 95% of intervals will contain the true parameter value. In situations where posterior distributions are skewed, highest posterior density (HPD) intervals - the smallest interval that has 95% posterior probability - are sensible alternatives to credible intervals.

There are many advantages of using Bayesian methodology, including the ability to obtain direct probability statements which can be useful in a decision making context. For example, interest may lie in the probability that a treatment effect of a new intervention exceeds that of the treatment effect of standard practice (Sutton and Abrams, 2001). Furthermore, Bayesian methodology allows the flexibility to model complex clinical scenarios. Incorporation of full parameter uncertainty in both the model and the data can be exploited to permit computation of parameters which would otherwise be difficult to estimate (Gilks, 2005). A Bayesian approach is also efficient in integrating all relevant information in the estimation procedure, which can be considered as a more ethical approach than that of the classical alternative, as it has the potential to make full use of the experience of past patients (Welton et al., 2012). In addition, a Bayesian framework is useful in providing predictive posterior distributions which can be used to inform the future design of trials or to assess expected clinical value in alternative patient populations (Spiegelhalter et al., 2004).

Disadvantages of adopting a Bayesian approach include the incorporation of prior beliefs which can potentially expose the estimates to an element of subjectivity. The magnitude of which can vary depending on the choice of prior distribution selected, though as the amount of data increases, the impact of any prior distribution will diminish (Spiegelhalter et al., 2000). In an article by Lambert et al. (2005) it has been shown that the choice of 'vague' prior distribution, particularly for variance parameters, can have a notable impact on the overall estimates.

Another disadvantage of using a Bayesian approach is the absence of easily interpretable, and a universal measure of statistical significance, such as p-values, though posterior probability statements can be made. Historically, there was an element of concern about the adoption of Bayesian methodology for fear of acceptance by regulatory bodies, and academic journals. However, this is beginning to change with the uptake of, and encouragement to use, Bayesian methodology by regulatory bodies such as NICE (Dias et al., 2011a).

3.2.1 Markov Chain Monte Chain (MCMC) and Gibbs Sampling

Calculation of marginal posterior distributions require high-dimensional integrals, and thus complex computational approaches are required (van Haasteren, 2014). Markov Chain Monte Carlo (MCMC) methods have developed as an efficient way of sampling from the posterior distribution (Hastings, 1970; Geman and Geman, 1984; Tanner and Wong, 1987; Gelfand and Smith, 1990). MCMC simulation repeatedly draws a value for each parameter such that, in time, the samples are drawn from the posterior distribution (Gelfand and Smith, 1990). To begin the simulation process, an initial guess of the parameter value is chosen and subsequent values are sampled based on the values of the previous iteration. The idea that current values depend on past iterations directly through the previous value is referred to as a 'Markov Chain' property. There are a number of methods with which to implement the Markov Chain property, and one of the simplest tools is the Gibbs Sampler (Gelfand and Smith, 1990). Gibbs sampling sequentially takes each parameter in the model and draws a sample from its posterior distribution, conditional on all other parameters being fixed at their current value. Therefore, for iteration x, the first parameter (θ_1) in the sequence of n parameters is defined as $f(\theta_1(x+1)|\theta_2(x), ..., \theta_n(x))$. This process is repeated for n parameters. These distributions are referred to as full conditional distributions, and tend to be easier

to sample from than that of joint distributions as they possess univariate properties (Welton et al., 2012).

Samples obtained *ab initio* may not accurately reflect the distribution of interest, however, it can be shown that with sufficient simulations the values will converge to the joint posterior distribution (Welton et al., 2012). Simulations prior to model convergence are considered a 'burn-in' phase, and are discarded from the analysis. Following model convergence, a large number of further samples are collected from the joint posterior distribution. With large enough samples, the summary measures obtained from these estimates should provide an accurate indication of the summary measures for the posterior marginal distributions (Gelman et al., 2011). Obtaining parameter estimates in this way is referred to as Monte Carlo simulation.

There are several software packages that are available for performing MCMC simulation. One of the most commonly used and freely available software packages is WinBUGS (Spiegelhalter et al., 2003). Alternative software programmes include OpenBUGS (Lunn et al., 2009), JAGS (Plummer, 2003), and newly emerging programmes such as Stan (Hoffman and Gelman, 2014; Carpenter et al., 2016). Throughout this thesis, the WinBUGS platform was used to perform all MCMC simulations as it is commonly used to perform Bayesian analyses in a HTA setting (Dias et al., 2011b; Sutton et al., 2008), and provides a stable environment with which to perform MCMC simulations (Lunn et al., 2012).

3.2.2 Convergence diagnostics

It is imperative to assess convergence of individual parameters to the posterior distribution, and this is often assessed informally by inspection of several diagnostics plots. In realistically complex models, there may be a large number of parameters

with which to assess convergence and it may not be feasible to evaluate all parameters. In this situation, parameters of interest together with a random selection of the remaining parameters are assessed (Lunn et al., 2012). For most models, the parameters of interest include basic parameters, such as log-odds ratios, and not those of functional parameters, such as odds ratios. Convergence diagnostic plots are founded on a null hypothesis of convergence and thus, by design, are only used to detect non-convergence. Different stochastic properties of Markov chains are examined through different diagnostic plots, and so it is vital to evaluate a range of methods (Lunn et al., 2012):

3.2.2.1 Brooks-Gelman-Rubin plots

One commonly used approach implemented in the WinBUGS software is the Brooks-Gelman-Rubin plot, which was originally proposed by Gelman and Rubin (1992) and soon after developed by Brooks and Gelman (1998). This plot is made up of 3 coloured lines: the green line (G) is a measure of between-chain variability and represents the width of an 80% credible interval from simulations pooled from all chains, the blue line (B) is a measure of within-chain variability and represents the average 80% credible interval of each chain separately, and the red line (R) is the ratio of both the between- and within-chain variability. Convergence is said to be achieved when R reaches 1, and B and G converge together to reach stable estimates (Lunn et al., 2012). However, Brooks-Gelman-Rubin plots can not be used to assess parameters with a binary property, for example, parameters used to assess the probability that a treatment effect surpasses that of another treatment effect.

3.2.2.2 Autocorrelation plots

Due to the properties of Markov Chain samplers, the MCMC chains will be inherently correlated. The degree of correlation between sampled values within chains can be assessed through autocorrelation plots. For a specified number of iterations

apart, also known as the lag, the autocorrelation plot illustrates the level of correlation between sampled values. With increasing lag the autocorrelation plot is expected to gradually diminish, if it does not do so, this can indicate slow mixing of the chains or non-convergence. In order to address unacceptable autocorrelation, models can be run for a longer sampling phase, the chains could be sampled with increased lag, or the model could be reparameterised (Lunn et al., 2012).

3.2.2.3 History/trace plots

History and trace plots can be used for one or more chains to assess convergence of the samples. To indicate convergence, history and trace plots should take the appearance of some random noise about a stable mean. If there is a notable structure in the plot, this could indicate slow convergence (Spiegelhalter et al., 2004). For two or more chains, history and trace plots can be useful to detect whether chains are very different.

3.2.2.4 Density plots

Density plots can be useful to assess whether or not parameters have taken the expected distributional form. For example, in time we would expect the posterior distribution for normally distributed parameters to have a characteristically bell-shaped appearance which is evenly distributed about some mean. A skewed density plot would suggest the need to use highest posterior density (HPD) credible intervals (Chen and Shao, 1999).

3.2.3 Goodness of fit and model selection

The posterior mean residual deviance is defined as the deviance for the fitted model minus the deviance of the saturated model, and is considered a measure of model fit (Spiegelhalter et al., 2002). A saturated model is considered to have a perfect fit to the data because it has as many parameters as there are values

to be fitted. If model predictions from the fitted model accurately fit the data, we would expect the posterior mean residual deviance to approximately equal the number of unconstrained data points (Spiegelhalter et al., 2002). Throughout this thesis, the posterior mean residual deviance was used to assess model fit. Though more complex models may improve model fit, in doing so, they may reduce interpretability; and so a balance between fit and complexity is sought. For model comparison and selection, the Deviance Information Criterion (DIC) was used. The DIC is an extension of Akaike's Information Criterion (AIC) and is thought of as a Bayesian measure of model complexity and fit (Spiegelhalter et al., 2002). The DIC is equivalent to the posterior mean deviance penalised by the effective number of parameters estimated in the model. The smaller the DIC, the better the model fit. It has been suggested that differences in DIC of 5 or more are considered important (Spiegelhalter et al., 2002).

3.2.4 Sampling and prior distributions

In a Bayesian analysis, sampling distributions form the foundation for likelihoods, and therefore, as in classical analyses, normal, binomial, and Poisson models, among others, are usual choices to reflect different types of data (e.g. continuous, count, event data etc.). As well as specification of sampling distributions for observable quantities, a Bayesian analysis requires specification of prior distributions to reflect the level of uncertainty for unknown quantities. Specification of prior distributions are more difficult than that of sampling distributions, and consideration of the distributional form is required to ensure that the range of possible values a parameter could take is within a plausible interval (Spiegelhalter et al., 2004). Throughout this thesis, a number of probability distributions with different distributional properties were used to specify both sampling and prior distributions. In this section, the selected distributions are discussed together with their common uses and distributional properties.

A) Discrete univariate distributions

3.2.4.1 Binomial distribution

A discrete binomial distribution is used to specify the total number of successes, Y, out of n independent Bernoulli trials with probability θ. When $n = 1$, the Binomial distribution is simplified to the Bernoulli distribution, denoted by $Y \sim \text{Bern}[\theta]$. The binomial distribution is denoted by $Y \sim \text{Bin}[n, \theta]$ with properties (Johnson et al., 2005):

$$p(y|n,\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}; \qquad y = 0, 1, ..., n \tag{3.3}$$
$$E(Y|n,\theta) = n\theta$$
$$Var(Y|n,\theta) = n\theta(1-\theta)$$

A binomial distribution is often used as a sampling distribution for the number of successes in a sample. In this thesis a binomial distribution was used in Chapters 6 and 7 as sampling distributions with which to model safety and tolerability data. Bernoulli distributions were used to impose ordering constraints on parameters in Chapters 4, 7 and 8.

B) Continuous univariate distributions with unrestricted range

3.2.4.2 Normal distribution

A normal distribution is used to specify a random quantity, Y, with mean, θ, and variance, σ^2. The normal distribution is denoted by $Y \sim \text{N}[\theta, \sigma^2]$ and has the properties (Johnson et al., 1994):

$$p(y|\theta,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(y-\theta)^2}{\sigma^2}\right); \qquad -\infty < y < \infty \tag{3.4}$$
$$E(Y) = \theta$$
$$Var(Y) = \sigma^2$$

Normal distributions are universally used as sampling and prior distributions for continuously distributed data due to its favourable properties and familiarity (Lunn et al., 2012). A normal distribution is used to specifiy a continuous likelihood for efficacy data in Chapters 6 and 7 of this thesis. It is also used generally for specifying prior distributions for treatment effects and regression coefficients.

C) Continuous univariate distributions restricted to be positive

3.2.4.3 Half-normal distribution

The half-normal distribution originates from wrapping the normal distribution about 0: if $X \sim \mathrm{N}[\theta, \sigma^2]$, then $|X| \sim \mathrm{HN}[\sigma^2]$. Thus a half-normal distribution can be denoted as $Y \sim \mathrm{HN}[\sigma^2]$ with distributional properties (Johnson et al., 1994):

$$
\begin{aligned}
&p(y|\sigma^2) = \sqrt{\frac{2}{\pi\sigma^2}} e^{\frac{-y^2}{2\sigma^2}}; \qquad y \in (0, \infty) \\
&E(Y|\sigma^2) = \sqrt{\frac{2}{\pi}}\sigma \\
&Var(Y|\sigma^2) = \sigma^2\left(1 - \frac{2}{\pi}\right)
\end{aligned}
\tag{3.5}
$$

Half-normal distributions are used throughout this thesis as an alternative prior distribution for functions of variance parameters such as standard deviations.

3.2.4.4 Exponential distribution

The exponential distribution is used to specify the likelihood for positive continuous data. A common use of the exponential model is to define time-to-event data. It is denoted by $Y \sim \mathrm{Exp}(\theta)$ with distributional properties (Johnson et al., 1994):

$$
\begin{aligned}
&p(y|\theta) = \theta e^{-\theta y} \text{ for } y > 0, \ \ \theta > 0 \\
&E(Y|\theta) = \frac{1}{\theta} \\
&Var(Y|\theta) = \frac{1}{\theta^2}
\end{aligned}
\tag{3.6}
$$

An exponential distribution is used in Chapter 4 of this thesis to specify the likelihood for survival analysis data.

3.2.4.5 Gamma distribution

Gamma distributions are also used for parameters that are constrained to positive values such as time-to-event data. A common use of a Gamma distribution is to define a prior for precision parameters. A Gamma distribution is defined by $Y \sim \text{Gamma}[a, b]$ with distributional properties (Johnson et al., 1994):

$$\begin{aligned} p(y|a,b) &= \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by}; \qquad y \in (0, \infty) \\ E(Y|a,b) &= \frac{a}{b} \\ Var(Y|a,b) &= \frac{a}{b^2} \end{aligned} \tag{3.7}$$

In this thesis Gamma distributions were used as prior distributions on the shape parameter of the Weibull model described in Section 3.2.4.6 below, and for variance parameters on the precision scale.

3.2.4.6 Weibull distribution

The Weibull distribution is an extension of the exponential distribution and is denoted by $y \sim \text{Weibull}(\gamma, \mu)$ with distributional properties (Johnson et al., 1994):

$$\begin{aligned} p(y|\gamma,\mu) &= \begin{cases} \mu\gamma y^{\gamma-1} \exp(-\mu y^\gamma) & x \geqslant 0 \\ 0 & x < 0 \end{cases} \\ E(Y|\gamma,\mu) &= \mu^{-1/\gamma} \Gamma\left(1 + \frac{1}{\gamma}\right) \\ Var(Y|\gamma,\mu) &= \mu^{-2/\gamma} \left[\Gamma\left(1 + \frac{2}{\gamma}\right) - \Gamma\left(1 + \frac{1}{\gamma}\right)^2\right] \end{aligned} \tag{3.8}$$

In this thesis, Weibull distributions were used to sample time-to-event data, and were used to form the basis of the models developed and applied in Chapter 4.

3.2.4.7 Log-normal distribution

The log-normal distribution is the log of a normally distributed random variable with mean, μ, and variance, σ^2. A log-normal distribution is often used to define time-to-event data and is denoted by $y \sim \text{log-normal}(\mu, \sigma^2)$ with distributional properties (Johnson et al., 1994):

$$\begin{aligned} p(y|\mu,\sigma^2) &= \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}} \\ E(Y|\mu,\sigma^2) &= e^{\mu+\frac{\sigma^2}{2}} \\ Var(Y|\mu,\sigma^2) &= (e^{2\mu+\sigma^2})(e^{\sigma^2}-1) \end{aligned} \tag{3.9}$$

Log-normal distributions were used in Chapter 4 as an alternative sampling distribution to model survival analysis data.

3.2.4.8 Generalised gamma distribution

If $X \sim \text{Gamma}(a, 1)$ then $Y = X^{1/c}/b$ has the generalised gamma distribution (Johnson et al., 1994) and denoted by $y \sim \text{ggamma}(a, b, c)$ with properties:

$$\begin{aligned} p(y|a,b,c) &= \frac{cb^{ca}y^{ca-1}e^{-(by)^c}}{\Gamma a} \text{ for } y>0,\ \ a,b,c>0 \\ E(Y|a,b,c) &= \mu = \frac{\Gamma\left(a+1/c\right)}{b\Gamma(a)} \\ Var(Y|a,b,c) &= \frac{\Gamma(a+2/c)}{b^2\Gamma(a)} - \mu^2 \end{aligned} \tag{3.10}$$

Generalised gamma distributions were used as sampling distributions in Chapter 4 of this thesis to model a RCT with time-to-event data.

D) Continuous univariate distributions restricted to a finite interval

3.2.4.9 Uniform distribution

A uniform distribution is used to define parameters that can only take values between a specific interval (a, b). It is commonly used to define variance parameters on the standard deviation scale, and proportional data such that the selected intervals are bound between (0,1). A uniform distribution is denoted by $Y \sim \mathrm{U}[a, b]$ with properties (Johnston et al., 1994):

$$\begin{aligned} p(y|a,b) &= \frac{1}{b-a}; \qquad y \in [a,b],\ b > a \\ E(Y|a,b) &= \frac{a+b}{2} \\ Var(Y|a,b) &= \frac{(b-a)^2}{12} \end{aligned} \tag{3.11}$$

Throughout this thesis, uniform distributions were used to specify prior distributions for variance parameters.

E) Continuous multivariate distributions

3.2.4.10 Multivariate normal distribution

Multivariate normal distributions are commonly used to model multiple, correlated outcome measures, or parameters such as random effects, which independently satisfy the normality assumption. A multivariate normal distribution for k-dimensional random vector $\mathbf{y} = [Y_1, ..., Y_k]$ is denoted by $\mathbf{y} \sim \mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with probability density function (Kotz et al., 2004),

$$p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{k}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}-\boldsymbol{\mu})} \tag{3.12}$$

mean vector,

$$\boldsymbol{\mu} = [E[Y_1], E[Y_2], ..., E[Y_k]]$$

and covariance matrix for $k \times k$ dimensions,

$$\boldsymbol{\Sigma} = [\text{Cov}[Y_i, Y_j]], \text{ for } i = 1, ..., k; \; j = 1, ..., k.$$

In this thesis, a multivariate normal distribution was used to specify sampling distributions for multiple outcome network meta-analyses described in Chapter 8.

3.3 Health technology assessment (HTA)

Health technology assessment (HTA), as a broad concept, aims to compare newly emerging and existing healthcare interventions with the intention of providing evidence as to what is the most clinically and cost effective intervention in order to enable the most efficient use of health resources at a population level. Interventions are usually compared in terms of clinical effectiveness, cost, and suitably to a patient population (Drummond et al., 2008). Governing bodies such as the National Institute for Health and Care Excellence (NICE) in the UK use HTA as a basis for developing treatment recommendations for the NHS (Drummond and Sorenson, 2009). Clinical value is at the heart of the HTA process as willingness to pay, and suitably to a given population are determined by the magnitude of clinical impact. This information is often obtained through the practice of evidence-based medicine. Evidence based medicine can be described as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett et al., 1996, p.71). In order to identify the best evidence for decision making, all relevant information needs to be collated.

3.4 Evidence-synthesis

In a HTA setting, systematic reviews and meta-analyses are one of the most commonly adopted forms of evidence synthesis. Systematic reviews are used to collate the totality of evidence on clinical effectiveness, safety, tolerability, and often

economic impact of healthcare interventions; and thus, systematic reviews form the foundation of evidence based medicine. Following specification of the scope of a systematic review, a predefined selection criteria is drawn up to identify a complete set of relevant studies using electronic databases and exploitation of references. Aside from collecting outcome data, it is important to collate information regarding exposure measures, and potential covariates. Empirical evidence identified from systematic literature reviews are extracted in a blinded, and reproducible manner. This data are synthesised and interpreted using meta-analyses.

Meta-analyses have previously been described as "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the finding" (Glass, 1976, p.3). A meta-analysis aims to quantify an average effect size and its relative uncertainty, from a combination of multiple trial results. There are several methods for performing meta-analyses including vote counting, synthesis of p-values, pooling individual participant data (IPD), and a combination of effect estimates. For data synthesis using a combination of effect estimates, two meta-analysis models are most commonly used: fixed and random effects models.

3.4.1 Fixed effect (FE) model

A fixed effect meta-analysis assumes no heterogeneity between studies, that is, the effect estimates from N studies are estimating the same underlying true effect, d, and effect estimates differ solely because of random sampling. Thus for the i^{th} study, the observed effect, Y_i, is assumed to be normally distributed about the true underlying value, d, with variance equal to the estimated within-study variance, se_i^2, assumed known (Sutton et al., 2000):

$$Y_i \sim \text{Normal}\left(d, se_i^2\right) \text{ for } i = 1, ..., N \tag{3.13}$$

If evaluated within a Bayesian framework, a prior distribution for the common effect size, d, is required. In the absence of external information, vague prior

knowledge may be specified using a Normal distribution. For comparative outcomes the normal distribution will be centred at no effect with a large variance relative to the scale of the outcome of interest. For example, if the outcome in question is change from baseline in the number of symptom episodes, the prior distribution

$$d \sim \text{Normal}(0, 10^3) \tag{3.14}$$

suggests that a priori the investigator would be 95% certain that the true value of d lies between $(0 - 1.96 \times 31.6)$ and $(0 + 1.96 \times 31.6)$. That is equivalent to an interval of (-61.9, 61.9) which more than spans the plausible range of values that would be expected for the change from baseline in the number of symptom episodes. However, care must be taken in specifying prior distributions for d, this is especially true when dealing with binary outcomes that are specified on the log-odds ratio scale (Welton et al., 2012). A normal prior distribution described in Equation (3.14) can be applied to the common effect d on the log-odds scale, with the assumption that log-transformed data are approximately normally distributed. On the odds-ratio (OR) scale, this is equivalent to specifying a prior interval in the range of $(1.31 \times 10^{-27}, 5.29 \times 10^{13})$, which although covers the range of all plausible values, may be too large for sensible and clinically interpretable posterior estimates. In this situation a smaller, yet still, vague prior distribution should be considered on the log-odds ratio scale.

In many HTA settings, the assumption that all studies are estimating a common truth may not be satisfied due to differences in trial design, patient population, conduct and context. A random effects model relaxes the assumption of a common truth and allows for between-study variability.

3.4.2 Random effects (RE) model

Random effects (RE) meta-analysis assumes that the true underlying effect is different between studies, and that the study-specific truth is subject to random

sampling. In order to capture this additional variability, a further level is incorporated in to the model:

$$\begin{aligned} Y_i &\sim \text{Normal}\left(\delta_i, se_i^2\right) \text{ for } i = 1, ..., N \\ \delta_i &\sim \text{Normal}\left(d, \tau^2\right) \text{ for } i = 1, ..., N \end{aligned} \tag{3.15}$$

where Y_i and se_i^2 have the same interpretation as described in Section 3.4.1, and δ_i denotes the study-specific true effect with true overall effect d and between study variance τ^2. In situations where there is no heterogeneity between studies i.e. $\tau^2 = 0$, random effects models reduce to fixed effect models. In a Bayesian framework, prior distributions are required for the overall pooled effect, d, and between-study variance, τ^2. As described in Section 3.4.1, a normal distribution centred at mean 0 with large variability (Equation (3.14)) is an appropriate choice of vague prior distribution for the overall pooled effect, d. For the between-study variance, τ^2, a restricted distribution is required to ensure that only positive values are sampled for variance parameters, such as a Uniform distribution (Spiegelhalter et al., 2004):

$$\tau^2 \sim \text{Uniform}(0, 5) \tag{3.16}$$

As previously mentioned in Section 3.2, choice of vague prior distributions for variance parameters can have a notable impact on the overall effect estimates. Consequently, different choices of vague prior distributions should be investigated through a series of sensitivity analyses (Lambert et al., 2005).

3.4.3 Heterogeneity

Whilst random effects models begin to account for differences between studies, it does not completely account for all sources of variation. Beyond that of chance, common sources of variation in treatment effects include differences in patient populations, administration of interventions, trial design and conduct, and changes in medical practice (i.e. differences in calendar year). Systematic differences in

treatment effects between studies, which is beyond that attributable to random sampling error alone, is termed statistical heterogeneity (Sutton et al., 2000). Heterogeneity between studies can be problematic in meta-analyses and sources of variation should be explored and accounted for in the analysis (Spiegelhalter et al., 2004). This can be done in one of two ways: the first is to undertake separate sub-group meta-analyses, for discrete characteristics, stratified by the covariate of interest; and the second involves incorporation of the covariate in to the analysis using meta-regression techniques. An advantage of the latter is that the differences between study-level characteristics can be quantified with associated uncertainty (Welton et al., 2012) and study-level characteristics, such as average participant age, can be incorporated as a continuous measure. Exploring sources of statistical heterogeneity can be useful in identifying reasons for variation in treatment effects, and as a result, it can help to tailor interventions for decision making.

3.4.3.1 Meta-regression

Meta-regression models can be used to relate the size of treatment effect estimates to measurable study characteristics using regression techniques. Meta-regression methods can be applied to both fixed and random effects meta-analysis, but random effect models are often preferred as they account for any residual heterogeneity that can not be explained by the covariate of interest (Whitehead, 2002). Extending Equation (3.15), and using the same notation previously described, the general model for a random effects meta-regression is given by:

$$\begin{aligned} Y_i &\sim \text{Normal}\left(\delta_i + \beta(x_i - \bar{x}), se_i^2\right) \text{ for } i = 1, ..., N \\ \delta_i &\sim \text{Normal}\left(d, \tau^2\right) \text{ for } i = 1, ..., N \end{aligned} \tag{3.17}$$

where β is the regression coefficient, x_i is the value of the covariate of interest for the i^{th} study, and $\bar{x}$ is the mean of the covariate across studies. By centering the covariate about the mean, d is interpreted as the mean treatment effect for the average study/aggregated patient characteristic. For example, if the covariate of interest was participant age, d could be interpreted as the mean treatment effect

for a patient of average age in the included studies. As the coefficient β is unconstrained, a normal distribution with large variability would be an appropriate choice of prior distribution.

As with all observational relationships, the meta-regression models may be subject to unknown confounding factors, as well as aggregation bias if the aggregated study-level data do not accurately reflect the relationship at the individual participant level. Thus, characteristics included in a meta-regression analysis should be considered to have an associative relationship rather than causative relationship on the overall treatment effect estimates (Sutton et al., 2000). Furthermore, meta-regression models have low power to detect moderator effects when there are few studies included in meta-regression analyses (Hempel et al., 2013), or there is little variation in covariate effects across studies (Debray et al., 2015). This approach further lacks power when using aggregated patient level covariates (Riley et al., 2010).

3.4.3.2 Adjusting for baseline risk

In many HTA settings, it is often desirable to adjust analyses to account for baseline risk in patient cohorts in order to tailor interventions to populations who will benefit most (Dias et al., 2012). Meta-regression techniques would seem a natural way to incorporate baseline risk in random effects models. Defining baseline risk as the average risk of an untreated patient i.e. the risk of an outcome for a patient under a control or placebo condition, then covariate values are obtained based on the outcome of patients in the control arm. This approach leads to structural dependence in the regression model as information from the control arm would contribute to both the outcome and covariate values. In addition, a meta-regression approach ignores uncertainty in the covariate effect, which together can lead to bias in the regression slope towards the null, or underestimation of the covariate effect, a phenomenon known as attenuation or regression dilution bias (Hutcheon et al., 2010). Thompson et al. (1997) proposed a model to appropriately include

baseline risk by incorporating the study-specific baseline means, μ_i, as a covariate in the meta-regression model. Using the notation described above, the baseline risk model is given by:

$$\begin{aligned} Y_i &\sim \text{Normal}\left(\delta_i + \beta(\mu_i - \bar{\mu}), se_i^2\right) \text{ for } i = 1, ..., N \\ \delta_i &\sim \text{Normal}\left(d, \tau^2\right) \text{ for } i = 1, ..., N \end{aligned} \tag{3.18}$$

where $\bar{\mu}$ denotes the mean baseline risk across all studies and is used to centre the covariate to the average baseline risk of the patient population.

3.4.4 Pairwise meta-analysis

In standard pairwise meta-analyses, two treatments will be compared on a head-to-head basis (Higgins et al., 2008); and thus the comparison of interest, such as a new healthcare intervention versus placebo, is common throughout the trials included. Pairwise meta-analyses are frequently used to provide a more precise estimate of a healthcare intervention relative to another (Whitehead, 2002). However, conducting a series of pairwise meta-analyses will not answer the research question often provoked from a HTA perspective - which intervention is most effective overall? Rather, it will identify a series of superior treatments in each pairwise analysis (Sutton et al., 2000).

3.4.5 Indirect treatment comparisons and network meta-analysis (NMA)

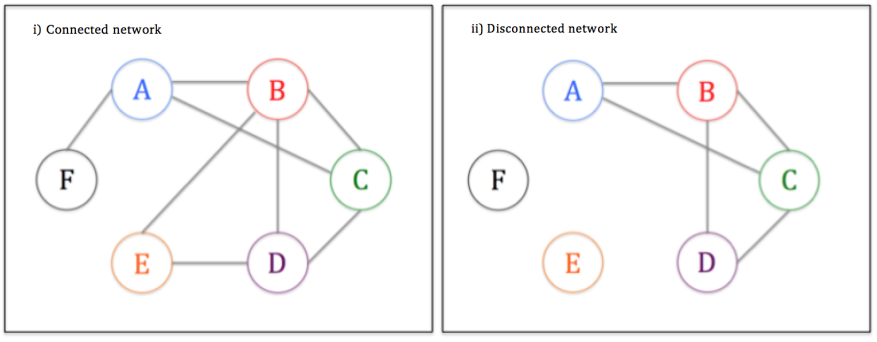
In situations where more than two treatments are of interest, a mixed treatment comparison (Lu and Ades, 2004; Caldwell et al., 2005) or network meta-analysis (NMA) (Lumley, 2002) is required. Network meta-analysis is an extension of meta-analysis and combines data from several clinical trials in similar patient populations with the aim to evaluate multiple interventions that may not have

been directly compared otherwise. This approach combines both direct information (obtained from head-to-head trials) and indirect information (obtained from trials that share a common comparator) to obtain relative treatment effects for all interventions whilst maintaining randomisation. Combining direct and indirect information in this way assumes an additive relationship between treatment effects. Meaning that for treatments A, B, and C, suppose that our primary interest is in comparing treatments C versus B. Trials evaluating B and C directly will provide information of a direct estimate of C versus B, d^{dir}_{BC}, and the remaining trials, namely trials directly comparing A and C, d^{dir}_{AC}, and A and B, d^{dir}_{AB}, are used to provide indirect information to estimate C versus B, d^{ind}_{BC} using the additivity assumption:

$$d^{ind}_{BC} = d^{dir}_{AC} - d^{dir}_{AB}. \tag{3.19}$$

For the additivity assumption to hold, treatments must form a connected network as shown in Figure 3.3i, and adhere to the consistency assumption (White et al., 2012) - that is, trials estimating the treatment effect in A versus B trials are the same (under the fixed effect assumption), or exchangeable (under the random effects assumption) with trials comparing B and C, and A and C, had they included treatments A and B, respectively. This is equivalent to saying that all studies are multi-arm studies including all treatments, but that some treatment effects are missing at random (MAR) i.e. their missingness is not related to the results.

FIGURE 3.3: Example of connected and disconnected networks



Nodes A-F represent treatments and interconnecting lines represent direct treatment comparisons

The general random effects NMA framework for a continuous outcome is described by Dias et al. (2011b), and expressed here in terms of specific interventions in order to easily extend this framework in subsequent chapters. For an intervention $j = 1, 2, ..., n_t$, in study i, a continuous outcome such as mean change from baseline in symptom episodes, y_{ij}, is assumed to follow a normal distribution with mean equal to the underlying intervention effect, θ_{ij}, and observed standard error, se_{ij}. Let μ_i represent the mean change from baseline for a reference intervention, t_{ib}, (often a placebo or control intervention) in arm 1 of the i^{th} study. Suppose that $\delta_{i,bk}$ represents the mean difference of the treatment in the k^{th} arm, t_{ik}, relative to study-specific reference intervention in arm 1 (the control arm), t_{ib}. For 2-arm trials, $\delta_{i,bk}$ is drawn from a normal distribution with mean equal to the relative effect of the treatment in arm k compared to the treatment in arm 1 of study i, $d_{t_{ib}t_{ik}}$, with between-study variance, τ^2. The overall model is based on a linear regression model on a natural additive scale and is given by:

$$
\begin{aligned}
&y_{ij} \sim \text{Normal}(\theta_{ij}, se_{ij}^2) \\
&\theta_{ij} = \mu_i + \delta_{i,bk} I_{\{j=k\}} \\
&I_{\{u\}} = \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \\
&\delta_{i,bk} \sim \text{Normal}(d_{t_{ib}t_{ik}}, \tau^2)
\end{aligned}
\tag{3.20}
$$

For example, for trials that compare interventions A and B, $d_{t_{ib}t_{ik}} = d_{AB}$. The pooled treatment effect of treatment A relative to treatment B is given by:

$$
d_{AB} = d_{1B} - d_{1A} \tag{3.21}
$$

The intervention effect of the reference treatment for the entire treatment network, $j = 1$, usually a placebo or control intervention, is set to 0 such that $d_{11} = 0$. Minimally informative prior distributions are placed on the relative effects of basic parameters, $d_{1j} \sim \text{Normal}(0, 10^3)$ for $j = 2, ..., n_t$, with d_{AB} (where $A > 1$ and $B > 1$) referred to as functional parameters, and expressed in terms of the basic

parameters described in Equation (3.21).

For trials of more than 2-arms, the correlation between study-specific treatment comparisons, $\delta_{i,bk}$, must be taken in to account. These trials will estimate a vector of random effects $\boldsymbol{\delta_i}$. The relative effect of the study-specific reference treatment in arm 1 (the control arm) relative to itself, $\delta_{i,b1}$, is set to 0. Assuming that the relative effects all have the same between-trial variance, $\boldsymbol{\delta_i}$ is given by (Dias et al., 2011b):

$$\boldsymbol{\delta_i} = \begin{pmatrix} \delta_{i,b2} \\ \vdots \\ \delta_{i,bna_i} \end{pmatrix} \sim \text{MVN}_{na_i-1} \left(\begin{pmatrix} d_{t_{ib}t_{i2}} \\ \vdots \\ d_{t_{ib}t_{ina_i}} \end{pmatrix}, \begin{pmatrix} \tau^2 & \tau^2/2 & \dots & \tau^2/2 \\ \vdots & \vdots & \ddots & \vdots \\ \tau^2/2 & \tau^2/2 & \dots & \tau^2 \end{pmatrix} \right) \tag{3.22}$$

where na_i denotes the number of arms in study i, and $d_{t_{ib}t_{ik}} = d_{1,t_{ik}} - d_{1,t_{ib}}$ as described in Equation (3.21). Equation (3.22) can be expressed in terms of an appropriate set of conditional univariate distributions. Thus, for the random effects for arm k ($k > 2$), conditional upon those for all other arms from 2 to $k-1$, the conditional distribution is given by (Dias et al., 2011b):

$$\delta_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{ik}}, \tau^2)$$

and for $k = 3 \dots na_i$, the k^{th} conditional distribution is defined by:

$$\delta_{i,bk} | \begin{pmatrix} \delta_{i,b2} \\ \vdots \\ \delta_{i,b(k-1)} \end{pmatrix} \sim \text{N} \left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\delta_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right) \tag{3.23}$$

To appropriately account for the between-arm correlations, either the multivariate distribution described in Equation (3.22) or the conditional distributions described in Equation (3.23) can be used. In this thesis, the conditional distributions were used.

Notably, when the between-study variance is zero, that is, $\tau^2 = 0$, the model reverts to a fixed effect model. The reference intervention means, μ_i, and between-study variance, τ^2, require specification of prior distributions. In the absence of external information a normal distribution with large variance (e.g. Equation (3.14)) is appropriate for the reference intervention means, μ_i, as $\mu_i \in (-\infty, \infty)$, and a uniform distribution spanning the range of plausible values for the between-study standard deviation, τ, can be specified. Typically a uniform (0, 5) prior distribution is selected for continuous outcomes, suggesting that the between-study standard deviation can take any value between, but not including, 0 and 5, and small values of τ are equally likely as large values (Spiegelhalter et al., 2004). A value of 5, for example, would indicate that for a random pair of studies, the difference in the mean change from baseline could be as large as 5.5. Choice of prior distribution for the between-study standard deviation, τ, can have a notable impact on the results, this is especially true for NMAs with few studies (Higgins et al., 2009). Therefore, as previously mentioned in Section 3.4.2, different selections of prior distributions for τ should be investigated through a series of sensitivity analyses.

3.4.5.1 Network meta-regression

The model described in Equation (3.20) can be extended to include regression coefficients in what is formally referred to as network meta-regression. Let x_i be the value of the covariate of interest for the i^{th} study, and $\bar{x}$ be the mean of the covariate across studies, then the general network meta-regression framework is described as:

$$
\begin{aligned}
y_{ij} &\sim \text{Normal}(\theta_{ij}, se_{ij}^2) \\
\theta_{ij} &= \mu_i + \delta_{i,bk} I_{\{j=k\}} + \beta(x_i - \bar{x}) \\
I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\tag{3.24}
$$

There are a number of methods for including regression coefficients in meta-analyses (Nixon et al., 2007; Dias et al., 2012; Cooper et al., 2009). Three commonly used meta-regression models include: 1) Seperate regression coefficients for each treatment; 2) Exchangeable regression coefficients between treatments; and 3) A common regression coefficient across treatments (Dias et al., 2012). A separate regression coefficient for each treatment assumes that all treatment by covariate interactions are unrelated to one another for all active treatments versus a comparator, i.e. $\beta_j \sim$Normal$(0, 0.001)$ for each intervention $j = 1, 2, ..., n_t$. However, it may not always be possible to estimate a separate regression coefficient for each treatment by covariate interaction in situations with which there are many interventions, or the data are limited (Cooper et al., 2009). Exchangeable regression coefficients assumes that each of the treatment by covariate interactions are different but related, such that $\beta \sim$ Normal(μ_b, σ_b^2) where μ_b is the overall mean, and σ_b^2 represents the between-treatment variability in the covariate effect of interest. For this model, vague prior distributions with large variability are specified for parameters μ_b and σ_b. Meta-regression models incorporating a common regression coefficient assumes that all treatment by covariate interactions are the same regardless of intervention, i.e. $\beta \sim$Normal$(0, 0.001)$; and thus this approach implies that the covariate effect is identical across all active treatments. In the context of HTA and decision making, it has been argued that a common interaction term is likely to be of most use (Dias et al., 2012).

3.4.5.2 Ranking interventions

One advantage of using NMA to compare multiple interventions is the ability to rank treatments based on their treatment effect estimates. In this thesis interventions were ranked based on the posterior distributions of the relative effect estimates, d, using the `rank`$(\boldsymbol{v}, s)$ function implemented in WinBUGS software. The `rank` function returns the number of elements of vector $\boldsymbol{v}$ whose value is less than or equal to the s^{th} element, such that for k treatments with vector of

treatment effects, $\boldsymbol{d}$ (Welton et al., 2012):

$$\text{rank}_k = \begin{cases} \texttt{rank}(\mathbf{d}, k) & \text{if a higher value is harmful} \\ (k+1) - \texttt{rank}(\mathbf{d}, k) & \text{if a higher value is beneficial} \end{cases} \tag{3.25}$$

Interventions with the largest relative treatment effect for efficacy outcomes, and the highest prevalence for safety and tolerability outcomes, were ranked in first place at each MCMC iteration. The estimated rankings overall were calculated as a summary of the individual ranks at each iteration. Thus, for efficacy outcomes, interventions were ranked from 'best' to 'worst', and a rank of 1 represents the most effective intervention. For safety and tolerability outcomes, interventions were ranked from 'worst' to 'best', and a rank of 1 indicates the most hazardous or least tolerable intervention. The probability that each treatment was the 'best' (for efficacy outcomes), or the 'worst' (for safety and tolerability outcomes) overall was calculated using the `equals` function implemented in WinBUGS. The `equals` function returns the value 1 when treatment k is ranked in first place, and a 0 otherwise (Lunn et al., 2012), thus for efficacy data:

$$\text{p}(\text{best}_k) = \texttt{equals}(\text{rank}_k, 1) \tag{3.26}$$

and for safety and tolerability data:

$$\text{p}(\text{worst}_k) = \texttt{equals}(\text{rank}_k, 1). \tag{3.27}$$

Alternative methods for ranking interventions exist, including the calculation of the surface under the cumulative ranking curve (SUCRA) (Salanti et al., 2011). However, calculating the probability that each intervention was the best or worst overall is generally considered more informative for decision making (Dias et al., 2011a; Salanti et al., 2011), and thus this method was used throughout the remainder of this thesis.

3.4.5.3 Assessing inconsistencies between direct and indirect information

As described in Section 3.4.5, consistencies between direct and indirect evidence are crucial for assumptions underpinning NMA to hold. Inconsistencies between direct and indirect information were assessed, where possible, using a method of node-splitting (Dias et al., 2010). This approach calculates two posterior distributions, one of which is derived from the direct information obtained from head-to-head trials comparing interventions B and C, d_{BC}^{dir}, and the other is indirectly derived, d_{BC}^{ind}, using the additivity assumption described in Equation (3.19). Agreement can be assessed by calculating the difference between direct and indirect estimates, together with the posterior probability that the direct estimate surpasses that of the indirect estimate. A 5% threshold could be applied to indicate a statistically important difference. This method, however, can only be applied to pairs of interventions within a closed loop in the network (i.e. there exists both direct and indirect information).

3.4.6 Advantages of a Bayesian approach to evidence synthesis

Aside from the general advantages of Bayesian methodology described in Section 3.2, there are many advantages of using a Bayesian approach to meta-analyse data (Sutton et al., 2000). As with all Bayesian hierarchical modelling methods, the assumption that treatment effects are exchangeable between trials leads to a 'borrowing of strength' across trial estimates. This results in a shrinkage of the trial-specific estimates towards the overall mean, and consequently a reduction in the width of the estimated credible intervals (Spiegelhalter et al., 2004). Though, incorporating full parameter uncertainty using a Bayesian approach generally results in wider intervals than that of the frequentist alternative. For example, in the most frequently used frequentist approach to pairwise random effects meta-analysis, DerSimonian and Laird (Higgins et al., 2009), uncertainty in the

between-study variance is ignored. Another advantage of a Bayesian approach is the opportunity to make predictions for future studies by making use of posterior predictive distributions drawn from meta-analyses of similar studies (Welton et al., 2012). Predictive distributions could also be valuable in estimating treatment effects in future populations for healthcare decision making. There is an argument for using predictive distributions to summarise treatment effects in random effects meta-analysis models as it is considered a more appropriate summary estimate that takes into account between-study heterogeneity (Spiegelhalter et al., 2004). However, in the interest of health policy decision making, a posterior distribution may be of more use when the participant populations of trials included in meta-analyses resemble the patient population for which the clinical decision is to be made. Thus, for the purpose of this thesis, posterior distributions were used unless otherwise stated.

3.5 Chapter summary

This chapter has introduced key concepts and statistical methodologies that are used throughout this thesis. Subsequent chapters highlight in more detail specific methodologies that are developed and applied in a Bayesian framework. The primary reason for adopting a Bayesian approach in this thesis was not necessarily to incorporate external information in the form of prior distributions, but to maximise the use of a flexible platform with which to model complex clinical scenarios. Chapter 4 illustrates the use of a flexible Bayesian framework for analysing a large RCT in patients with DO, using a novel application of Bayesian methodology to analyse repeat treatment in patients with interval censored data.

Chapter 4

Bayesian Methods for Clinical Trial Evaluation

4.1 Chapter overview

Decisions regarding healthcare policy are frequently determined by trials of alternative treatments; and thus, clinical trials are at the forefront of medical decision making. In a HTA setting, interest often lies in the evaluation of expected patient benefit and cost over a lifetime horizon. Estimation of long-term and continued use of treatment is therefore of great importance in a decision making context. This chapter describes a novel application of Bayesian methods to evaluate repeat treatment of onaBoNT-A in patients with interval censored data. An overview of the motivating example, together with an introduction to the design and conduct of the RELAX trial by Tincello et al. (2012) are described. The proposed statistical models are discussed, and the impact of model misspecification together with disparate proportions of missing data are explored through a series of simulation studies. The work described in this chapter has been published in *Neurourology and Urodynamics* (Owen et al., 2016a) with the full manuscript given in Appendix F.1. Further details and an extension to the analyses are described in this chapter.

4.2 Introduction

Recurrent events are common in clinical trials of chronic diseases. From a HTA perspective, we are often interested in the efficacy of repeat treatment to ensure efficient use of healthcare resources to maximise patient benefit. Due to the ongoing nature of chronic conditions, and consequently clinical trials, patient follow-up can be intermittent and information regarding the time of symptom development is often unreported or poorly reported.

Motivated by the RELAX trial - a large randomised trial in OAB (Tincello et al., 2012), interest lies in evaluating the duration of treatment effect defined as the time to patient-reported return of symptoms (urinary incontinence, voiding, urgency and nocturia) following repeat injection of onaBoNT-A. The primary motivation for this analysis was to explore whether there is a potential waning effect of repeat treatment. Participants from the RELAX trial self-selected repeat treatment of onaBoNT-A and so, there may be an element of selection bias, with which patients who previously responded to onaBoNT-A injections requested further treatment. This phenomenon can lead to a false perception of treatment benefit for repeated injections. During the open label extension study, patients were intermittently lost to follow-up, and thus for these patients, the time of symptom development is uncertain. However, patients would frequently return for repeat treatment, and complete information was obtained for the interval between the date of last complete follow-up and re-treatment. This is often referred to as interval censored data and arises when a failure time can not be observed or is missing, but is known to fall in an interval obtained from a sequence of examinations (Peto, 1973).

In these situations, traditional approaches such as frequentist parametric and semi-parametric models require strong assumptions to be made regarding missing event times. Typically, the first, last or mid point of the interval is taken (Lindsey and Ryan, 1998), though the mid point is most commonly used (Sun, 2007). Such an approach may underestimate standard errors of estimated parameters as the

uncertainty surrounding missing event times is ignored (Radke, 2003). In addition to maximum likelihood estimation, imputation using an expectation-maximisation (EM) algorithm is another option for dealing with interval censored data. Goggins et al. (1998) describes this method using a MCMC sampler for the Cox proportional hazards model, but this procedure is not as easily implemented for models such as the Weibull distribution as it is not a member of the exponential family (Odell et al., 1992). However, parametric distributions such as the Weibull model are more powerful when the underlying assumptions are satisfied (Goggins et al., 1998).

As described in Chapter 3, a Bayesian approach has several advantages, including the flexibility to model complex data structures, as well as the ability to obtain posterior predictive distributions for missing data. Development of Gibbs samplers, such as the WinBUGS software described in Section 3.2.1, have greatly eased computation and made routine fitting of complex models feasible.

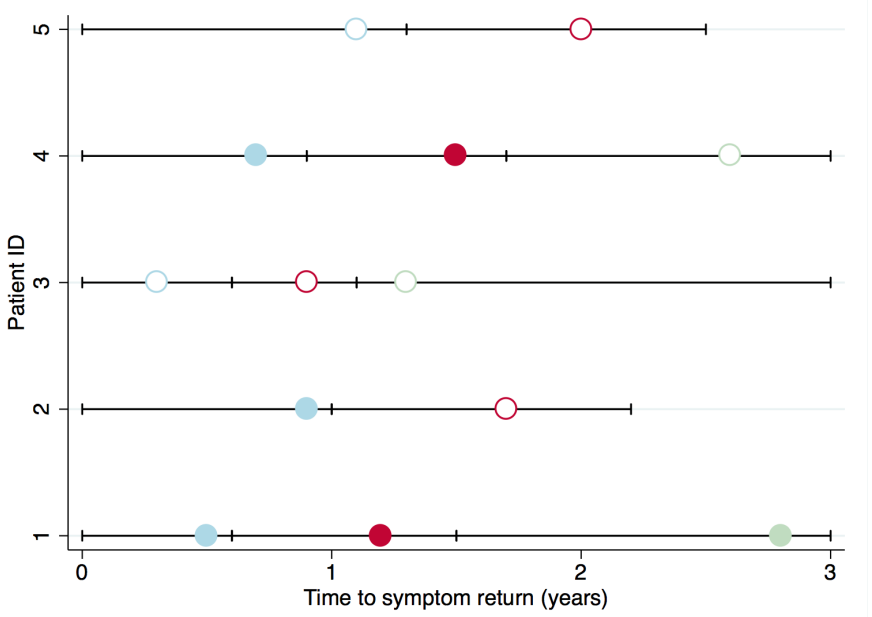
4.2.1 The RELAX trial

The RELAX study was a randomised, double-blind, placebo-controlled trial of 240 women with proven DO on urodynamic testing within two years of recruitment and refractory to standard treatment (Tincello et al., 2012). Women were included if they experienced eight or more voiding episodes and two or more moderate to severe urgency episodes per 24 hours, with or without incontinence. Women with urodynamic stress incontinence, neurologic DO, or voiding dysfunction were excluded. Participants were recruited between July 2006 and November 2009 from eight UK hospitals. The trial was approved by the Scottish Multicentre Research Ethics Committee (04/MRE10/67), and registered on Current Controlled Trials (ISRCTN26091555) on May 26^{th} 2005.

Participants were randomised on a 1:1 basis to receive 200 units of onaBoNT-A or placebo, injected in 20 sites, sparing the trigone. Blinded outcome data were collected at baseline, 6 weeks, 3 months, and 6 months. Following completion of the blinded trial, participants entered a 5 year open-label extension study after further informed, written consent, and were offered a maximum of 2 further onaBoNT-A injections. Thus patients could have a maximum of 2 active injections for participants initially randomised to placebo, and a maximum of 3 active injections for participants initially randomised to onaBoNT-A, which for the remainder of this chapter will be termed active injection 1, 2, and 3. The provision of 2 injections was determined by the level of drug provision provided by Allergan (BOTOX®, Allergan USA). The final treatment for the final patient was administered at the end of May 2013.

Outcome data, including bladder diaries, urgency episode frequency, and quality of life questionnaires, were collected by post every 6 months from the date of the first (randomised) injection throughout the extension study. Patients were in regular contact with the local continence nurse specialists. With each follow up, patients were asked the question "have your symptoms returned?" and a complete record of symptom development was recorded. Repeat injection requests could be initiated by patient request in response to the question "do you wish to have a repeat injection at this time?". Participants could also request treatment at any time between follow-up contact. Patients were sent a follow-up pack including urinary bladder diaries, I-QOL questionnaire (Wagner et al., 1996), and ICIQ short form (ICIQ-SF) (Avery et al., 2004), as previously described in Section 2.4 of Chapter 2, as well as the Indevus Urgency Severity Scale (IUSS) (Nixon et al., 2005), and Patient Global Impression of Improvement (PGI-I) (Tincello et al., 2013), 6 weeks after every subsequent injection, in addition to the scheduled 6 monthly review. Throughout the assessment period, patients were intermittently lost to follow-up, and thus for these patients the time of symptom recurrence is uncertain, though, complete information was obtained for the time of re-injection. In this analysis, duration of treatment effect was based solely on self-reported return of symptoms.

FIGURE 4.1: Example of patient intervals



Capped lines represent specific patient intervals, solid circles represent known event times, and hollow circles represent unknown event times for active injection 1(blue), 2(red), and 3(green).

A schematic illustration of patient follow-up is given in Figure 4.1, where the lines represent the specific patient intervals. For active injection 1 (blue), 2 (red), and 3 (green), the solid circles represent documented times of symptom recurrence, and the hollow circles represent patients with missing event times.

4.3 Methods

Models were fitted in a Bayesian framework in which the correlation between repeated events within individuals were accounted for by incorporating a shared frailty term (Pennell and Dunson, 2006). The time origin was defined as the date participants received active injection of 200 units of onaBoNT-A. A clock reset model was used where the clock was reset to zero for every subsequent active injection. A stochastic approach was adopted to model the time of symptom return following active injection using Bayesian MCMC simulation and implemented in WinBUGS 1.4.3. The basic parametric models currently available as

stochastic distributions in WinBUGS include the exponential, generalised gamma, log-normal, and Weibull models, which are further described in Chapter 3, Section 3.2.4. Other survival models such as the Cox model (Clayton, 1994) and Royston-Parmar model (Royston et al., 2011) can be implemented in WinBUGS using the 'zeros' or 'ones trick' (Spiegelhalter et al., 2003; Lunn et al., 2012). In the absence of built-in distributions, the 'zeros' or 'ones trick' allows the user to specify a general likelihood. This approach, however, expresses survival times as a deterministic parameter and thus, survival times are assumed known. In the case of the RELAX trial, there exists interval censored data for which survival times are not known. Therefore, all stochastic parametric models available in WinBUGS 1.4.3 were evaluated for model fit. An advantage of using a stochastic approach includes the ability to obtain posterior predictions for missing data, in this case, participants active injection failure times, whilst incorporating full parameter uncertainty.

As shown in Appendix A.1, all of the basic parametric distributions poorly represent the data. It is worth noting that the exponential model appears to have a relatively better fit to the data (Figure A.1) than that of the generalised gamma model (Figure A.2). However, further inspection of model fit, specifically at earlier points in time using log-log plots of survival (Figure A.5), illustrates that the exponential model severely overestimates survival at earlier time points. Whilst the generalised gamma model also overestimates survival at earlier time points it appears to have a relatively better fit to the data than that of the exponential model. Overall, it is clear that all of the basic parametric distributions poorly capture the shape of the data and this is largely because the underlying hazards for each injection have more than one turning point. A poly-Weibull model presented by Berger and Sun (1993), is an extension of the Weibull model and naturally lends itself to this type of scenario. Typically, poly-Weibull models are used to describe competing risks in survival analysis data, where the underlying hazard may have an increasing or decreasing hazard rate at different time points causing a bathtub- or U-shape distribution. In this situation, the poly-Weibull model

provides a flexible framework in which to model the underlying hazard rate. In the case of the RELAX trial, this additional flexibility was required to capture the different hazard rates over time. Recently, Demiris et al. (2015) have implemented the poly-Weibull model as a stochastic distribution under the WinBUGS Development software.

4.3.1 Model 0: Bayesian parametric frailty model

The survivor function of the Weibull model for time to symptom development t, can be written as

$$S(t) = \exp(-\mu t^{\gamma}) \tag{4.1}$$

with hazard function

$$h(t) = \mu\gamma t^{\gamma-1} \tag{4.2}$$

and probability density function

$$f(t) = \mu\gamma t^{\gamma-1} \exp(-\mu t^{\gamma}) \tag{4.3}$$

where, μ is defined as the exponent of the hazard model, more commonly known as the scale parameter, and γ the shape of the underlying hazard. To account for the correlation between repeated events within the k^{th} individual, an additive random effect b_k was incorporated so that for observation $i = 1, ..., n$, $\mu_i = \exp^{\boldsymbol{\beta}\mathbf{Z_i}+b_k}$, where $\mathbf{Z_i}$ is a vector of covariates and $\boldsymbol{\beta}$ a vector of regression coefficients. To appropriately account for interval censored patients (i.e. patients with missing event times) and right censored patients (i.e. patients completely lost to follow-up) a truncated Weibull distribution was assumed, $t_i \sim \text{Weibull}(\gamma, \mu_i)\text{I}(C_i, R_i)$, with lower bound corresponding to the time of last complete follow-up (or censoring time), C_i, and upper bound equal to the time of re-treatment, R_i. For right censored patients, the upper bound R_i was assumed at the end of the study. Plausibly vague prior distributions were assigned for the vector of regression coefficients $\boldsymbol{\beta} \sim \text{Normal}(0, 10^2)$, and random effects $b_k \sim \text{Normal}(0, \tau^2)$ where $\tau \sim \text{Uniform}(0, 2)$. A Gamma prior, described in Section 3.2.4, was applied for the

shape parameter, $\gamma \sim$ Gamma $(10^2, 10^2)$. Sensitivity analyses assessed the impact of the choice of plausibly vague prior distributions on the shape parameters, γ, where a Uniform$(0, 1)$ distribution, and a Normal$(0, 10^2)$ distribution on the log scale, were considered.

4.3.2 Model 1: Bayesian flexible parametric frailty model

If event times t under m rates are assumed to be independently distributed and the individual components follow a Weibull distribution as in Section 4.3.1, then the survivor function of the poly-Weibull model can be expressed as:

$$S(t) = \exp\left(-\sum_{j=1}^{m} \mu_j t^{\gamma_j}\right). \tag{4.4}$$

The hazard function can be described as the sum of m independent Weibull hazards:

$$h(t) = \sum_{j=1}^{m} \gamma_j \mu_j t^{\gamma_j - 1} \tag{4.5}$$

with probability density function:

$$f(t) = \exp\left(-\sum_{j=1}^{m} \mu_j t^{\gamma_j}\right) \sum_{j=1}^{m} \gamma_j \mu_j t^{\gamma_j - 1} \tag{4.6}$$

Following the notation outlined in Section 4.3.1, for $j = 1, ..., m$ rates, μ_j is defined as the exponent of the hazard model, and γ_j the shape of the underlying hazard. For observation $i = 1, ..., n$, the hazard model is defined as $\mu_{ij} = \exp^{\boldsymbol{\beta}\mathbf{Z_i} + b_k}$, where b_k is an additive random effect for the k^{th} individual, $\mathbf{Z_i}$ is a vector of covariates and $\boldsymbol{\beta}$ is a vector of regression coefficients.

Note that when the shape parameters $\gamma_1, \gamma_2, ...\gamma_m$ are equal, the poly-Weibull model reverts to a standard Weibull distribution, and the parameters $\mu_1, \mu_2, ..., \mu_m$ become unidentifiable (Ibrahim et al., 2005). To ensure identifiability of the hazard

components, an ordering constraint can be placed on the shape parameters as described by Demiris et al. (2015). To apply ordering constraints, the joint prior is multiplied by an indicator function λ, equal to 1, given by:

$$\lambda = \prod_{j=1}^{m-1} I(\gamma_{j+1} - \gamma_j) \tag{4.7}$$

$$I(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Equation (4.7) forces $\gamma_j < \gamma_{j+1}$, and consequently imposes increasing ordering constraints on the multiple shape components, thus $\gamma_j \neq \gamma_{j+1}$.

Using the same notation outlined in Section 4.3.1, a truncated poly-Weibull model was assumed, $t_i \sim$ poly-Weibull$(\gamma_j, \mu_{ij})I(C_i, R_i)$, where the lower bound of the interval corresponds to the time of last complete follow-up (or censoring time), C_i, and the upper bound is equal to the time of re-treatment, R_i. Plausibly vague prior distributions were assigned for the vector of regression coefficients $\boldsymbol{\beta} \sim$ Normal$(0, 10^2)$, and random effects $b_k \sim$ Normal$(0, \tau^2)$ where $\tau \sim$ Uniform$(0, 2)$. Normal priors were specified for the shape parameters on the log scale, $\log(\gamma_j) \sim$ Normal$(0, 10^2)$ for $j = 2, ..., m$.

4.3.3 Model 2: Bayesian flexible parametric frailty model incorporating covariates

In the case of the RELAX trial, it may not be reasonable to assume the same underlying hazard for each active injection. To allow for greater flexibility, covariate effects can be incorporated onto the shape parameters γ_j. As mentioned above in Section 4.3.2, constraints must be applied to the shape parameters to ensure identifiability of the hazard component (Ibrahim et al., 2005). This is more difficult in situations where covariate effects are applied to both the shape and rate components (Demiris et al., 2015). In order to impose constraints on the shape

parameters such that $\gamma_j \neq \gamma_{j+1}$, constraints must be placed on all possible combinations of covariates. As described by Demiris et al. (2015), these constraints can be placed by assigning the indicator function λ as follows:

$$\lambda = \prod_{l=1}^{2} \prod_{j=1}^{m-1} C_{lj}, \tag{4.8}$$

where,

$$C_{1j} = I(min_q \left\{w_{(j+1)q}\right\} - min_q \left\{w_{jq}\right\})$$

$$C_{2j} = I(max_q \left\{w_{(j+1)q}\right\} - max_q \left\{w_{jq}\right\})$$

and

$$I(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

If there are p covariates, then w_{jq} represents the q^{th} element of the 2^p-dimensional vector $w_j = W\alpha_j$, where $\alpha_j = (\alpha_{0j}, \alpha_{1j}, ..., \alpha_{pj})'$ and W is defined as a $2^p \times (p+1)$ matrix containing all possible combinations of the minimum and maximum values of each covariate (Demiris et al., 2015). In this clinical example, there are 2 additional covariates (active injection 2 and 3), both of which act as indicator variables taking the values 0 or 1, thus α_j can be defined as $\alpha_j = (\alpha_{0j}, \alpha_{1j}, \alpha_{2j})'$ and W can be described as:

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

For example, a model with 2 shape components, w_1 and w_2 can be written as:

$$w_1 = \begin{pmatrix} \alpha_{01} \\ \alpha_{01} + \alpha_{21} \\ \alpha_{01} + \alpha_{11} \\ \alpha_{01} + \alpha_{11} + \alpha_{21} \end{pmatrix}, w_2 = \begin{pmatrix} \alpha_{02} \\ \alpha_{02} + \alpha_{22} \\ \alpha_{02} + \alpha_{12} \\ \alpha_{02} + \alpha_{12} + \alpha_{22} \end{pmatrix}.$$

4.3.4 Restricted mean survival time (RMST)

The primary motivation of this analysis, as described in Section 4.2, was to assess a potential waning effect of repeat treatment of onaBoNT-A. In order to compare active injections, restricted mean survival time (RMST) was calculated for all models. One advantage of using RMST is that it is valid under all distributional forms of time to event data. Another advantage is that RMST has a more clinically meaningful interpretation than that of hazard ratios (Royston and Parmar, 2013) and can be interpreted as the average time to symptom return of participants between treatment administration (e.g. $t = 0$) and a specified time horizon ($t = t* > 0$). RMST is equal to the area under the survival curve $S(t)$ from $t = 0$ to $t = t*$ and thus for active injection $x = 1, 2, 3$, RMST is given by:

$$\text{RMST}_x = \int_0^{t*} S_x(t)dt \tag{4.9}$$

The area under the survival curve was calculated in WinBUGS using the trapezium rule so that,

$$\text{RMST}_x = \int_0^{t*} S_x(t)dt = \sum_{t=0}^{t=t*-h} \frac{1}{2}h[S_x(t) + S_x(t+h)] \tag{4.10}$$

where h is the distance between uniform time intervals. In the case of the RELAX trial, h=0.1 years and time horizon $t*$=3 years was used for all analyses.

To calculate the difference between active injections 1, 2 and 3 with survival functions $S_1(t)$, $S_2(t)$ and $S_3(t)$, the difference in RMST, $\Delta_{(y+1)y}$ was calculated, where $\Delta_{(y+1)y}$ is the area between survival curves for active injection $(y+1)$ relative to active injection $y = 1, 2$:

$$\Delta_{(y+1)y} = \int_0^{t*} S_{(y+1)}(t)dt - \int_0^{t*} S_y(t)dt \tag{4.11}$$

The probability that RMST for repeat treatment was greater than that of the previous treatment were calculated by monitoring the number of MCMC iterations for which the RMST for active injection $(y+1)$ surpasses that of active injection y, i.e. $p(\text{RMST}_2 > \text{RMST}_1)$ and $p(\text{RMST}_3 > \text{RMST}_2)$.

4.3.5 Model comparison and computation

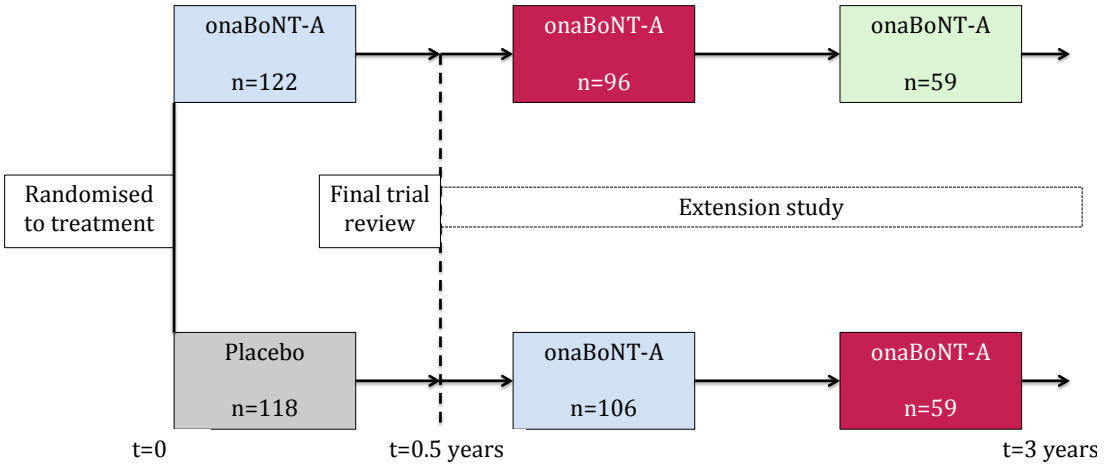
Poly-Weibull models with 2 or 3 shape components, with and without covariate effects, were assessed and compared to naive Weibull and poly-Weibull models assuming mid-point censoring. In this instance, the DIC was not utilised for model comparison due to poor estimation of the effective number of parameters. This is often the case when several of the posterior distributions are skewed (Demiris et al., 2015). As the total number of parameters are difficult to calculate for random effects models, the difference in mean posterior deviance, calculated by monitoring the deviance node in WinBUGS, penalised by the difference in the effective number of parameters were used for model comparison.

Results are based on 20,000 samples for 2 chains with disparate starting values, where the first 10,000 samples were discarded in the form of a burn-in. Model convergence was assessed through visual inspection of Brook-Gelmin-Rubin, autocorrelation, history, and density plots as described in Section 3.2.2 of Chapter 3. Example WinBUGS code for Models 0, 1 and 2 are given in Appendix A.2, A.3 and A.4, respectively.

4.4 Results

A total of 240 women were enrolled and treated; 122 women were initially randomised to onaBoNT-A and 118 to placebo. During the randomised study and 5-year open-label extension study, 442 active injections were administered: 228 participants received first active (122 of which were originally randomised to onaBoNT-A), 155 received second active (96 of which were originally randomised to onaBoNT-A), and 59 received third active injections (all 59 of which were originally randomised to onaBoNT-A) (Figure 4.2).

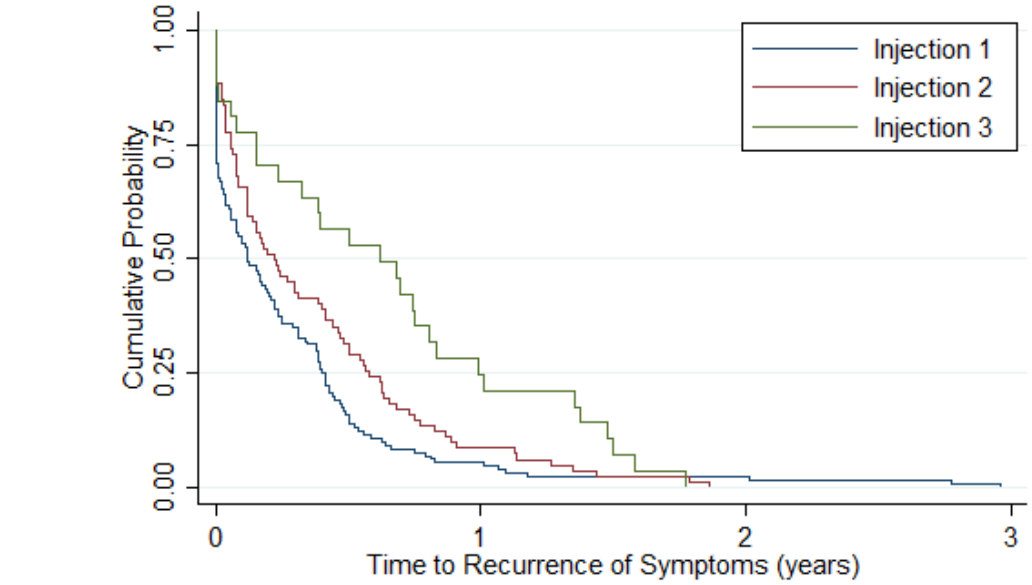
FIGURE 4.2: Repeat treatment timeline



Overall, 189 (83%), 112 (72%) and 31 (53%) participants experienced symptom return after active injection 1, 2 and 3, respectively. Of these, 47 (25%) and 25 (22%) patients were interval censored for active injections 1 and 2. For active injection 3 it was difficult to distinguish between participants that failed to report symptom return, and those without symptoms. This is largely because participants seldom presented to a healthcare professional following the last available treatment, and these participants were thus treated as right censored. In total, 39 (17%), 43 (28%), and 28 (47%) participants were lost to follow-up for injection 1, 2 and 3, and these participants were right censored at the date of last complete follow-up. Figure 4.3 displays the Kaplan-Meier curve for patient-reported return of symptoms for each number of active injection. Patients with symptom return

but with missing event times were omitted from this exploratory analysis and thus 181, 130 and 59 patients were at risk of symptom return from the start of treatment for injection 1, 2, and 3, respectively.

FIGURE 4.3: Kaplan-Meir curve for patient reported time to recurrence of symptoms



Number at risk

Injection 1	181	7	3	0
Injection 2	130	7	0	0
Injection 3	59	7	0	0

Figure 4.4 illustrates that a flexible parametric distribution, such as a poly-Weibull distribution, is required to appropriately model the data compared to that of more basic parametric distributions given in Appendix A.1. This is further illustrated through model fit statistics given in Table 4.1. As shown through the difference in penalised deviance, a flexible parametric model such as a 2-component poly-Weibull model assuming mid-point censoring was of a better fit to the data than that of the Weibull equivalent. A poly-Weibull model without a censoring assumption had a better fit to the data than that of a poly-Weibull model assuming mid-point censoring. Overall, the 3-component poly-Weibull model without covariate effects and without a censoring assumption (Model 1) was the best fitting model, with the largest difference in penalised deviance of -1189.7. Incorporating covariate effects on the shape parameters (Model 2) led to no further benefit in model fit. However, incorporating this additional complexity led to computational difficulties which resulted in the models repeatedly sampling extreme values for parameter estimates. Thus, models failed to reach convergence and model fit statistics should be interpreted with caution. Due to issues with model convergence for Model 2, the remaining results will focus on Models 0 and 1.

TABLE 4.1: Model fit statistics

Model		**Deviance**	**Difference in deviance**	**Difference in no. of parameters**	**Difference in penalised deviance**
Weibull model (Model 0) with mid-point censoring	1 component	725.1	Reference	Reference	Reference
poly-Weibull model without covariates (Model 1) with mid-point censoring	2 component	434.4	-290.7	1	-289.7
poly-Weibull model without covariates and without censoring assumption (Model 1)	2 component	-244.2	-969.3	1	-968.3
	3 component	-466.6	-1191.7	2	-1189.7
poly-Weibull model with covariates and without censoring assumption (Model 2)	2 component	-240*	-965.1*	5	-960.1*
	3 component	-237.7*	-962.8*	7	-955.8*

*To be interpreted with caution due to issues with model convergence

FIGURE 4.4: Predicted survival of the 2-component poly-Weibull model without covariates stratified by active injection

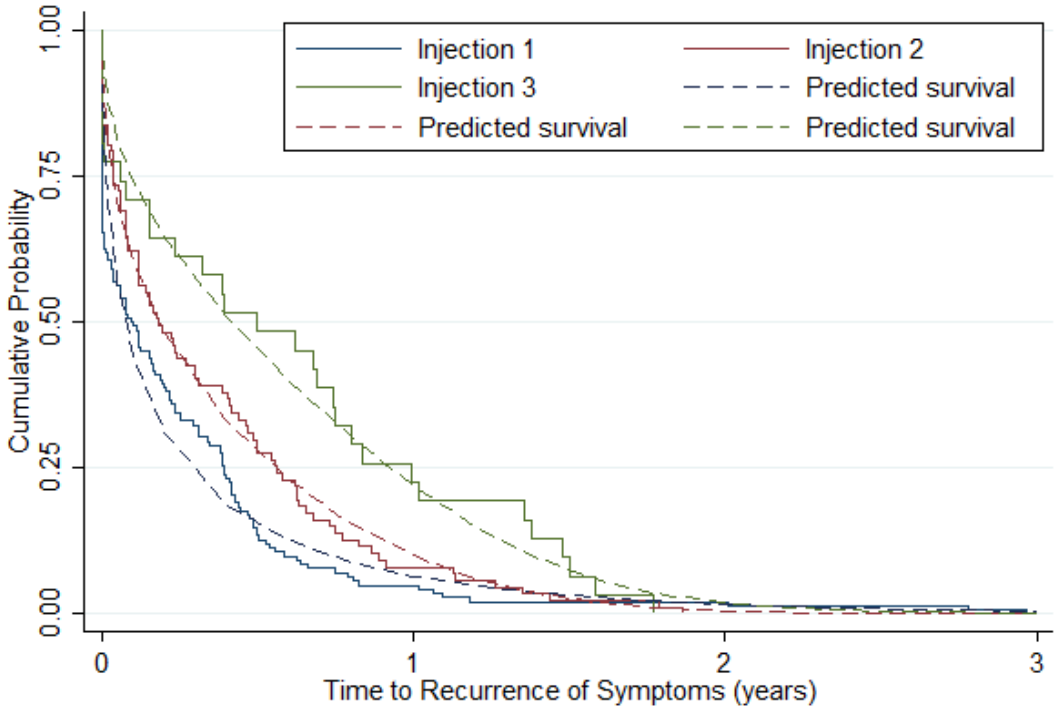


Table 4.2 displays the table of results for all models without covariate effects. It is apparent that the 3-component poly-Weibull model has reduced uncertainty in the estimated RMST compared to all other models. Across all models, uncertainty in the estimates of a 3-component poly-Weibull model without a mid-point censoring assumption was reduced in the range of 30%-250%, 21%-186% and 6%-40% for active injection 1, 2 and 3. It is apparent that estimates of RMST from models assuming mid point censoring are much larger than estimates obtained from models without a censoring assumption. This increase in RMST is likely to be a result of the wide time intervals between participants last complete follow-up and administration of repeat injection, and thus models assuming mid-point censoring vastly overestimates the average treatment effect. The RMST for active injections 1, 2, and 3 were 0.11 (95%CrI: 0.07, 0.17), 0.19 (95%CrI: 0.13, 0.27) and 0.35 (95%CrI: 0.20, 0.55) years, respectively. Though the RMST increased with active injection, it is apparent that the 95% credible intervals for active injections 1 and 2, and 2 and 3, are overlapping, suggesting that there is no important difference in efficacy between increased numbers of repeat treatment. However, the probabilities that RMST for repeat treatment surpasses that of previous treatment were 99%. There appears to be a considerable difference in efficacy between active injection 1 and active injection 3. On average, active injection 3 appears to delay symptom return for approximately 87 additional days (0.24 years) compared to active injection 1.

TABLE 4.2: Table of results for models without covariate effects

	Weibull with mid-point censoring		**2-component poly-Weibull with mid-point censoring**		**2-component poly-Weibull without censoring assumption**		**3-component poly-Weibull without censoring assumption**	
Parameter	**Mean**	**95%CrI**	**Mean**	**95%CrI**	**Mean**	**95%CrI**	**Mean 95%CrI**	**95%CrI**
$RMST_1$	0.93	(0.81, 1.05)	1.22	(1.06, 1.41)	0.22	(0.16, 0.29)	0.11	(0.07, 0.17)
$RMST_2$	0.97	(0.83, 1.13)	1.24	(1.05, 1.45)	0.32	(0.24, 0.41)	0.19	(0.13, 0.27)
$RMST_3$	0.98	(0.75,1.24)	1.07	(0.87, 1.26)	0.53	(0.36, 0.73)	0.35	(0.20, 0.55)
γ_1	0.58	(0.53, 0.63)	0.41	(0.36, 0.46)	0.51	(0.44, 0.58)	0.68	(0.55, 0.79)
γ_2	-	-	15.92	(13.52, 18.86)	2.2	(1.44, 7.13)	1.62	(0.67, 3.10)
γ_3	-	-	-	-	-	-	2.79	(2.18, 13.58)
$\mathrm{p}(RMST_2 > RMST_1)$	0.67	-	0.56	-	0.97	-	0.99	-
$\mathrm{p}(RMST_3 > RMST_2)$	0.53	-	0.10	-	0.99	-	0.99	-

$RMST_x$ denotes restricted mean survival time in years for active injection x, γ_n denotes the shape parameter for n components, $\mathrm{p}(RMST_{(y+1)} > RMST_y)$ denotes the probability that RMST for active injection $(y+1)$ surpasses that of active injection y.

In the case of the RELAX trial, there may be a potential selection effect in the patient population consenting to further treatment. To investigate this further, analysis of variance (ANOVA) adjusting for differences in baseline symptoms were used to assess the variability of mean diary data at 6 weeks post-injection for each of the active injections. Table 4.3 displays the average patient symptoms at 6 weeks following each of the active onaBoNT-A injections. Notably, women who opted for further injection had less severe symptoms at the preceding 6 week follow-up compared to the entire cohort, though these differences cannot be examined using statistical t-tests as these patients contribute to both statistics (i.e., the entire cohort of patients and the subgroup continuing treatment). The difference in symptom severity between those who did, and did not opt for further injection is particularly apparent for incontinence episodes. Patients receiving a third active injection experienced on average 1.71 (3.44) incontinence episodes daily at 6 weeks following the first onaBoNT-A injection compared to 2.53 (3.76) incontinence episodes daily for the entire cohort, suggesting that there was a clear selection effect for women continuing treatment.

Further exploratory analyses of known events found that for active injection 1 (Figure 4.5), patients initially randomized to onaBoNT-A in the double-blinded trial had a shorter duration of effect (RMST: 0.07, 95%CrI: 0.04, 0.13), on average, compared to patients initially randomized to placebo (RMST: 0.16, 95%CrI: 0.09, 0.24) i.e. patients blinded to treatment had a more rapid rate of symptom return, on average, compared to unblinded patients. Furthermore, for active injection 2 (Figure 4.6) - where the entire cohort of patients were unblinded to treatment - there is no difference in duration of treatment effect between patients initially randomised to onaBoNT-A and those randomised to placebo (RMST: 0.15, 95%CrI: 0.08, 0.25, versus RMST: 0.15, 95%CrI: 0.09, 0.24, respectively). Thus suggesting that there may be an element of an extended placebo effect.

TABLE 4.3: Average patient symptoms at 6 weeks following each active onaBoNT-A injection

			Mean episode frequency after each active injection			
	Active injection*	**No. of patients**	**1**	**2**	**3**	**p-value†**
Incontinence episodes	1	201	2.53 (3.76)	-	-	-
	2	133	2.35 (3.85)	1.91 (2.61)	-	0.14
	3	50	1.71 (3.44)	1.4 (2.60)	1.4 (2.68)	0.08
Urgency episodes	1	197	3.74 (4.15)	-	-	-
	2	136	3.28 (3.92)	2.77 (3.18)	-	0.66
	3	50	3.34 (4.02)	2.63 (2.84)	2.77 (2.47)	0.49
Voiding episodes	1	204	8.41 (3.39)	-	-	-
	2	138	8.24 (3.43)	7.43 (2.42)	-	0.27
	3	50	7.99 (2.91)	7.06 (2.02)	7.35 (1.99)	0.11

Results are presented as mean(SD) where each row includes all non-missing data for patients from the cohort receiving each number of injection

*Refers to the sequence of active injection received, as outlined in Figure 4.2

†Significance test is comparing across columns (i.e. comparing injection 3 with injection 2 and 1 in the same cohort of patients)

The overall clinical finding that the fact that there is no important difference between increasing numbers of active injection would not change with model choice. However, models assuming mid-point censoring appeared to have an increased treatment effect, with the average time of symptom return estimated at approximately 1 year for each stratum. This is likely to be the result of extensive time intervals between the last patient follow-up and time to re-treatment i.e. large intervals for interval censored data; and thus assuming mid-point censoring would exaggerate the treatment effects.

4.4.1 Convergence diagnostics

Convergence diagnostics for RMST and shape parameters (γ_m) estimated from the 3-component poly-Weibull model are given in Section A.5 of Appendix A and described in Section 3.2.2. Examination of Brook-Gelman-Rubin plots showed that the ratio of between- and within-chain variability, R, had converged to 1, which would suggest convergence of the samples. The autocorrelation plots showed successive iterations of RMST and γ_m appeared to be sampled from independent posterior distributions leading to adequate mixing and quicker convergence. The density plots for RMST looked reasonably smooth displaying the characteristic bell-shaped appearance of a normal distribution. The history and trace plots appeared to be reasonably stable, displaying random noise in the chains. Overall, the diagnostic plots provided evidence to suggest that the posterior estimates for RMST and γ_m were obtained from samples with little evidence of non-convergence.

FIGURE 4.5: Time to patient reported symptom recurrence for active injection 1 stratified by
randomised treatment allocation

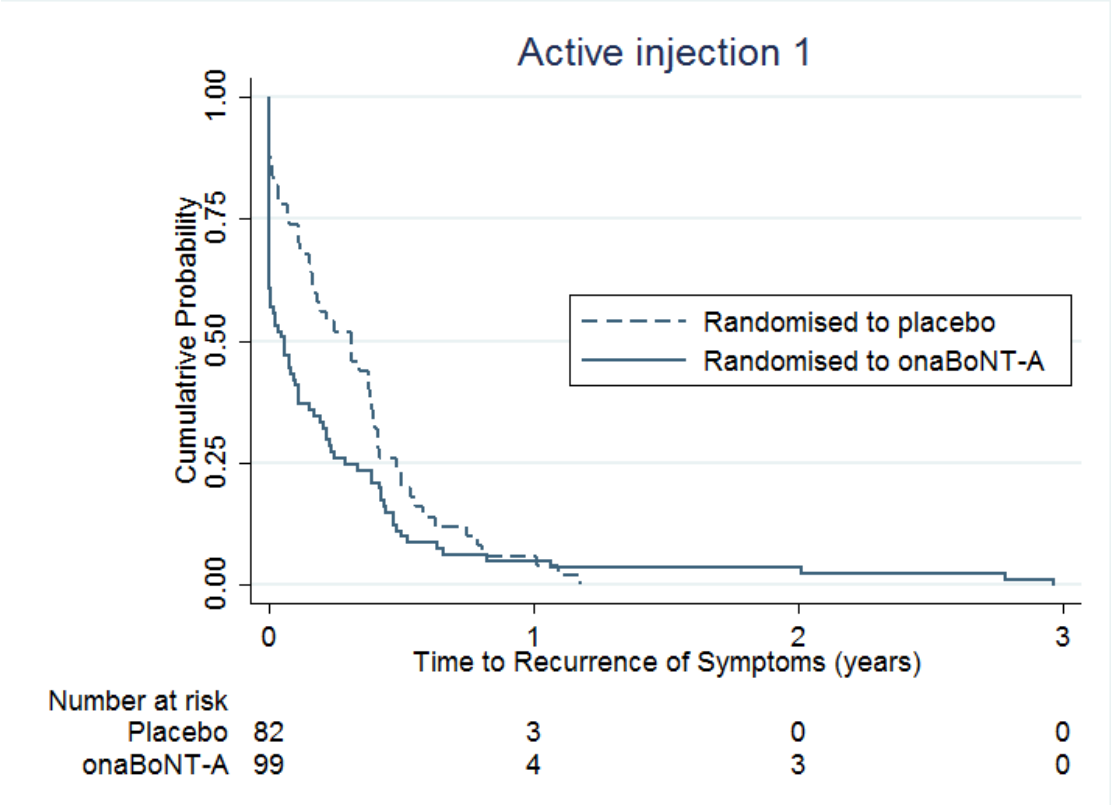
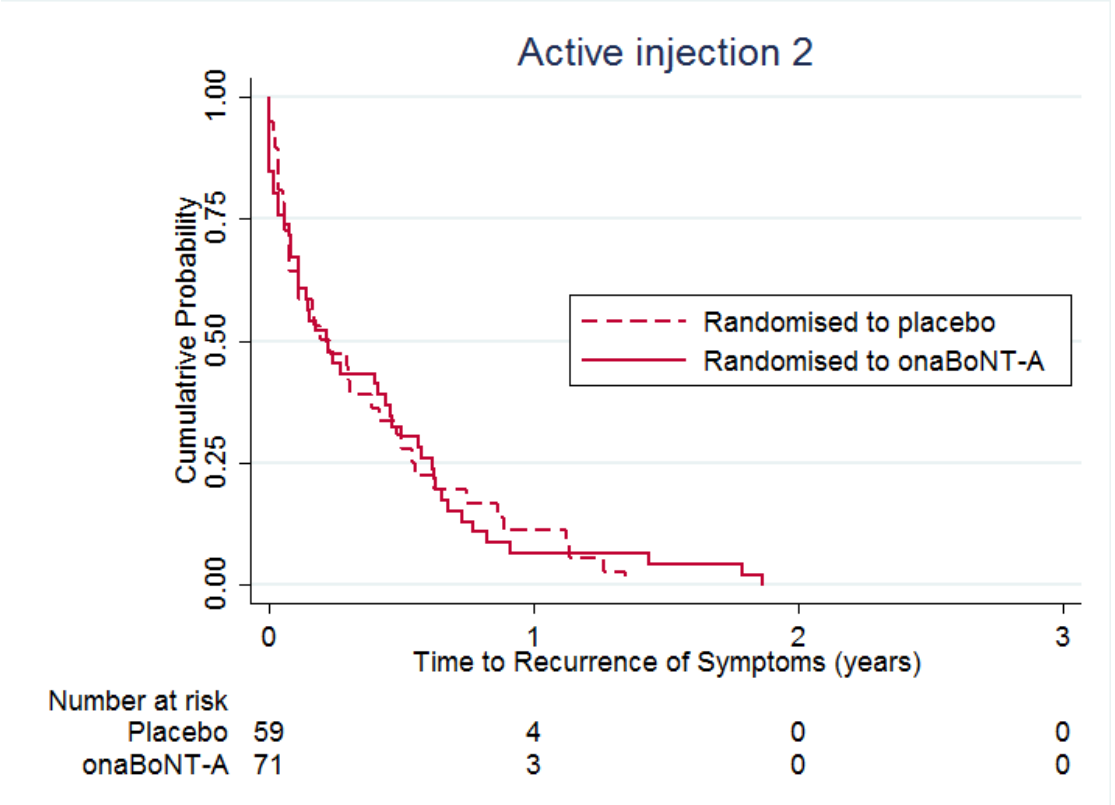


FIGURE 4.6: Time to patient reported symptom recurrence for active injection 2 stratified by
randomised treatment allocation



4.5 Simulation study

4.5.1 Introduction

Simulation studies use computationally intensive methods to evaluate the performance of statistical models relative to a known truth. Such an approach enables empirical estimation of the sampling distribution of the parameters of interest, which simply could not be obtained through evaluation of a single study (Burton et al., 2006). As the true parameter value is known, simulation studies allow calculation of performance measures such as the bias in estimated parameters, and the coverage of 95% confidence intervals. The simulation study described in this Chapter was designed to reflect a similar clinical scenario to that of the RELAX trial. Performance of Bayesian prediction modelling were assessed for increasing proportions of missing data i.e. increased numbers of interval censored data, together with the impact of model misspecification.

4.5.2 Methods

4.5.2.1 Simulation procedures

As described earlier in this chapter, computation of poly-Weibull models were difficult with the inclusion of covariates and large proportions of missing data. For computational ease, a more basic Weibull distribution was selected to assess the performance of Bayesian prediction modelling with different missing data structures. Three missing data structures with increased proportions of missing event times, all assumed to be missing at random (MAR), were considered: 10%, 25%, and 50%. Model performance using Bayesian predictive posterior distributions to sample missing data were assessed. For comparison, a complete case assessment was undertaken i.e. the corresponding missing event times were removed from the analysis. Impact of model misspecification using the Weibull model was further assessed by simulating data from a 2-component poly-Weibull distribution, which

allows for a more complex underlying hazard than that of the Weibull distribution. Thus, in total, 9 different scenarios were considered:

1. Bayesian prediction modelling with 10% of event times missing at random
2. Bayesian prediction modelling with 25% of event times missing at random
3. Bayesian prediction modelling with 50% of event times missing at random
4. Complete case analysis with 10% of event times missing at random
5. Complete case analysis with 25% of event times missing at random
6. Complete case analysis with 50% of event times missing at random
7. Bayesian prediction modelling under model misspecification with 10% of event times missing at random
8. Bayesian prediction modelling under model misspecification with 25% of event times missing at random
9. Bayesian prediction modelling under model misspecification with 50% of event times missing at random

To limit the sampling variation, the simulation study was run over 1000 iterations. At each iteration, WinBUGS 1.4.3 (Spiegelhalter et al., 2003) was called from within `STATA` 14 (StataCorp., 2015) to fit the proposed model and dataset from each scenario. The first 10,000 MCMC iterations were discarded in the form of a burn-in period, and the following 10,000 MCMC iterations were collected in the sampling phase. A summary of the estimated parameters monitored in the sampling phase were stored at each iteration of the simulation study, and summarised over the total number of simulated runs. Summary results were then compared to the 'true' values used to simulate the data. In keeping with the motivating clinical question, log hazard ratios (logHR) of repeat active injection were selected as the estimands of interest as they can feasibly be fixed in the data generating mechanism, and are valid under all distributional forms of the

simulation study. In this instance, RMSTs were not used as estimands in the simulation study because they were more difficult to fix in the data generating mechanism. This is because RMSTs were functional parameters in the model, represented in terms of basic parameters such as log hazard ratios. In addition, clinical interpretation was of less importance in analysing model performance, and therefore log hazard ratios were preferred.

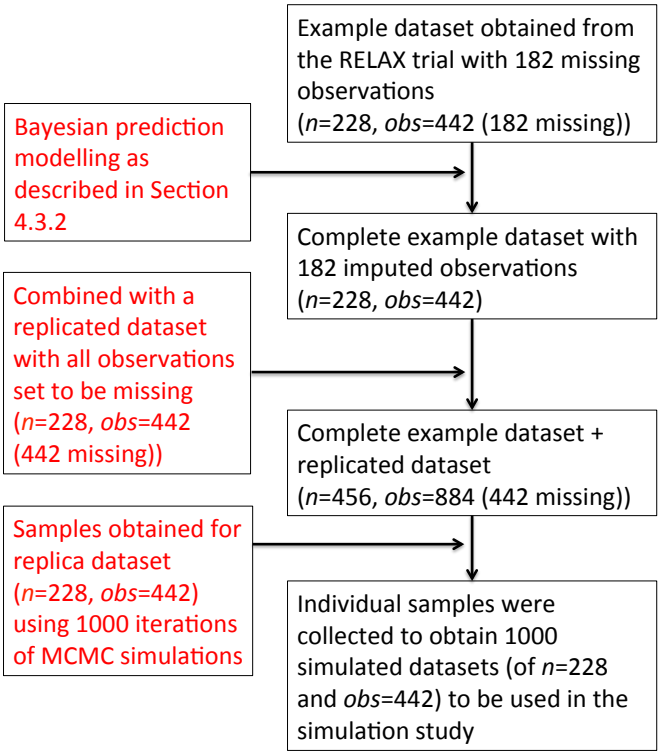
4.5.2.2 Data generating mechanism

Replicating complex situations as seen in clinical practice can often be difficult. For example, a series of covariates are rarely fully independent of one another, and a data generating mechanism incorporating correlation structures for multivariate data is required (Burton et al., 2006). For this reason, a novel Bayesian approach was adopted to replicate the complexity of the RELAX trial. In this example, event times are expected to be associated with patient baseline severity and between repeated events within individuals. In order to capture the complexity of correlation structures in the simulated data, covariates were placed on the log hazard regression models ($\log(\mu)$), further described in Section 4.3.1.

Figure 4.7 illustrates the flow diagram for the data generating mechanism of simulated datasets. Complete data including 228 participants receiving 442 active injections were obtained from the predictive modelling methods described in Section 4.3.2. This data were combined with a replicated dataset of complete patient information but with missing event times. Thus, in total there were 884 observations, with which, 442 were set to be missing. Shape (γ) and scale (μ) parameters of the Weibull and where appropriate, the poly-Weibull model, were fixed based on data from the RELAX trial. Using a Bayesian framework, predictive posterior distributions were used to sample missing event times for the replica dataset using MCMC simulation. Samples were collected from 1000 MCMC iterations to form 1000 different simulated datasets. The replica datasets were then used in the simulation study, where MAR techniques were applied by drawing an indicator

function from a corresponding Bernoulli distribution. All simulations were performed in WinBUGS 1.4.3 (Spiegelhalter et al., 2003) and `STATA` 14 (StataCorp., 2015).

FIGURE 4.7: Flow diagram illustrating the data generating mechanism



n denotes the total number of patients; *obs* denotes the total number of observations

4.5.2.3 Performance measures

Model performance was assessed on bias, accuracy, and coverage. Bias is a measure of deviation in the estimate from the true parameter value and provides an indication of model performance. For the number of simulations m, bias can be calculated as the average difference between the summary estimate of the logHR drawn from the simulation study ($\widehat{logHR}$) and the true logHR ($logHR$) (Burton et al., 2006):

$$\text{Bias} = \frac{1}{m}\sum_{i=1}^{m}\widehat{logHR} - logHR \tag{4.12}$$

Percentage bias can be calculated providing that the true parameter value does not equal zero:

$$\text{Percentage bias} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{\widehat{logHR} - logHR}{logHR}\right) \times 100 \tag{4.13}$$

Collins et al. (2001) highlights the importance of measuring a multitude of performance measures in addition to bias, as results may deviate under different criterion. Though the expectation of simulated estimates is often of primary interest, Burton et al. (2006) discusses the trade-off between assessing the amount of bias and the variability induced with different modelling methods. In order to assess the accuracy, or variability, of the estimates, Burton et al. (2006) suggests the use of the mean square error (MSE) as it incorporates both measures of bias and variability:

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}\left(\widehat{logHR} - logHR\right)^2 + \left(SE\left(\widehat{logHR}\right)\right)^2 \tag{4.14}$$

Coverage is defined as the proportion of times that the estimated confidence interval contains the true parameter value. For $Z_{1-\alpha/2}$ denoting the $1-(\alpha/2)$ quantile of the standard normal distribution, the coverage can be defined as:

$$\begin{gathered}\text{Coverage} = \text{Proportion of times the 100(1-}\alpha\text{)\% credible interval:} \\ \widehat{logHR}_i \pm Z_{1-\alpha/2}SE(\widehat{logHR}_i) \\ \text{includes the true } logHR, \text{ for } i = 1, ..., m \text{ simulations.}\end{gathered} \tag{4.15}$$

Assuming normality of the samples, the coverage should approximately equal the nominal coverage rate. Thus, for 95% confidence intervals, 95% of samples should contain the true parameter estimate. Over-coverage - where the coverage rates are more than 95% -indicates that the variability is too large and the estimated effects are too conservative. On the other hand, under-coverage - where the coverage rates are less than 95% suggests that the variability in the parameter estimates are too small, and there is an over-confidence in the estimated effect (Burton et al., 2006).

4.5.3 Results

Results from the simulation studies are presented in Table 4.4. Generally, the average point estimates for the logHRs of active injection 2 and 3, relative to active injection 1 decreased, and uncertainty increased, with larger proportions of MAR data. This result indicates that for increasing levels of missing data structures, treatment effect may be exaggerated. However, the complete case analysis underestimated the treatment effect for active injection 3. This is likely to be a consequence of fewer patients being administered a third active injection and thus, in a complete case analysis (with 10, 25 and 50% of these observations MAR) the overall mean of the logHR shrinks towards no difference. For increasing levels of missingness, Bayesian predictive analyses increased bias from 2.67% to 23.33% for the estimated logHR of active injection 2, and from 0.27% to 2.97% for the estimated logHR of active injection 3. However, this is still a lot less than the complete case analysis which increased bias from 4% to 54% for the estimated logHR of active injection 2 and from -1.62% to -8.65% for the estimated logHR of active injection 3. On average, bias increased 2 fold for complete case analyses. The average model-based standard errors, decreased with Bayesian predictive methods. However, the MSE were broadly similar across both the predicted and complete case analyses. It is evident that the 95% credible intervals of the parameter estimates appeared to yield coverage above 95%, with the exception of complete analyses on larger proportions of missing data (25 and 50% MAR).

In situations where model choices were incorrectly specified, bias in the estimates were in the order of 20-30%. Overall, treatment effects for injection 2 were underestimated and treatment effects for injection 3 were over-estimated. The MSE was sufficiently larger than that of correctly specified models, and the coverage of parameter estimates were poor with coverage ranging from 0.72 to 0.88. Notably for active injection 3, coverage improved, and percentage bias marginally decreased with increased proportions of missing data.

TABLE 4.4: Simulation results

Model		***logHR* injection 2 = -0.15**						***logHR* injection 3 = -1.11**					
MAR	**Method**	**E($\widehat{logHR}$)**	**SE($\widehat{logHR}$)**	**Bias**	**%Bias**	**MSE**	**Coverage**	**E($\widehat{logHR}$)**	**SE($\widehat{logHR}$)**	**Bias**	**%Bias**	**MSE**	**Coverage**
10%	Predicted	-0.15	0.28	-0.004	2.67	0.18	1	-1.11	0.375	-0.003	0.27	5.12	1
25%	Predicted	-0.16	0.284	-0.014	9.33	0.19	1	-1.12	0.383	-0.014	1.26	5.20	1
50%	Predicted	-0.19	0.291	-0.035	23.33	0.21	1	-1.14	0.397	-0.033	2.97	5.26	1
10%	Complete case	-0.16	0.288	-0.006	4.00	0.13	1	-1.09	0.38	0.018	-1.62	4.94	0.98
25%	Complete case	-0.18	0.307	-0.028	18.67	0.16	1	-1.07	0.397	0.044	-3.96	4.84	0.91
50%	Complete case	-0.23	0.351	-0.081	54	0.22	1	-1.01	0.438	0.096	-8.65	4.66	0.75
		***logHR* injection 2 = -0.56**						***logHR* injection 3 = -0.68**					
10%	Model misspecification	-0.46	0.12	0.10	-17.9	1.07	0.87	-0.9	0.18	-0.22	32.35	2.56	0.72
25%	Model misspecification	-0.46	0.13	0.104	-18.6	1.07	0.87	-0.89	0.19	-0.215	31.62	2.55	0.75
50%	Model misspecification	-0.45	0.15	0.11	-19.64	1.06	0.88	-0.87	0.22	-0.193	28.38	2.50	0.82

E($\widehat{logHR}$) denotes the average estimated log hazard ratio and SE($\widehat{logHR}$) denotes the average model-based standard error. Bias denotes the bias in the point estimate and %Bias denotes the percentage bias in the point estimates. MSE denotes the mean square error, and coverage denotes the proportion of nominal 95 per cent confidence intervals that cover the true value.

4.6 Discussion

The novel use of a Bayesian framework has proven to be valuable for modelling complex clinical scenarios. The stochastic nature of a Bayesian approach allows the user to avoid strong censoring assumptions, and rather, allows computation of predictive posterior distributions in which to sample unreported event times. Recent developments by Demiris et al. (2015) in the freely available WinBUGS development software, have made routine fitting of flexible parametric models accessible. Consequently, survival data with increasingly complex data structures can be adequately modelled in a Bayesian setting. In the case of the RELAX trial, repeated injections of onaBoNT-A appeared to have a similar effect in patients with refractory DO. The average time to symptom return for active injections 1, 2 and 3 were 0.11(95%CrI: 0.07, 0.17), 0.19(95%CrI: 0.13, 0.27), and 0.35(95%CrI: 0.20, 0.55) years, respectively. This data suggests that either repeat injections of onaBoNT-A had a slow cumulative effect, or there is a potential selection process whereby only those patients who observed treatment benefit returned for further treatment. Indeed, participants who opt for re-injection appeared to have better symptom profiles at the preceding 6 week follow up compared to the entire cohort. This would suggest that patients with less severe symptoms, or those who obtain a greater treatment benefit from earlier injections, chose further treatment. To my knowledge, this finding has not been reported in the OAB and DO literature and represents a potential selection effect for each subsequent treatment.

Moreover, patient-reported duration of effect appeared to be influenced by initial treatment randomization. Patients randomized to placebo had a considerably decreased rate of symptom return for their first onaBoNT-A injection (received in the open label extension) compared to patients initially randomized to onaBoNT-A (who received their first active drug in a blinded fashion). This result may represent an extended placebo effect, which has not been noted before in trials of interventions for DO and OAB. It is known from migraine research that the placebo effect in randomized studies of onaBoNT-A treatment can remain at a

steady rate for up to 6 months (Mathew et al., 2005; Silberstein et al., 2005) and that the placebo effect is greater for more invasive treatments (Diener, 2010). However, in the current onaBoNT-A literature, there is no evidence to suggest that patients initially randomised to placebo, report greater treatment benefit upon administration of active treatment. Patients who received active injection initially, subsequently reported a greater duration of effect with the second active injection during the extension phase. Thus, it would seem that both groups (those randomized to both active and placebo injections) reported greater efficacy for the subsequent injection received during the extension study. This over-reporting by both groups suggests that open label extension studies following randomization may be biased toward more positive outcomes compared to the true (randomized and blinded) effects. This observation has wide implications in a HTA setting as nearly all drug studies for OAB and DO have a pooled open-label extension included to generate additional data in support of licensing, reimbursement and product use.

In a desirable analysis, potential selection effects and confounding factors, such as initial treatment allocation, would be accounted for by incorporating covariates in the log-hazard regression model. In situations with few known events, incorporating additional covariates can lead to computational difficulties. As a general rule of thumb, at least 10 events are required per parameter to be estimated in the model (Concato et al., 1995; Peduzzi et al., 1995, 1996) and, for prediction purposes, rules requiring 20 or more events may be more appropriate (Harrell et al., 1996). Data with fewer than 10 events per parameter can often run in to problems with parameter estimation and model convergence. A limitation of this study is that, in the case of the RELAX trial, there were few known events relative to the number of parameters to be estimated in the model; and thus incorporating covariates such as randomised treatment allocation, patient severity, and potential interactions with active injection, proved to be problematic. In order to adjust for potential confounding factors, more data documenting patient failure times would

be required.

The number of known events in the dataset relative to the number of parameters to be estimated in the model was a recurring problem for models incorporating covariate effects on the shape parameters (Model 2). The proportion of events relative to the number of parameters to be estimated in the models were considerably low, this was especially true for active injection 3. To ensure that the lack of events was the reason for computational difficulties, additional, hypothetical events were added to the dataset in an exploratory analysis. Increasing the number of events, in the form of a sensitivity analysis, considerably improved computation, parameter estimation, and model convergence; thus suggesting that more information regarding patient events would be necessary to allow for a more sophisticated model choice.

The impact of missing data on model performance was further investigated through a series of simulation studies. Data based on the RELAX trial was used to predict survival times of patients with repeated events by drawing samples from a Weibull distribution. Simulating data from complex correlation structures can often be difficult, so a Bayesian approach was adopted to simulate patient event times using sampling techniques from predictive posterior distributions. Model performance was assessed on 10%, 25% and 50% of interval censored event times. Overall, the proposed models performed well in terms of bias, accuracy and coverage. As expected, bias increased, and accuracy decreased with increasing proportions of missing data. However, using a Bayesian predictive analysis method reduced bias by approximately two fold compared to a complete case analysis. Generally, coverage appeared to be too high for all estimates of the logHRs for Bayesian predictive models, and for estimates of injection 2 for complete case methods. Modelling both the predicted and complete case analysis in a Bayesian framework allows for uncertainty in all of the parameter estimates, where plausibly vague prior distributions were assigned to all unknown and nuisance parameters. Incorporating this additional uncertainty will potentially increase the 95% credible intervals of

the estimates compared to the true value. Therefore, increased coverage would be expected to a certain extent. To improve coverage of the estimates, informative prior distributions could be incorporated to reduce the variability in sampled parameter estimates. The complete case analysis (with the exception of 10% MAR) showed under-coverage of estimates of the logHR for injection 3. This is likely to be due, in part, to the increased bias in the complete case analysis.

In principle, a limitation of all simulation studies is that model performance is likely to be associated with the chosen data generating mechanism. Assessing model performance based on data simulated from the same underlying distributional assumptions are likely to yield affirmative results. Consequently, impact of model misspecification was further assessed. Bias, accuracy, and coverage were assessed in situations where simplistic survival models - in this case, Weibull models - were chosen to represent time-to-event data with more complex underlying hazards generated from a 2-component poly-Weibull. Overall, Weibull models appeared to have reduced coverage in the parameter estimates with approximately a 20-30% observed bias. It is therefore important that model choice for Bayesian prediction modelling is selected with care.

Other parametric (Banerjee and Carlin, 2004), non-parametric (Calle and Gómez, 2001; Komárek et al., 2005) and semi-parametric models (Sinha et al., 1999; Jara et al., 2010; Hanson and Johnson, 2012) have been proposed to model interval censored data in a Bayesian framework. To my knowledge, this is the first study to adopt a fully Bayesian framework to model interval censored data using poly-Weibull models. However, in dental research, Wong et al. (2005) used a similar approach to model interval censored data in patients with multiple failure times using a Weibull model. With an increasing need to assess the time to recurrent events in chronic medical conditions, and the difficulties faced with intermittent follow-up, the novel use of a flexible Bayesian framework would appear to be promising. It is envisaged that a future application of this work could be used

to estimate disease recurrence where information regarding the event interval is obtained from hospitalization data.

4.7 Chapter summary

This chapter demonstrated a novel application of Bayesian methodology to evaluate repeat treatment in patients with interval censored data. Use of a Bayesian approach provided a flexible framework with which to model complex clinical scenarios, and recent developments by Demiris et al. (2015) permitted the use of flexible survival distributions. Motivated by the RELAX trial, application of Bayesian flexible parametric frailty models found that there may be a small cumulative effect of onaBoNT-A injections in patients with refractory DO, but this difference was not of clinical or patient importance. However, a potential selection effect and extended placebo-effect should be noted in open-label extension studies of patients with OAB and DO. Simulation studies found that Bayesian prediction models generally perform well with up to 50% of interval censored data, but care should be taken when selecting an appropriate distributional form.

Chapter 5

Systematic Review of Clinical Trials in Overactive Bladder

5.1 Chapter overview

To bridge the gap between clinical trials published in medical journals and the integration of scientific knowledge to clinical practice, data from *all* relevant trials must be acquired. In a regulatory setting, governing bodies such as NICE in the UK will examine all pertinent evidence to inform healthcare decision making. This process requires two steps; the first step involves data acquisition of all relevant trials of healthcare interventions. This is often obtained through a systematic literature review of the available evidence. The second step involves synthesising the data in a coherent analysis. This chapter describes the first of these two steps, and presents data from a systematic literature review of clinical trials of healthcare interventions for OAB and DO. The chapter begins by outlining the search criteria and methods for trial identification, data extraction and data manipulation, before describing study characteristics, and quality assessment of all eligible studies. Data from clinical trials identified and extracted in this chapter will be used throughout the remainder of this thesis.

5.2 Searching and identification methods

To obtain a comparable patient population – that is, a population with a similar distribution of potential treatment effect modifiers, stringent inclusion criteria were designed to include community dwelling adults with OAB symptoms, idiopathic DO, urge incontinence, and mixed urinary incontinence, where the predominant cause of incontinence was urgency. Individuals with neurogenic disorders, pregnancy related incontinence, stress incontinence, benign prostatic hyperplasia, bladder outlet obstruction, lower urinary tract symptoms, and mentally impaired individuals were excluded from the analysis due to differences in the underlying aetiology.

Studies were screened for randomised controlled trials (RCT) in adults with OAB symptoms. The process of randomisation prevents systematic differences in the baseline characteristics of trial participants in comparative groups. With a sufficiently large sample size, and successful random allocation of participants to intervention groups, the impact of potential known and unknown confounders on the study specific treatment effect estimates will be limited. Study selection was restricted to RCTs only to ensure that the highest quality of evidence was included and randomisation was maintained throughout the analysis. NICE have formulated guidance on methods for technology appraisals within the context of HTA in the UK, and state that:

> "In all cases when evidence is combined using adjusted indirect comparisons or network meta-analysis frameworks, trial randomisation must be preserved, that is, it is not acceptable to compare results from single treatment arms from different randomised trials. If this type of comparison is presented, the data will be treated as observational in nature and associated with increased uncertainty" (National Institute for Health and Care Excellence, 2013a, p.40).

Empirical evidence suggests that, non-randomised studies produce over inflated intervention effects, on average, than that of RCTs (Ioannidis et al., 2001). Methodological advancements have, however, allowed for the incorporation of non-randomised studies in evidence synthesis frameworks by adjusting for potential biases (Prevost et al., 2000). Such methodologies are still contentious in the current literature (Fleurence et al., 2010; Marko and Weil, 2010; Mullins and Sanchez, 2011; Goulart et al., 2014). In this thesis, inclusion of non-randomised data was beyond the scope of the research objective outlined in Chapter 1, which aimed to aid decision makers such as NICE in the UK using the highest level of study quality. Thus, as per NICE guidance, study inclusion was restricted to RCTs comparing two or more interventions, including placebo.

In keeping with NICE guidance, all minimally invasive interventions for the conservative management of OAB were included in the systematic literature review and evidence synthesis. NICE state that:

> "Ideally, the network meta-analysis should contain all treatments that have been identified either as an intervention or as appropriate comparators in the scope. Therefore, trials that compare at least 2 of the relevant (intervention or comparator) treatments should be incorporated, even if the trial includes comparators that are not relevant to the decision problem" (National Institute for Health and Care Excellence, 2013a, p.39),

Major surgical interventions such as augmentation cystoplasty and urinary diversion were excluded from the review on the basis that patients requiring highly invasive surgeries such as these, will only be recommended surgical treatment in the most extreme cases, and these patients will often require life-long follow-up. Thus, invasive surgeries are not considered appropriate comparators for the management of a more general OAB patient population, and are often not compared with other less invasive interventions.

Primary outcomes were mean change from baseline in urinary incontinence, voiding, urgency, and nocturia episodes per 24 hours. If mean change from baseline was not reported, it was calculated from the difference between baseline and follow-up values. Weekly mean episodes were averaged per 24 hours. Secondary outcomes included safety and tolerability measures including the number of patients experiencing adverse events, the number of patients who discontinue due to adverse events, and the number of patients who discontinue due to a lack of efficacy. In the pre-study protocol, trials were included in the systematic review if they reported quality of life outcomes. However, these outcomes were not analysed in an evidence synthesis framework in this thesis due to vast differences in study reporting and a varied use of quality of life measures.

Search strategies were developed using the Cochrane Collaboration's randomised controlled trials filter (Higgins et al., 2008) in combination with search terms covering all terminologies for OAB and DO. Key search terms included all commonly used names of interventions for overactive bladder, as well as "randomised controlled trial", "urinary incontinence", "overactive bladder", "detrusor overactivity" and their synonyms. A full search strategy is given in Appendix B.1. There were no restrictions on date, publication status, or language. Search engines included Medline, EMBASE, and Clinicaltrials.gov. Relevant references of eligible studies as well as published literature reviews investigating OAB management were searched through to 5th January 2016. Following trial identification, titles and abstracts were screened for eligibility. Full texts were obtained for all potentially relevant articles. Expert opinion on relevant trials were sought from Professor Douglas Tincello, a clinical specialist in Urogynaecology. Two reviewers independently assessed all studies for eligibility, and disagreement was resolved by joint review and arbitration by Professor Douglas Tincello if necessary. The systematic review protocol was registered with the International Prospective Register of Ongoing Systematic Reviews (PROSPERO) database (CRD42015024002) (Booth et al., 2011).

5.3 Data extraction and estimation methods

Trial information, patient characteristics, and all relevant outcomes were extracted from trial reports. Where possible, data were extracted at 12-weeks following treatment. If data were not reported at 12 weeks, the nearest, fully reported follow-up was taken. Twelve week follow-up was chosen for the main analysis as it is considered a sufficient length of time to show improvement in OAB symptoms (Geoffrion et al., 2012), and it is the most commonly reported trial follow-up time. Data extraction was independently performed and cross-validated for 10% of RCTs.

5.3.1 Continuous data

For trials that reported outcomes in a subgroup of the patient population, a pooled average (μ) and standard deviation (sd) were calculated using Equation (5.1) (Higgins et al., 2008). This might be the case, for example, if a study reported the sample size (n), mean (θ) and standard deviation (σ) separately for males and females in each of the intervention groups.

$$\begin{aligned} \mu &= \frac{n_1\theta_1 + n_2\theta_2}{n_1 + n_2}, \\ sd &= \sqrt{\frac{(n_1 - 1)\,\sigma_1^2 + (n_2 - 1)\,\sigma_2^2 + \frac{n_1 n_2}{n_1 + n_2}\,(\theta_1^2 + \theta_2^2 - 2\theta_1\theta_2)}{n_1 + n_2 - 1}} \end{aligned} \tag{5.1}$$

It was important to capture uncertainty in the trial effect estimates and incorporate this uncertainty in the evidence synthesis framework. However, trials often reported uncertainty in various ways using standard deviations (SD), standard

errors (SE), confidence intervals (CI), ranges, interquartile ranges (IQR) and p-values. Missing standard errors (SE) were calculated from standard deviations (SD) using Equation (5.2):

$$SE = \frac{SD}{\sqrt{n}} \tag{5.2}$$

Where the assumption of normality held, SEs were calculated from CIs, ranges, IQRs, and p-values. Standard errors were calculated from 95% CIs using a critical Z-value of 1.96 as described in Equation (5.3) (Higgins et al., 2008). Standard errors were calculated from ranges assuming an approximate 99.9% CI with corresponding critical Z-value of 3.291 (Equation (5.4)), and similarly from IQRs assuming an approximate 50% CI with critical Z-value of 0.674 (Equation (5.5)):

$$SE = \frac{\text{Upper 95\% CI limit} - \text{Lower 95\% CI limit}}{2Z} \quad \text{where } Z = 1.96 \tag{5.3}$$

$$SE = \frac{\text{Upper range limit} - \text{Lower range limit}}{2Z} \quad \text{where } Z = 3.291 \tag{5.4}$$

$$SE = \frac{\text{Upper IQR limit} - \text{Lower IQR limit}}{2Z} \quad \text{where } Z = 0.674 \tag{5.5}$$

Standard errors were calculated from p-values using Equation (5.6) where MD is the mean difference between treatment effect estimates of two comparators and t is calculated from the t-distribution with appropriate degrees of freedom (Higgins et al., 2008).

$$SE = \frac{MD}{t} \tag{5.6}$$

If information regarding uncertainty was not available or the assumption of normality did not hold, standard errors were estimated using the correlation between change from baseline, baseline, and follow-up, where reported (Abrams et al., 2005). This methodology is further described in Chapter 6.

5.3.2 Binary data

Sparse outcomes are predominantly encountered with safety and tolerability measures, where zero events can often occur in one or more treatment arms. In situations where zero events are encountered, calculation of the log-odds ratio will include division by 0. Consequently, the odds ratio (OR), together with it's variance, will remain undefined. One approach is to apply a continuity correction, where a small positive value (typically 0.5) will be added to the zero event arm (Higgins et al., 2008). A Bayesian MCMC approach can be used without modification to the zero event arm as it uses a binomial likelihood. However, in situations where there are low event data or zero events in both treatment arms, application of vague prior distributions can be more informative than intended (Lambert et al., 2005). Thus, for trials with zero events in both treatment arms a continuity correction of 0.5 were applied (Friedrich et al., 2007).

5.3.3 Predictors of response

As mentioned in Section 3.4.3.1, meta-analysis models may be subject to unknown confounding factors, such as differences in patient characteristics, that need to be explored and accounted for in any evidence synthesis model. It is, therefore, crucial

that all potential confounding factors are identified and collected during the data extraction phase of a systematic review. In order to identify potential factors that influence response to treatment, individual patient data obtained from the RELAX trial (Tincello et al., 2012) (further described in Section 4.2.1) were analysed using univariate and multivariate logistic regression models. Patient characteristics, demographic factors and baseline clinical covariates were evaluated. Univariate and multivariate regression models found that baseline symptom severity, smoking status, and participant age were all associated with response to treatment. Further details are described in the full manuscript published in *Neurourology and Urodynamics* (Owen et al., 2016b) given in Appendix F.2. In addition to gaining expert opinion from Professor Douglas Tincello, patient characteristics including participant age, BMI, proportion of females, and baseline symptom severity were extracted. Smoking status of participants were often not reported in the original trials and thus were not able to be extracted for analysis.

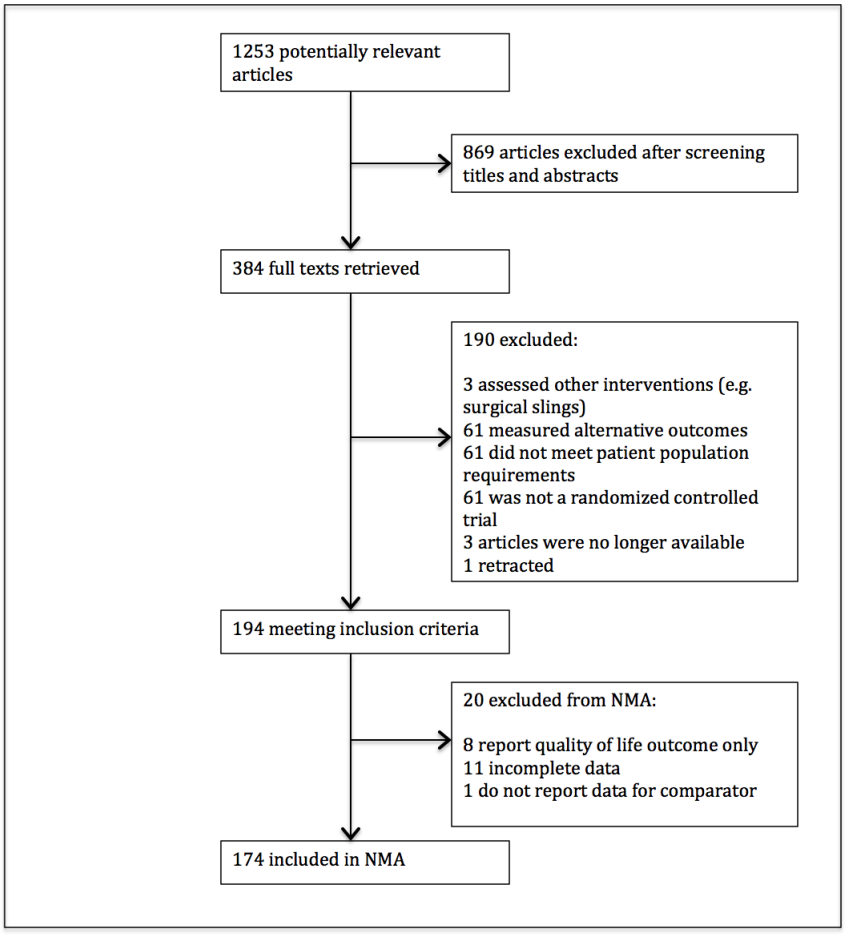
5.4 Study characteristics

Figure 5.1 shows the flow diagram for study inclusion. 1253 articles were identified after initial screening. Following review of titles and abstracts, full texts were obtained for 384 potentially relevant studies. 194 trials met the inclusion criteria$^{A1-A194}$ and 174 were eligible for analysis$^{A1-A174}$. Study references $A1 - A194$ are given in Table B.1 of Appendix B. Table 5.1 illustrates the characteristics of included studies. Eighty-two studies evaluated patients with a diagnosis of OAB, 55 with urge urinary incontinence (UUI), 16 with idiopathic detrusor overactivity (IDO), 14 with urge- and mixed incontinence, 6 with IDO and UUI, and 1 in mixed incontinence. A total of 75,355 patients were randomised to receive 1 of 140 different interventions. The most commonly reported follow-up was 12 weeks, with 60% of studies reporting 12 week follow-up in their primary analysis. Table 5.2 illustrates the patient characteristics of included studies. Mean age of participants was 57.5 (SD: 6.3) years and 84.9% (SD: 16.8) were female. On average, participants at baseline experienced 2.3 (SD: 1.7) urinary incontinence episodes,

11.4 (SD: 1.6) voiding episodes, 9.2 (SD: 2.7) urgency episodes and 2 (SD: 0.01) nocturia episodes per 24 hours.

Table 5.3 illustrates the severity of symptoms at baseline in participants receiving first-line, second-line, and third-line therapies as recommended by The National Institute for Health and Care Excellence (2013b). On average, participants receiving first-line therapies experienced 1.20 (SD: 0.68) incontinence episodes, 10.92 (SD: 2.73) voiding episodes, 5.99 (3.08) urgency episodes, and 2.22 (SD: 0.58) nocturia episodes at baseline. For participants receiving second-line therapies, patients experienced an average of 2.87 (SD: 1.35) incontinence episodes, 11.39 (SD: 1.94) voiding episodes, 6.69 (SD: 3.00) urgency episodes, and 2.19 (SD: 0.92) nocturia episodes. For participants receiving alternative second line therapies, usually because other (more commonly administered) second-line therapies are contraindicated, clinically ineffective, or deemed unacceptable in terms of side effects, experienced an average of 2.48 (SD: 0.51) incontinence episodes, 11.82 (SD: 0.56) voiding episodes, 5.64 (SD: 0.71) urgency episodes, and 2.17 (SD: 0.53) nocturia episodes at baseline. Participants receiving third-line therapies experienced a slightly higher rate of symptoms at baseline compared to participants receiving first- and second-line therapies, with an average of 4.57 (SD: 2.75) incontinence episodes, 12.98 (SD: 4.13) voiding episodes, 7.98 (SD: 0.98) urgency episodes, and 2.3 (SD: 0.17) nocturia episodes. On average, participants receiving interventions that are not currently recommended by NICE (National Institute for Health and Care Excellence, 2013b) experienced 2.83 (SD: 1.61) incontinence episodes, 11.31 (SD: 1.85) voiding episodes, 6.88 (SD: 2.30) urgency episodes, and 2.27 (SD: 1.17) nocturia episodes.

FIGURE 5.1: Flow diagram of studies identified in the systematic review



One third of studies compared interventions solely with placebo or control. The majority of studies (87%) identified in the systematic review assessed pharmacotherapies. Overall, 151 studies assessed oral drug therapies, 23 studies assessed multicomponent therapies, 16 studies assessed physical therapies, 12 assessed minimally invasive surgical management, 4 assessed behavioural therapy, and 4 assessed bladder training. This finding is largely driven by the vast number of industry led publications in the field of OAB, and the popularity of pharmacotherapies with patients and clinicians.

Table 5.1: Study characteristics

Reference	First author	Year	Control	Interventions	No. randomised	Design	Location	Allocation concealment	Diagnosis	Duration of treatment (wks)	Follow-up (wks)	No. of arms
A1	Abdelbary	2015	.	Pelvic floor electrical stimulation, vaginal estrogen, electrical stimulation + estrogen	315	Parallel	Egypt	Open Label	OAB	6	12	3
A2	Abrams	1998	Placebo	Tolterodine 2mg 2xdaily, Oxybutynin 5mg 3xdaily	293	Parallel	UK, ROI, Sweden	Double-blind	IDO	12	12	3
A3	Anderson	1993	Placebo	Terodiline 25mg 2xdaily	98	Parallel	USA	Double-blind	UUI + Mixed	4	12	2
A4	Anderson	1999	.	Oxybutynin ER (5-30mg), Oxybutynin IR 5mg (1-4xdaily)	105	Parallel	USA	Double-blind	UUI + Mixed	Variable	Variable (EOT)	2
A5	Appell	2001	.	Oxybutynin chloride ER 10mg 1xdaily, Tolterodine Tartrate 2mg 2xdaily	378	Parallel	USA	Double-blind	UUI	12	12	2
A6	Barkin	2004	.	Oxybutynin ER 15mg 1xdaily, Oxybutynin IR 5mg 3xdaily	125	Parallel	Canada	Double-blind	UUI	6	6	2
A7	Batista	2015	.	Mirabegron 50mg 1xdaily, Solifenacin 5mg 1xdaily	1887	Parallel	International	Double blind	OAB	12	6	2
A8	Bent	2008	Placebo	Duloxetine 40mg 2xdaily	156	Parallel	Canada, UK + USA	Double-blind	UUI	8	8	2
A9	Burgio	2008	.	Tolterodine tartrate ER 4mg 1xdaily, Tolterodine tartrate ER 4mg 1xdaily + Behaviour therapy (Education + PFE + Fluid management)	307	Parallel	USA	Open Label	UUI + Mixed	10	10	2
A10	Burgio	2010	.	Oxybutynin chloride ER 5mg-30mg daily, Oxybutynin chloride ER 5mg-30mg + Behaviour therapy (PFE, Urge supression)	64	Parallel	USA	Open Label	UUI + Mixed	8	8	2
A11	Burgio	2011	.	Behaviour therapy (PFE, Urge supression + delayed voiding), Oxybutynin ER 5-30mg/day	143	Parallel	USA	Single-blind	OAB	8	8	2
A12	But	2012	.	Solifenacin 5mg 1xdaily, Darifenacin 7.5mg 1xdaily	77	Parallel	Slovenia	Open Label	OAB	12	12	2
A13	Cardozo	2004	Placebo	Solifenacin Succinate 5mg 1xdaily, Solifenacin Succinate 10mg 1xdaily	911	Parallel	Europe	Double-blind	OAB	12	12	3
A14	Cardozo	2005	Placebo	Darifenacin 30mg 1xdaily	72	Parallel	UK	Double-blind	OAB	2	2	2
A15	Cardozo	2008	Placebo	Solifenacin (5mg or 10mg (wk8) 1xdaily)	865	Parallel	Europe	Double-blind	OAB	16	12	2
A16	Cartwright	2011	Placebo	Oxybutynin transdermal 3.9mg/day	96	Parallel	UK	Double-blind	OAB	4	12	2
A17	Chancellor	2000	Placebo	Tolterodine 2mg 2xdaily	1022	Parallel	Europe, N America + Australia	Double-blind	UUI	12	12	2
A18	Chancellor	2008	.	Darifenacin 7.5mg (up to 15mg) 1xdaily, Darifenacin 7.5mg (up to 15mg) + Behaviour management programme (diet modification/ PFE/ time voiding)	395	Parallel	USA	Open Label	OAB	12	12	2
A19	Chapple	2004	Placebo	Solifenacin 2.5mg, 5mg, 10mg, 20mg 1xdaily, Tolterodine IR 2mg 2xdaily	225	Parallel	Europe	Double-blind	IDO	4	4	6
A20	Chapple	2004	Placebo	Solifenacin 5mg, 10mg 1xdaily, Tolterodine 2mg 2xdaily	1081	Parallel	International	Double-blind	OAB	12	12	4
A21	Chapple	2005	.	Solifenacin 5mg (up to 10mg) 1xdaily	1200	Parallel	Europe	Double-blind	OAB	12	12	2
A22	Chapple	2006	Placebo	ZD0947IL 25mg/day (ATP-Sensitive Potassium Channel Opener)	191	Parallel	UK	Double-blind	OAB	12	12	2
A23	Chapple	2007	Placebo	Fesoterodine 4mg, Fesoterodine 8mg, Tolterodine ER 4mg	1135	Parallel	International	Double-blind	OAB	12	12	4
A24	Chapple	2013	Placebo	OnabotulinumtoxinA 100U (avoiding the trigone)	548	Parallel	Europe + USA	Double-blind	UUI	NA	12	2
A25	Chapple	2013	Placebo	Mirabegron 100mg, 150mg 2xdaily, Tolterodine ER 4mg 1xdaily	262	Parallel	International	Double-blind	OAB	4	4	4
A26	Chapple	2013	Placebo	Mirabegron oral controlled absorption system 25, 50, 100, 200mg, Tolterodine ER 4mg 1xdaily	928	Parallel	International	Double-blind	OAB	12	12	6
A27	Chapple	2013	.	Mirabegron 50mg, 100mg, Tolterodine ER 4mg 1xdaily	2444	Parallel	International	Double-blind	OAB	52	12	3
A28	Chapple	2014	Placebo	Fesoterodine 4mg, 8mg 1xdaily	2012	Parallel	International	Double-blind	UUI	12	12	3
A29	Chapple	2014	Placebo	ONO-8539 30mg, 100mg, 300mg 2xdaily, Tolterodine 4mg 1xdaily	435	Parallel	Europe	Double-blind	OAB	12	12	5
A30	Choo	2008	.	Solifenacin succinate 5mg 1xdaily, solifenacin succinate 10mg 1xdaily, tolterodine IR 2mg 2xdaily	357	Parallel	Korea	Double-blind	OAB	12	12	3
A31	Chu	2009	Placebo	Solifenacin 10mg 1xdaily	672	Parallel	USA	Double-blind	UUI + Mixed	12	12	2
A32	Chuang	2014	Saline	Liposome encapsulated onabotulinumtoxinA 200U (80mg sphingomyelin liposomes)	62	Parallel	Taiwan/USA (not stated)	Double blind	OAB	12	12	2
A33	Davila	2001	.	Oxybutynin 1.3mg transdermal 1xdaily + Placebo oral capsule, Oxybutynin IR 2.5mg titrated (half of 5mg) + Placebo patch	76	Parallel	USA	Double-blind	UUI	6	6	2
A34	Digesu	2012	Placebo	Elocalcitol 150mg, elocalcitol 75mg	257	Parallel	Europe	Double-blind	OAB	4	4	3
A35	Diokno	2003	.	Oxybutynin ER 10mg 1xdaily, Tolterodine ER 4mg 1xdaily	790	Parallel	USA	Double-blind	UUI	12	12	2
A36	Dmochowski	2002	Placebo	Oxybutynin transdermal 1.3mg, 2.6mg, 3.9mg 2xweekly	520	Parallel	USA	Double-blind	UUI + Mixed	12	12	4
A37	Dmochowski	2003	Placebo	Oxybutynin transdermal 3.9mg/day, Tolterodine ER 4mg 1xdaily	361	Parallel	USA	Double-blind	OAB	12	12	3
A38	Dmochowski	2008	Placebo	Trospium 60mg 1xdaily	564	Parallel	USA	Double-blind	UUI	12	12	2
A39	Dmochowski	2010	Placebo	OnabotulinumtoxinA 50u, 100u, 150u, 200u, 300u avoiding trigone and dome	313	Parallel	International	Double-blind	UUI	NA	12	6
A40	Dmochowski	2010	Placebo	Fesoterodine 4mg (up to 8mg) 1xdaily	896	Parallel	USA	Double-blind	OAB	12	12	2
A41	Dmochowski	2014	Placebo	Tolterodine IR 2mg + Pilocarpine ER 9mg, Tolterodine IR 2mg 2xdaily	138	Cross-over	Austrailia, NZ and South Korea	Double-blind	UUI + Mixed	12	12	3
A42	Drutz	1999	Placebo	Tolterodine 2mg 2xdaily, Oxybutynin 5mg 3xdaily	277	Parallel	USA + Canada	Double-blind	IDO	12	12	3
A43	Dubeau	2014	Placebo	Fesoterodine 4mg (up to 8mg) 1xdaily	562	Parallel	USA	Double-blind	UUI	12	12	2
A44	Enzelsberger	1991	Placebo	Estradiol intravaginally 1mg, 3mg/day	40	Parallel	Austria	Open Label	UUI	3	4	3
A45	Enzelsberger	1991	.	Lidocaine Gel (2x6ml transurethrally), Emperonium bromide 3x200mg/day	30	Parallel	Austria	Open Label	UUI + Mixed	3	4	2
A46	Enzelsberger	1995	Placebo	Oxybutynin 20mg intravesically 1xdaily	52	Parallel	Austria	Blind	UUI	2	2	2
A47	Enzelsberger	1995	Placebo	Oxybutynin 20mg intravesically 1xdaily	39	Parallel	Austria	Double-blind	UUI	1.5	2	2
A48	Finazzi-Agro	2010	Control	Percutaneous tibial nerve stimulation	35	Parallel	Italy	Double-blind	IDO	4	4	2
A49	FitzGerald	2008	.	Tolterodine ER 4mg 1xdaily, Tolterodine ER 4mg + Behaviour (PFE + education)	307	Parallel	USA	Open Label	UUI	8	8	2
A50	Franzen	2010	.	Electrostimulation 20 min 5-10Hz 2xdaily, Tolterodine ER 4mg 1xdaily	72	Parallel	Sweden	Open Label	UUI	6	6	2
A51	Frenkl	2010	Placebo	Seropitant 0.25mg, 1mg, 4mg, Tolterodine ER 4mg 1xdaily	557	Parallel	International	Double-blind	UUI	8	8	5
A52	Fukuda	2013	.	Solifenacin 5mg 1xdaily, Propiverine 20mg 1xdaily	66	Parallel	Japan	Single-blind	OAB	8	8	2
A53	Giannitsas	2004	.	Oxybutynin 5mg 3xdaily, Tolterodine 2mg 2xdaily	128	Cross-over	Greece	Open Label	IDO	6	6	2
A54	Gittelman	2014	Placebo	Oxybutynin vaginal ring 4, 6mg 1xdaily	445	Parallel	USA + Canada	Double-blind	UUI	12	12	3
A55	Goldfischer	2015	Placebo	Oxybutynin 84mg gel, oxybutynin 56mg gel	626	Parallel	USA	Double blind	UUI + Mixed	12	12	3
A56	Gotoh	2011	Placebo	Propiverine 20mg 1xdaily	567	Parallel	Japan	Double-blind	OAB	12	12	2
A57	Haab	2014	Placebo	Netupitant 50mg, 100mg, 200mg 1xdaily	246	Parallel	Europe	Double-blind	OAB	8	8	4
A58	Halaska	2003	.	Trospium Chloride 20mg (2xdaily), Oxybutynin 5mg (2xdaily)	358	Parallel	Europe	Double-blind	IDO + UUI	52	2	2

TABLE 5.1: Study characteristics (cont.)

A59	Hassouna	2000	SMT	Subcutaneous Sacral Nerve Stimulation	51	Parallel	International	Open Label	IDO	24	24	2
A60	Herschorn	2004	.	Tolterodine, Tolterodine + Behaviour therapy	84	Parallel	Canada	Open Label	OAB	16	16	2
A61	Herschorn	2008	Placebo	Tolterodine ER 4mg 1xdaily	617	Parallel	International	Double-blind	OAB	12	12	2
A62	Herschorn	2010	.	Solifenacin 5mg 1xdaily, Oxybutynin IR 5mg 3xdaily	132	Parallel	Canada	Double-blind	OAB	8	8	2
A63	Herschorn	2010	Placebo	Fesoterodine 8mg, Tolterodine ER 4mg 1xdaily	1712	Parallel	USA	Double-blind	UUI	12	12	3
A64	Herschorn	2013		Mirabegron 25mg, Mirabegron 50mg 1xdaily	1306	Parallel	Europe + N America	Double-blind	OAB	12	12	3
A65	Hill	2006	Placebo	Darifenacin ER 7.5mg, 15mg, 30mg 1xdaily	439	Parallel	International	Double-blind	UUI	12	12	4
A66	Ho	2010	.	Solifenacin 5mg 1xdaily, Tolterodine 4mg 1xdaily	75	Parallel	Taiwan	Open Label	OAB	12	12	2
A67	Holmes	1989	.	Oxybutynin 5mg 3xdaily, Propantheline Bromide 15mg 3xdaily	23	Cross-over	UK	Single-blind	IDO	4	4	2
A68	Homma	2003	Placebo	Tolterodine ER 4mg 1xdaily, Oxybutynin IR 3mg 3xdaily	608	Parallel	Japan + Korea	Double-blind	UUI	12	12	3
A69	Homma	2008	Placebo	Imidafenacin 0.05mg 2xdaily, Imidafenacin 0.1mg 2xdaily, Imidafenacin 0.25mg 2xdaily	401	Parallel	Japan	Double-blind	UUI	12	12	4
A70	Homma	2009	Placebo	Imidafenacin 0.1mg 2xdaily, Propiverine 20mg 1xdaily	781	Parallel	Japan	Double-blind	UUI	12	12	3
A71	Hsiao	2011	.	Solifenacin 5mg 1xdaily, Tolterodine ER 4mg 1xdaily	48	Parallel	Taiwan	Open Label	OAB	12	12	2
A72	Huang	2012	Placebo	Fesoterodine 4mg (up to 8mg) 1xdaily	645	Parallel	USA	Double-blind	UUI	12	12	2
A73	Huo	2013	.	Solifenacin Succinate 5mg 1xdaily, Naftopidil 25mg 1xdaily, Solifenacin Succinate 5mg 1xdaily + Naftopidil 25mg 1xdaily	67	Parallel	China	Single-blind	OAB	4	4	3
A74	Jabs	2013	Placebo	Botulinum toxin A 100u sparing trigone	21	Parallel	Canada	Double-blind	IDO + UUI	NA	4	2
A75	Jacquetin	2001	Placebo	Tolterodine 1mg 2xdaily, Tolterodine 2mg 2xdaily	251	Parallel	France + Belgium	Double-blind	OAB	4	4	3
A76	Johnson	2005	Placebo	Behaviour training (Education, PFE + Biofeedback), Oxybutynin IR 2.5mg (up to 15mg) 1xdaily	131	Parallel	USA	Double-blind (drugs)	UUI	8	8	3
A77	Junemann	2005	.	Propiverine 15mg 2xdaily, Tolterodine 2mg 2xdaily	201	Parallel	Europe	Double-blind	IDO	4	4	2
A78	Junemann	2006	Placebo	Propiverine hydrochloride IR 15mg 2xdaily, Propiverine hydrochloride ER 30mg 1xdaily	988	Parallel	International	Double-blind	UUI	4	4	3
A79	Kafri	2013	.	Tolterodine SR 4mg, Bladder training, PFE, BT+PFE+Behaviour education	164	Parallel	Israel	Single blind	OAB	12	12	4
A80	Kaplan	2011	Placebo	Tolterodine ER 4mg 1xdaily, Fesoterodine 4mg (up to 8mg) 1xdaily	2417	Parallel	International	Double-blind	UUI	12	12	3
A81	Kaplan	2014	Placebo	Fesoterodine 8mg 1xdaily (11 wks) fesoterodine 4mg 1xdaily (week1)	609	Parallel	International	Double blind	OAB	12	12	2
A82	Karram	2009	Placebo	Solifenacin 5mg (up to 10mg) 1xdaily	739	Parallel	USA	Double-blind	OAB	12	12	2
A83	Kaya	2011	.	Trospium Chloride 15mg 3xdaily (45mg total), Physiotherapy (Electrotherapy, PFE + BT), Trospium chloride 15mg 3xdaily + Physiotherapy	46	Parallel	Turkey	Open Label	IDO	8	8	3
A84	Khullar	2004	Placebo	Tolterodine ER 4mg 1xdaily	854	Parallel	Europe	Double-blind	Mixed UI	8	8	2
A85	Khullar	2013	Placebo	Mirabgeron 50mg, Mirabegron 100mg, Tolterodine ER 4mg 1xdaily	1978	Parallel	Europe + Austrailia	Double-blind	OAB	12	12	4
A86	Kosilov	2014	Placebo	Trospium 60mg/day + Solifenacin 20mg/day (cyclic), Trospium 30mg/day + Solifenacin 10mg/day (cyclic), Trospium 30mg/day + Solifenacin 10mg/day (continuous)	239	Parallel	Russia	Double blind	OAB	52	24	4
A87	Kuo	2011	.	OnabotulinumtoxinA 100IU bladder body sparing trigone, bladder body (75U) + trigone (25U), bladder base (50U) + trigone (50U)	105	Parallel	Taiwan (not stated)	Single-blind	IDO	NA	12	3
A88	Kuo	2014	Saline	OnabotulinumtoxinA 200IU	24	Parallel	Taiwan (not stated)	Double-blind	UUI	4	4	2
A89	Kuo	2015	Placebo	Mirabegron 50mg 1xdaily, Tolterodine 4mg 1xdaily	1126	Parallel	Taiwan/Korea/China/India	Double blind	OAB	12	12	3
A90	Kuo	2015	Placebo	Mirabegron 50mg 1xdaily, Tolterodine 4mg 1xdaily	248	Parallel	Taiwan	Double blind	OAB	12	12	3
A91	Kurz	1993	Placebo	Estriol 1mg intravesical	42	Parallel	Germany	Double-blind	UUI	3	3	2
A92	Lauti	2008	.	Oxybutynin 2.5mg (up to 3x5mg) 1xdaily, Bladder training (by a Physiotherapist), Oxybutynin 2.5mg + Bladder training	57	Parallel	New Zealand	Open Label	UUI	12	12	3
A93	Lee	2002	.	Tolterodine 2mg 2xdaily, Oxybutynin 5mg 2xdaily	228	Parallel	Korea	Double-blind	OAB	8	8	2
A94	Lee	2010	Placebo	Propiverine 20mg 1xdaily	264	Parallel	Korea	Double-blind	OAB	12	12	2
A95	Lee	2013	.	Imidafenacin 0.1mg (2xdaily), Fesoterodine 4mg (1xdaily)	206	Parallel	Korea	Double-blind	OAB	12	12	2
A96	Lehtoranta	2002	Placebo	Oxybutynin 5mg 3xdaily intravesically	9	Cross-over	Finland	Double-blind	IDO	2	2	2
A97	Madersbacher	1999	Placebo	Propiverine 15mg 3xdaily, Oxybutynin 5mg 2xdaily	366	Parallel	Europe	Double-blind	OAB	4	4	3
A98	Mak	2007	.	Reflexology, Foot massage (45min 1xdaily)	120	Parallel	China	Single-blind	OAB	3	3	2
A99	Malone-Lee	2001	Placebo	Tolterodine 1mg 2xdaily, Tolterodine 2mg 2xdaily	177	Parallel	UK, Ireland, France	Double-blind	OAB	4	4	3
A100	Malone-Lee	2001	.	Tolterodine 2mg 2xdaily, Oxybutynin (2.5mg -5mg) 2xdaily	379	Parallel	UK & ROI	Double-blind	OAB	10	10	2
A101	Malone-Lee	2009	Placebo	Tolterodine ER 4mg 1xdaily	308	Parallel	UK	Double-blind	OAB	12	12	2
A102	Marencak	2011	Placebo	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d, Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d, Pregabalin 150mg b.i.d, Tolterodine ER 4mg q.d	188	Cross-over	International	Double-blind	OAB	4	4	5
A103	Martinez-Garcia	2009	Placebo	Cizolirtine citrate 400mg 2xdaily, cizolirtine citrate 200mg 2xdaily	79	Parallel	Spain	Double-blind	UUI	12	12	3
A104	Mattiasson	2003	.	Tolterodine 2mg 2xdaily, Tolterodine 2mg 2xdaily +BT	501	Parallel	Sweden, Norway, Denmark	Open Label	OAB	24	24	2
A105	Mattiasson	2010	.	Solifenacin 5mg (up to 10mg) 1xdaily + BT, Solifenacin 5mg (up to 10mg) 1xdaily	644	Parallel	Europe + Austrailia	Open Label	OAB	16	16	2
A106	Mazur	1995	.	Propiverine hydrochloride 15, 30mg (15mg 2xdaily), 45mg (15mg 3xdaily), 60mg 1xdaily	185	Parallel	Germany	Open Label	UUI	3	3	4
A107	Meyhoff	1983	Placebo	Emepronium bromide 200mg, Flavoxate chloride 200mg 1xdaily	19	Cross-over	Denmark	Double-blind	UUI + Mixed	2	2	3
A108	Millard	1999	Placebo	Tolterodine 1mg, 2mg 2xdaily	316	Parallel	Austrailia	Double-blind	IDO	12	12	3
A109	Millard	2004	.	Tolterodine 2mg 2xdaily, Tolterodine 2mg 2xdaily + PFE	480	Parallel	Australia	Open Label	UUI	24	12	2
A110	Nitti	2007	Placebo	Fesoterodine 4mg, 8mg 1xdaily	836	Parallel	USA	Double-blind	OAB	12	12	3
A111	Nitti	2010	Placebo	Fesoterodine 4mg, 8mg, 12mg 1xdaily	99	Parallel	USA	Double-blind	IDO	8	8	4
A112	Nitti	2013	Placebo	Mirabegron 50mg, 100mg 1xdaily	1329	Parallel	USA + Canada	Double-blind	OAB	12	12	3
A113	Nitti	2013	Placebo	OnabotulinumtoxinA 100U	557	Parallel	USA + Canada	Double-blind	UUI	NA	12	2
A114	Norton	1994	Placebo	Terodiline 50mg (25mg 2xdaily)	93	Parallel	USA	Double-blind	UUI	4	4	2
A115	O'Reilly	2008	Sham Therapy	Trans-sacral stimulation of S3 and S4 sacral nerve 5-20Hz pulse width 1ms (Electromagnetic stimulation)	63	Parallel	Unknown	Double-blind	OAB	12	12	2
A116	Ohlstein	2012	Placebo	Solabegron 50mg 2xdaily, Solabegron 125mg 2xdaily	258	Parallel	International	Double-blind	UUI	8	8	3

Table 5.1: Study characteristics (cont.)

A117	Olmo	2013	.	Percutaneous nerve stimulation (neuromodulation) 20hz frequency, 320ms pulse, Electrostimulation of SP6	24	Parallel	Spain	Double-blind	UUI	12	12	2
A118	Oreskovic	2012	Placebo	Solifenacin 5mg 1xdaily	171	Parallel	Croatia, Slovenia	Double-blind	OAB	4	4	2
A119	Orri	2014	Placebo	Tolterodine ER 4mg 1xdaily	18	Parallel	USA	Double-blind	OAB	12	12	2
A120	Ouslander	1993	Placebo	Terodiline 25mg 2xdaily	98	Parallel	USA	Double-blind	UUI + Mixed	4	4	2
A121	Ozdedeli	2010	.	Trospium chloride 15mg 3xdaily, Physiotherapy electrotherapy, PFE, BT), Trospium chloride 15mg 3xdaily + Physiotherapy	46	Parallel	Turkey	Open Label	OAB	8	8	3
A122	Park	2014	.	Imidafenacin (0.1mg 2 x daily), Propiverine 20mg 1xdaily	162	Parallel	Korea	Double-blind	UUI	12	12	2
A123	Peters	2009	.	Percutaneous tibial nerve stimulation, Tolterodine tartrate ER 4mg 1xdaily	100	Parallel	USA	Open Label	OAB	12	12	2
A124	Peters	2010	Sham Therapy	Percutaneous tibial nerve stimulation 20Hz	220	Parallel	USA	Double-blind	OAB	12	12	2
A125	Preik	2004	.	Oxybutynin ER (5mg - 30mg), Oxybutynin IR (5mg - 20mg)	105	Parallel	Unknown	Double-blind	UUI + Mixed	12	12	2
A126	Preyer	2015	.	Percutaneous tibial nerve stimulation, Tolterodine IR 2mg b.i.d	36	Parallel	Austria, Germany	Open Label	OAB	12	12	2
A127	Rentzhog	1998	Placebo	Tolterodine 0.5, 1, 2 or 4mg 2xdaily	81	Parallel	UK & Sweden	Double-blind	OAB	2	2	5
A128	Rios	2007	Placebo	Resiniferatoxin 50nM	58	Parallel	Brazil (not stated)	Double-blind	IDO + UUI	4	4	2
A129	Rogers	2008	Placebo	Tolterodine ER 4mg 1xdaily	413	Parallel	USA	Double-blind	UUI	12	12	2
A130	Rudy	2006	Placebo	Trospium chloride 20mg (2xdaily)	658	Parallel	USA	Double-blind	OAB	12	12	2
A131	Rufford	2003	Placebo implant	Estradiol implants 25mg	40	Parallel	UK	Double-blind	OAB	Unclear	12	2
A132	Sahai	2007	Placebo	Boutlinum toxin A 200U (trigone sparing)	36	Parallel	UK	Double-blind	UUI + Mixed	NA	12	2
A133	Sancaktar	2010	.	Tolterodine 4mg 1xdaily, Tolterodine 4mg 1xdaily + Neurostimulation	50	Parallel	Turkey	Open Label	UUI	12	12	2
A134	Schmidt	1999	SMT	Sacral Nerve Stimulation	98	Parallel	International	Open Label	UUI	24	24	2
A135	Schreiner	2010	.	PFE (Kegel exercises -15contractions 3xday) + Bladder training + Electrical stimulation (10-Hz, 200ms 30mins 1xweek), PFE + Bladder training	52	Parallel	Brazil	Open Label	UUI	12	12	2
A136	Song	2006	.	Bladder training, Tolterodine 2mg 2xdaily, Bladder training + Tolterodine 2mg 2xdaily	139	Parallel	Korea (not stated)	Open Label	OAB	12	12	3
A137	Song	2015	Placebo	Tarafenacin 0.2mg, 0.4mg 1xdaily	235	Parallel	Korea	Double blind	OAB	12	12	3
A138	Soomro	2001	.	Oxybutynin 2.5mg 2xdaily (up to 5mg 3xdaily), Transcutaneous electrical nerve stimulation	43	Cross-over	UK	Open Label	IDO	6	6	2
A139	Staskin	2007	Placebo	Trospium Chloride 60mg 1xdaily	601	Parallel	USA	Double-blind	OAB	12	12	2
A140	Staskin	2009	Placebo	Oxybutynin chloride topical gel 1g 1xdaily	789	Parallel	USA	Double-blind	OAB	12	12	2
A141	Steers	2005	Placebo	Darifenacin ER 7.5mg (or 15mg) 1xdaily	395	Parallel	Canada + USA	Double-blind	OAB	12	12	2
A142	Steers	2007	Placebo	Duloxetine 80mg/day for 4weeks to 120mg/day for 8weeks	306	Parallel	Australia, Canada + USA	Double-blind	OAB	12	12	2
A143	Subak	2002	No instruction	Behaviour therapy 6 weekly 20min group instruction on bladder training	152	Parallel	USA (California)	Open Label	UUI + Mixed	6	6	2
A144	Swift	2003	Placebo	Tolterodine ER 4mg 1xdaily, Tolterodine IR 2mg 2xdaily	1235	Parallel	Europe, N America, Austrailia, NZ	Double-blind	UUI	12	12	3
A145	Tang	2014	.	Tolterodine ER 2mg 1xdaily, Tolterodine ER 2mg + intermittent percutaneous needle sacral nerve stimulation	240	Parallel	China	Open Label	OAB	12	12	2
A146	Tapp	1989	Placebo	Terodiline 50mg (25mg 2xdaily)	91	Parallel	Europe	Double-blind	IDO + UUI	8	8	2
A147	Tincello	2000	.	Oxybutynin 2.5mg 2xdaily (up to Oxybutynin 5mg 3xdaily), Oxybutynin 2.5mg 2xdaily (up to Oxybutynin 5mg 3xdaily) + Salivary stimulant pastilles	67	Parallel	UK	Open Label	IDO	8	8	2
A148	Tincello	2012	Placebo	OnabotulinumtoxinA 200IU	240	Parallel	UK	Double-blind	IDO	NA	12	2
A149	Tseng	2009	.	Tolterodine ER 2mg 2xdaily, Tolterodine ER 2mg 2xdaily + vaginal conjugated equine estrogen 0.625mg 2xweek	80	Parallel	Unkown	Open Label	OAB	12	12	2
A150	Ulshofer	2001	Placebo	Trospium Chloride 15mg 3xdaily	46	Parallel	Germany	Double-blind	UUI	4	4	2
A151	Van Kerrebroeck	2001	Placebo	Tolterodine ER 4mg 1xdaily, Tolterodine IR 2mg 2xdaily	1529	Parallel	Australasia, Europe and N America	Double-blind	UUI	12	12	3
A152	Van Kerrebroeck	2009	Placebo	Tolterodine ER 4mg 1xdaily	987	Parallel	Australasia, Europe and N America	Double-blind	UUI	12	12	2
A153	Vardy	2009	Placebo	Solifenacin 5mg 1xdaily (up to 15mg, 5mg ˆevery 4 weeks optional)	768	Parallel	USA	Double-blind	OAB	12	12	2
A154	Versi	2000	.	Oxybutynin ER 5mg (up to 20mg) q.d, Oxybutynin IR 5mg (up to 20mg) q.d	226	Parallel	USA	Double-blind	UUI	Unclear	4	2
A155	Visco	2012	.	Anticholinergic (Solifenacin 5mg >Solifenacin 10mg >Trospium XR 60mg 1xdaily) + saline injection, placebo + onabotulinumtoxinA 100U	249	Parallel	USA (not stated)	Double-blind	UUI	24	24	2
A156	Wagg	2013	Placebo	Fesoterodine 4mg 1xdaily	794	Parallel	International	Double-blind	OAB	12	12	2
A157	Wang	2006	Placebo	Electrical stimulation (10-Hz, 400-ms pulse width, 10/5 duty cycle), Oxybutynin 2.5mg 3xdaily	68	Parallel	Taiwan	Open Label	OAB	12	12	3
A158	Wang	2009	Placebo	Electrical stimulation, Oxybutynin 2.5mg 3xdaily	73	Parallel	Taiwan (not stated)	Open Label	OAB	12	12	3
A159	Weiss	2013	Placebo	Fesoterodine 4mg (up to 8mg) 1xdaily	963	Parallel	USA	Double-blind	OAB	12	12	2
A160	Yamaguchi	2007	Placebo	Solifenacin 5mg 1xdaily, Solifenacin 10mg 1xdaily, Propiverine 20mg 1xdaily	1593	Parallel	Japan	Double-blind	OAB	12	12	4
A161	Yamaguchi	2011	Placebo	Fesoterodine 4mg, 8mg 1xdaily	951	Parallel	Japan, Taiwan, Korea, Hong Kong	Double-blind	OAB	12	12	3
A162	Yamaguchi	2014	Placebo	Mirabegron 50mg 1xdaily, Tolterodine 4mg 1xdaily	1139	Parallel	Japan	Double-blind	OAB	12	12	3
A163	Yamaguchi	2014	Placebo	Oxybutynin patch 73.5mg (35cm2), Propiverine 20mg 1xdaily	1530	Parallel	Japan	Double-blind	OAB	12	12	3
A164	Yamaguchi	2014	Placebo	Mirabegron 25mg, 50mg, 100mg 1xdaily	842	Parallel	Japan	Double blind	OAB	12	12	4
A165	Yamanishi	2000	Sham Therapy	Electrical stimulation	68	Parallel	Japan (not stated)	Double-blind	IDO + UUI	4	4	2
A166	Yokoyama	2013	.	Imidafenacin 0.1mg 2xdaily, Solifenacin 5mg 1xdaily	109	Parallel	Japan	Open Label	OAB	52	52	2
A167	Yokoyama	2014	Placebo	Fesoterodine 4mg 1xdaily, Fesoterodine 8mg 1xdaily	555	Parallel	Asia	Double-blind	UUI	12	12	3
A168	Yoon	2003	Untreated control	Bladder training, PFE (Kegel exercises)	50	Parallel	South Korea (not stated)	Open Label	UUI	8	8	3
A169	Zat'ura	2010	Placebo	Cizolirtine citrate 800mg/d (2 x 400mg), Oxybutynin 15mg/d (3 x 5mg)	135	Parallel	Czech Republic	Double-blind	IDO + UUI	12	12	3
A170	Zellner	2009	.	Oxybutynin 7.5mg (2.5mg 3xday up to 5mg t.i.d), Trospium chloride 45mg (15mg 3xday up to 30mg t.i.d)	1659	Parallel	Germany	Double-blind	UUI	12	12	2
A171	Zimmern	2010	.	Tolterodine (mg not specified) + Fluid intake management	307	Parallel	USA	Open Label	UUI	10	10	2
A172	Zinner	2002	Placebo	Tolterodine ER 4mg	1015	Parallel	International	Double-blind	UUI	12	12	2
A173	Zinner	2005	Placebo	Darifenacin 15mg 1xdaily, Darifenacin 30mg 1xdaily, Oxybutynin 5mg 3xdaily	76	Cross-over	USA (not stated)	Double-blind	UUI	2	2	4
A174	Zinner	2006	Placebo	Darifenacin 15mg 1xdaily	445	Parallel	USA (not stated)	Double-blind	UUI	12	12	2

5.4.1 Trials not included in evidence synthesis methods

In total 20 studies identified in the systematic review were excluded from the evidence synthesis. Burgio et al. (2002) analysed different forms of behavioural therapy (behavioural therapy with biofeedback, behavioural therapy with verbal feedback, and behavioural therapy with self-help booklets) and found that different forms of behavioural therapy have comparable efficacy for the management of urge incontinence. Behavioural therapy is multifaceted and can often take many different approaches. As behavioural therapy was not defined in the majority of studies, all forms of behavioural therapy were considered as a single intervention in the evidence synthesis methods described in Chapters 6, 7, and 8; and thus by this definition, Burgio et al. (2002) did not present data on a comparator arm. Eight studies reported quality of life outcomes only (Homma and Kawabe (2004), Kelleher et al. (2002), Nascimento-Correia et al. (2012), Newman et al. (2010), Pleil et al. (2001), Rogers et al. (2009), Sahai et al. (2009) and Wang et al. (2004)).

Eleven studies did not provide data on effect estimates and/or uncertainty (Amaro et al. (2005), Bellette et al. (2009), Dowson et al. (2011), Karademir et al. (2005), Peters et al. (2005), Robinson et al. (2007), Souto et al. (2014), Zinner et al. (2004), Abrams et al. (2015), Kosilov et al. (2014) and Kosilov et al. (2015)). Amaro et al. (2005) investigated the effect of electrical stimulation on pelvic floor muscle strength and found that both active and sham treatment significantly improved the number of voiding episodes, though there was no statistically significant difference between interventions (mean change from baseline: -3.5 and -2.5 episodes, respectively. Estimates of uncertainty were not presented). Bellette et al. (2009) also evaluated electrical stimulation versus sham therapy and found improvements in voiding, urgency, and nocturia, but did not quantify the intervention effects. Dowson et al. (2011) presented data from an interim analysis of a larger trial that was prematurely terminated due to poor perceived patient benefit. For the 21 participants analysed, 10 were treated with onaBoNT-A and 11 were treated with placebo. There appeared to be no difference in urge urinary incontinence

(mean difference:-1, 95%CI: -3.7, 1.7), or voiding episodes (mean difference:-0.9, 95%CI:-4.9, 3.1) between interventions, but data on individual intervention effects was not reported. Karademir et al. (2005) compared neurostimulation versus neurostimulation combined with oxybutynin 5mg, and concluded that there was no apparent difference between the intervention groups for voiding (p-value=0.48) or urgency (p-value=0.43) episodes. Peters et al. (2005) evaluated pudendal nerve stimulation and SNS, and found that there were significant differences in voiding episodes from baseline for both interventions, but there was little difference between interventions. Robinson et al. (2007) investigated the efficacy of 4 different doses of tamsulosin (an α_1-adrenoceptor antagonist), 4mg of tolterodine extended release, and placebo, and found that the difference between tamsulosin, and tolterodine, versus placebo were not significantly different in reducing voiding episodes (p-value=0.189 and p-value=0.353, respectively). However, Karademir et al. (2005), Peters et al. (2005) and Robinson et al. (2007) did not provide data on treatment effect estimates. Souto et al. (2014) compared electrical stimulation, oxybutynin 10mg extended release, and their combination. Voiding episodes decreased from 12.7 to 8 for electrical stimulation, from 11 to 7.9 for oxybutynin 10mg extended release, and 11.2 to 7.6 for combination therapy. There appeared to be no significant difference between treatment groups (p-value=0.75) but data on the variability of change from baseline in individual interventions was not provided. Zinner et al. (2004) compared trospium 20mg twice daily with placebo. Trospium 20mg twice daily reduced incontinence episodes on average by 59% whilst placebo reduced incontinence by 44.2%. Urgency episodes were reduced on average by -2.30 and -1.08, and nocturia episodes were reduced by -0.47 and -0.29, respectively, but estimates of uncertainty were not provided. Abrams et al. (2015) compared different doses of mirabegron, solifenacin, and their combinations, with placebo in a 12-arm trial. All treatments demonstrated a reduction in incontinence and urgency episodes but none showed a statistically significant difference to placebo. For voiding episodes, statistically significant differences were observed for combinations of solifenacin 10mg + mirabegron 25mg, solifenacin 5mg + mirabegron

50mg and solifenacin 10mg + mirabegron 50mg, compared with placebo, but treatment effect estimates were not quantified. Kosilov et al. (2014) compared cyclic versus continuous administration of trospium 60mg + solifenacin 20mg, trospium 30mg + solifenacin 10mg, and placebo. There did not appear to be a statistically significant difference between cyclic and continuous therapy, but data were not available for intervention effects or their associated uncertainty. Kosilov et al. (2015) evaluated mirabegron 50mg, solifenacin 10mg, mirabegron 50mg + solifenacin 10mg, and placebo. Mirabegron 50mg reduced incontinence by -2.3 episodes on average. Solifenacin 10mg reduced incontinence and voiding episodes by -2.2 and -3.4 episodes on average. Combination therapy decreased incontinence and voiding episodes by an average of -3.8 and -4 episodes, respectively. However, estimates of variability were not provided for any of the outcomes.

TABLE 5.2: Patient characteristics

Reference	First author	Year	Mean age (yrs)	Female (%)	Mean BMI (kg/m2)	Baseline mean incontinence	Baseline mean voids	Baseline mean urgency	Baseline mean nocturia
A1	Abdelbary	2015	48.70	100		0.5	6.7	5.7	2.2
A2	Abrams	1998	56.80	76		2.8	11.0		
A3	Anderson	1993	70.65	100		2.2	9.6		1.6
A4	Anderson	1999	59.40	92		3.9	7.1		
A5	Appell	2001	59.10	83		4.0	13.0		
A6	Barkin	2004	59.30	90		3.3	11.0		
A7	Batista	2015	57.00	76		2.1	11.0	7.8	2.3
A8	Bent	2008	53.70	100	30.20	1.8			
A9	Burgio	2008	56.90	100	32.75	3.3			
A10	Burgio	2010	58.35	100	33.15	2.4	11.0		
A11	Burgio	2011	64.25	0		0.2	11.0		2.2
A12	But	2012	54.80	100	27.60		9.1	5.8	2.5
A13	Cardozo	2004	55.80	82			12.0		
A14	Cardozo	2005	54.00	71				10.0	
A15	Cardozo	2008	57.76	88		1.7	11.0	5.2	
A16	Cartwright	2011	51.80	100		0.8	7.9	3.3	1.5
A17	Chancellor	2000	60.50	80		3.3	11.0		
A18	Chancellor	2008	57.95	89		2.8	11.0	10.0	1.8
A19	Chapple	2004	53-59	60		1.5	11.0	5.4	
A20	Chapple	2004	57.50	75		2.5	12.0	5.5	
A21	Chapple	2005	56.45	87		2.6	11.0	5.9	1.9
A22	Chapple	2006	50.00	72		2.2	14.0	7.9	
A23	Chapple	2007	56.60	80	27.33	3.7	11.0	11.	1.9
A24	Chapple	2013	59.35	86	29.09	5.6	11.0	8.9	2.1
A25	Chapple	2013	57.00	85		2.8	11.0	5.8	1.8
A26	Chapple	2013	57.15	89					
A27	Chapple	2013	59.63	74		2.5	11.0		2.0
A28	Chapple	2014	59.36	81	30.46	3.9	12.0	11.0	
A29	Chapple	2014	57.60	77		3.6	12.0	6.8	
A30	Choo	2008	52.92	79	23.65	2.0	11.0	4.0	1.7
A31	Chu	2009	58.50	82		3.0	11.0	7.0	
A32	Chuang	2014	65.00	53		0.8	12.0	9.5	
A33	Davila	2001	63.50	92		7.2			
A34	Digesu	2012	55.40	100					
A35	Diokno	2003	60.00	100		6.1	13.0		
A36	Dmochowski	2002	61.38	92		5.2	12.0		
A37	Dmochowski	2003	63.50	93		4.9	12.0		
A38	Dmochowski	2008	59.80	85		4.0	12.0		
A39	Dmochowski	2010	58.80	92					
A40	Dmochowski	2010	59.90	83		2.1	12.0	9.2	2.6
A41	Dmochowski	2014	56.00	100		0.4	8.9		
A42	Drutz	1999	64.14	77	28.62	3.5	11.0		
A43	Dubeau	2014	75.05	82	31.20	4.0	12.0	10.0	3.1
A44	Enzelsberger	1991	50.50	100			8.2		5.0
A45	Enzelsberger	1991	51.50	100	42.93		10.0		5.6
A46	Enzelsberger	1995	60.00	100			7.8		4.8
A47	Enzelsberger	1995	56.55	100			7.9		4.7
A48	Finazzi-Agro	2010	45.20	100		1.3	14.0		
A49	FitzGerald	2008	56.90	100	32.75				1.6
A50	Franzen	2010	58.00	100			11.0		
A51	Frenkl	2010	61.74	94		3.1	10.0		
A52	Fukuda	2013	71.30	79		1.6	12.0	2.4	2.6
A53	Giannitsas	2004	56.00	100			8.5		
A54	Gittelman	2014	57.03	100	31.27	3.9	11.0		
A55	Goldfischer	2015	58.77	87	31.17	6.5	11.0		
A56	Gotoh	2011	57.65	75		1.4	11.0	4.2	1.3
A57	Haab	2014	55.63	62		1.7	13.0	6.4	
A58	Halaska	2003	53.70	86	26.37	1.8	11.0	10.0	
A59	Hassouna	2000	39.00	90			16.0		
A60	Herschorn	2004	64.40	88					
A61	Herschorn	2008	57.67	72		3.3	11.0	6.3	
A62	Herschorn	2010	61.00	78					
A63	Herschorn	2010	58.20	81		2.3	11.0	9.3	2.2
A64	Herschorn	2013	59.00	69	29.50	2.5	11.0	5.6	
A65	Hill	2006	54.73	85		2.3	10.0	8.4	1.6
A66	Ho	2010	57.10	67	24.20	4.7	13.0	4.1	
A67	Holmes	1989	42.05	100			11.0		2.0
A68	Homma	2003	59.32	70		2.9	11.0		
A69	Homma	2008	63.13	70		2.4	11.0	4.9	
A70	Homma	2009	58.60	85		2.5	11.0	5.0	
A71	Hsiao	2011	53.50	100	23.90	0.9	13.0	4.2	1.6
A72	Huang	2012	56.05	100		4.6	9.9		1.2
A73	Huo	2013	39.80	100			17.0		
A74	Jabs	2013	63.40	100		5.5	9.2		2.2
A75	Jacquetin	2001	55.67	79	25.57	2.7	11.0		
A76	Johnson	2005	67.90	100					1.9
A77	Junemann	2005	56.30	79					
A78	Junemann	2006	56.10	90	26.96	3.3	12.0	6.2	
A79	Kafri	2013	56.70	100	28.20	0.9	12.0		
A80	Kaplan	2011	58.30	85		2.5	11.0	9.6	2.2
A81	Kaplan	2014	57.75	82	29.90	3.8	12.0	11.0	
A82	Karram	2009	57.00	84		2.6	11.0	6.0	
A83	Kaya	2011	47.00	100	31.00	1.6	10.0		2.8
A84	Khullar	2004	58.20	100		2.9	10.0	5.8	
A85	Khullar	2013	59.10	72	27.78	2.7	11.0		
A86	Kosilov	2014	69.40	55		5.2	9.3	6.4	
A87	Kuo	2011	66.63	54		2.4	10.0	7.4	

TABLE 5.2: Patient characteristics (cont.)

A88	Kuo	2014	67.00	58		1.0	10.0	9.9	
A89	Kuo	2015	54.50	59		2.3	12.0	5.3	2.4
A90	Kuo	2015	57.93	54		2.2	12.0	5.6	2.4
A91	Kurz	1993	61.15	100			7.7		3.8
A92	Lauti	2008	55.10	100		1.6	8.0	3.4	1.1
A93	Lee	2002	52.00	77	23.25	2.5	12.0		
A94	Lee	2010	52.62	74			12.0	7.4	1.7
A95	Lee	2013	57.79	70		1.5	12.0	5.5	1.5
A96	Lehtoranta	2002	37.00	56		1.3	6.9		
A97	Madersbacher	1999	49.47	93	25.95		11.0	11.0	
A98	Mak	2007	56.20	100	24.45	0.7	10.0	2.1	2
A99	Malone-Lee	2001	75.00	65	26.60	2.9	11.0		
A100	Malone-Lee	2001	65.10	67	27.80	2.6	11.0		
A101	Malone-Lee	2009	56.40	80		1.2	10.0		
A102	Marencak	2011	52.90	100		2.4	10.0	7.7	
A103	Martinez-Garcia	2009	55.10	70		2.4	9.0	5.4	
A104	Mattiasson	2003	62.50	27		2.1	10.0	6.3	
A105	Mattiasson	2010	58.40	86		1.5	11.0	5.0	
A106	Mazur	1995	48.00	98	24.95		11.0		
A107	Meyhoff	1983	51.00	100		0.6	8.0		1
A108	Millard	1999	60.22	75	27.48	3.7	11.0		
A109	Millard	2004	53.40	75	25.26	3.3	12.0	4.2	
A110	Nitti	2007	59.00	76		3.8	12.0	11.0	2.0
A111	Nitti	2010	56.03	83		3.0	11.0		
A112	Nitti	2013	60.10	74	30.20	2.8	11.0	5.8	2.0
A113	Nitti	2013	61.35	89		5.3	11.0	8.2	2.1
A114	Norton	1994	55.05	100		2.0	9.3		1.2
A115	O'Reilly	2008	61.20	100			9.5		
A116	Ohlstein	2012	53.63	100		4.6	11.0		
A117	Olmo	2013	60.00	100		5.0	11.0	6.7	3.1
A118	Oreskovic	2012	56.90	0	29.41				
A119	Orri	2014	47.30	100			10.0		
A120	Ouslander	1993	70.65	100		2.2	9.6		1.6
A121	Ozdedeli	2010	47.00	100	31.00	2.0	9.3	4.9	
A122	Park	2014	57.22	85		3.3	10.0	5.2	
A123	Peters	2009	57.85	94		2.8	12.0	6.7	2.5
A124	Peters	2010	61.35	79	29.60	2.4	12.0	8.1	2.9
A125	Preik	2004	59.40	92	29.37	3.6			
A126	Preyer	2015	56.60	100	26.40	1.7	11.0		
A127	Rentzhog	1998	57.00	76		2.4	10.0		
A128	Rios	2007	56.00	100		4.0	9.7		2.2
A129	Rogers	2008	48.00	100		2.3	12.0		
A130	Rudy	2006	61.05	81		2.8	13.0		2
A131	Rufford	2003	NA	100	27.75	0.0	10.0		
A132	Sahai	2007	50.30	56		4.4	14.0	9.5	
A133	Sancaktar	2010	46.40	100	26.75	3.2	12.0	12.0	
A134	Schmidt	1999	46.60	81		8.9			
A135	Schreiner	2010	68.30	100	29.00	2.3	7.1		2.6
A136	Song	2006	46.52	100			11.0	2.8	1.7
A137	Song	2015	59.18	66		1.0	11.0	5.9	1.5
A138	Soomro	2001	50.00	70			11.0		
A139	Staskin	2007	59.45	85		4.1	12.0		
A140	Staskin	2009	59.40	89	31.25	5.4	12.0		2.5
A141	Steers	2005	58.00	84		2.2	10.0	8.2	1.6
A142	Steers	2007	54.60	100	29.70	1.4	10.0		
A143	Subak	2002	69.25	100		1.6	8.2		
A144	Swift	2003	59.33	100	28.87	3.2	11.0		
A145	Tang	2014	53.00	100			22.0		
A146	Tapp	1989	43.00	100		3.8	10.0		1.4
A147	Tincello	2000	47.00	100		2.0	9.0	6.0	2.0
A148	Tincello	2012	59.45	100		6.2	10.0	7.8	
A149	Tseng	2009	65.35	100	24.90	2.0	14.0	4.4	3.4
A150	Ulshofer	2001	51.50	92	25.55				
A151	Van Kerrebroeck	2001	60.30	81		3.2	11.0		
A152	Van Kerrebroeck	2009	60.50	82		3.2	11.0		
A153	Vardy	2009	59.50	82		2.8	11.0	5.7	1.6
A154	Versi	2000	59.20	89	29.04				
A155	Visco	2012	58.00	100	32.50	5.0			
A156	Wagg	2013	72.70	53	28.05	1.5	12.0	8.6	2.8
A157	Wang	2006	NA	100					
A158	Wang	2009	53.18	100	24.25				
A159	Weiss	2013	57.75	65	30.35	2.2	12.0	9.9	3.1
A160	Yamaguchi	2007	60.18	81		2.1	11.0	4.2	1.8
A161	Yamaguchi	2011	57.60	80	23.37				
A162	Yamaguchi	2014	58.27	84	22.64	1.9	11.0	4.2	1.7
A163	Yamaguchi	2014	55.73	89		1.2	11.0	3.6	1.3
A164	Yamaguchi	2014	55.93	82	22.84	1.9	11.0	4.6	1.6
A165	Yamanishi	2000	70.00	57	34.19				
A166	Yokoyama	2013	71.20	62					
A167	Yokoyama	2014	60.77	75	23.33				1.9
A168	Yoon	2003	NA	100			16.0		1.9
A169	Zat'ura	2010	52.27	93		1.9	16.0		
A170	Zellner	2009	61.55	91					
A171	Zimmern	2010	56.50	100		3.7	7.2		
A172	Zinner	2002	62.50	81		3.2	11.0		
A173	Zinner	2005	59.90	93		2.9	10.0	9.3	
A174	Zinner	2006	59.10	87		2.8	11.0	12	

TABLE 5.3: Baseline characteristics of patients receiving first, second, and third-line therapies as recommended by the NICE in England and Wales

	Baseline mean incontinence episodes (SD)	Baseline mean voiding episodes (SD)	Baseline mean urgency episodes (SD)	Baseline mean nocturia episodes (SD)
First line therapy	1.20 (0.68)	10.92 (2.73)	5.99 (3.08)	2.22 (0.58)
Second line therapy	2.87 (1.35)	11.39 (1.94)	6.69 (3.00)	2.19 (0.92)
Second line therapy*	2.48 (0.51)	11.82 (0.56)	5.64 (0.71)	2.17 (0.53)
Third line therapy	4.57 (2.75)	12.98 (4.13)	7.98 (0.98)	2.3 (0.17)
Not currently recommended	2.83 (1.61)	11.31 (1.85)	6.88 (2.30)	2.27 (1.17)

*Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)

5.5 Efficacy data

Table B.2 of Appendix B displays the efficacy data extracted from eligible trials. A total of 117, 124, 62, and 57 studies reported outcomes for urinary incontinence, voiding, urgency and nocturia episodes, respectively. Generally the RCTs were performed on large patient populations, where the median number of participants randomised to an intervention group was 104 (range: 8, 942), 96 (range: 6, 942), 107 (range: 11, 942), and 107 (range: 10, 942) patients for incontinence, voiding, urgency, and nocturia episodes, respectively. Nearly one third of all studies failed to report measures of variability in the mean treatment effects. In total, 44 (38%), 42 (34%), 18 (29%) and 17 (30%) studies solely reported mean effects and gave no measure of uncertainty or variability for incontinence, voiding, urgency, and nocturia episodes, respectively. In total, 143 (82%) studies reported one of the three cardinal symptoms of OAB (incontinence, urgency, and increased voiding). Of these, only 51 (36%) reported treatment effects for all 3 outcomes. It is apparent that there is an element of selective reporting in the trials assessing OAB as a complete symptom syndrome, and this is further assessed in Section 5.8 below.

5.6 Safety and tolerability data

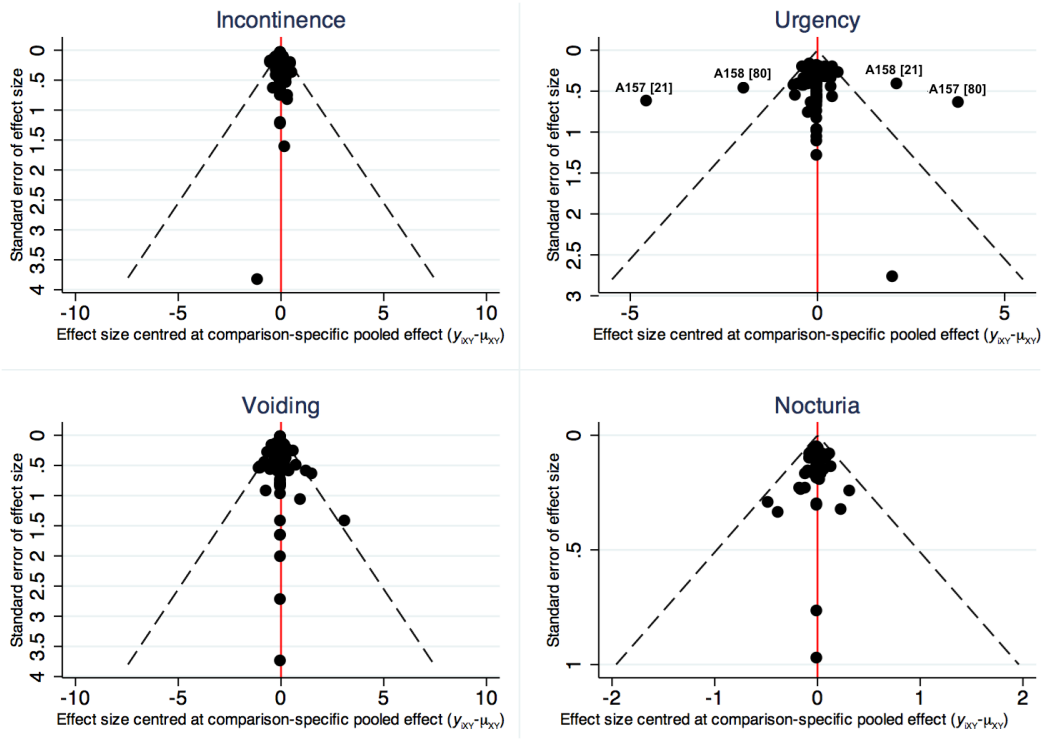
Table B.3 of Appendix B displays the extracted data from eligible trials reporting safety and tolerability measures. In total, data were collected from 94, 103, and 57 studies for number of patients experiencing adverse events, discontinuation due to adverse events, and discontinuation due to a lack of efficacy. For the number of patients with adverse events, 4 studies reported zero events in both treatment arms. The median number of participants randomised to an intervention group was 108 (range: 6, 936) for number of patients with adverse events, 153 (range: 6, 936) and 169 (range: 27, 936) for discontinuation due to adverse events and a lack of efficacy, respectively.

5.7 Publication bias

Publication bias occurs when studies with favourable or positive results are more likely to be published in the medical literature than that of studies displaying negative or unfavourable results. As with all analyses relying on the published literature, publication bias can play a major role in inducing bias in the overall results, resulting in misleading conclusions. This is particularly true of meta-analyses. Meta-analysing data with an element of publication bias on the premise that all evidence, positive and negative are available, will lead to a distorted picture of the true overall treatment effect. It is therefore necessary to assess publication bias and adjust for biases in meta-analyses that are being used to inform decision models (Welton et al., 2012).

Funnel plots are a visual tool that are primarily used to assess publication and other biases in meta-analysis. For pairwise meta-analyses, standard funnel plots are simple scatter plots of treatment effect size for each individual study versus the size of that study. If the study estimates are lying symmetrically around the line of the meta-analysis summary then the funnel plot suggests little evidence of small study effects or publication bias (Rothstein et al., 2005). For NMA, interest lies in the differences in relative treatment effects between small and large trials. As in NMA, multi-arm trials comparing 2 or more treatments are included and thus different treatment comparisons are evaluated with each pair of comparisons displaying its own summary effect; and so a common reference line of symmetry does not exist. Chaimani et al. (2013) have recently developed a 'comparison adjusted' funnel plot for NMA which is implemented in `STATA`. For treatment comparison XY, the x-axis of the comparison-adjusted funnel plot shows the difference between the study-specific effect sizes y_{iXY} for study i and the respective comparison-specific summary effect θ_{XY}, $(y_{iXY} - \theta_{XY})$. The y-axis represents a measure of dispersion of y_{iXY}. In the absence of small-study effects and publication bias, all studies are expected to lie symmetrically about the line centred at zero.

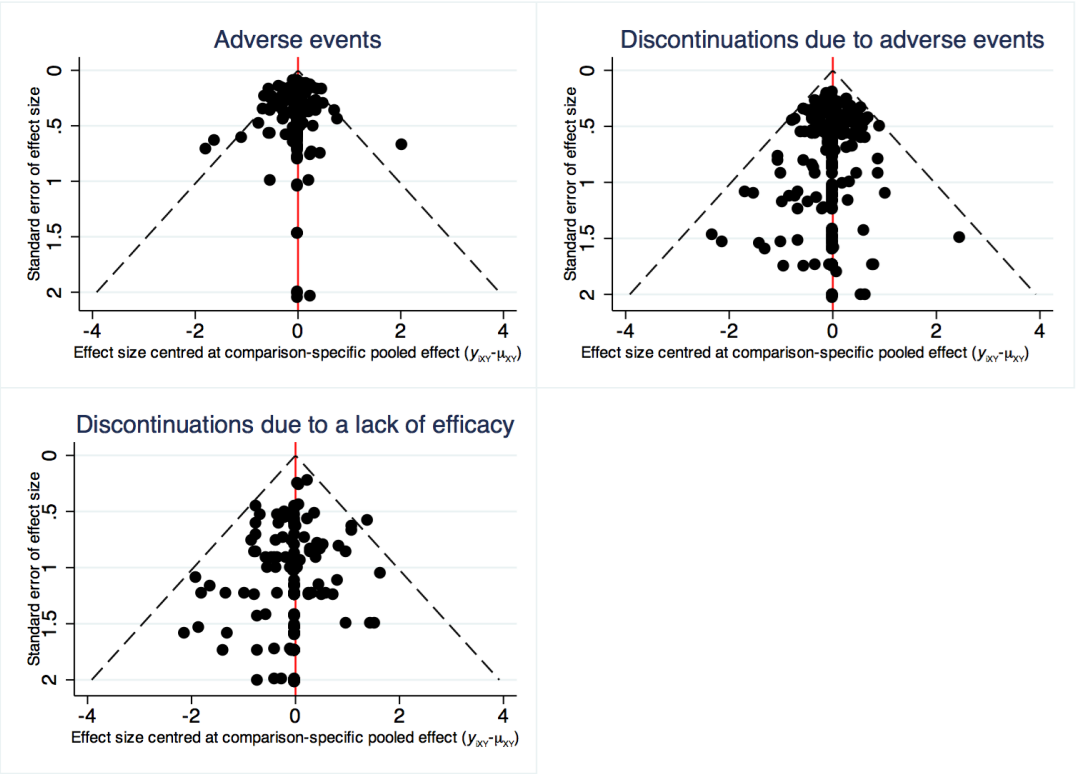
FIGURE 5.2: Assessment of publication bias: efficacy data



5.7.1 Efficacy data

Figure 5.2 displays the comparison adjusted funnel plot for efficacy data. All studies appeared to lie symmetrically around the zero line suggesting that there is no evidence of publication bias. For urgency episodes, two studies by Wang et al. (2006) (study id: A157) and Wang et al. (2009) (study id: A158) appeared to have heterogeneously reported treatment effects for both oxybutynin IR 2.5mg t.i.d [21] and electrostimulation [80] relative to a placebo intervention, compared to the comparison specific summary estimate. Wang et al. (2006) and Wang et al. (2009) are amongst the smaller studies evaluating urgency episodes with randomised treatment groups in the range of 21 to 26 patients. Despite the small scale study, the reported standard errors seem to be unexpectedly small for urgency, which may explain the implicit biases highlighted in Figure 5.2. This did not appear to be the case for all other efficacy measures reported by Wang et al. (2006) and Wang et al. (2009).

FIGURE 5.3: Assessment of publication bias: safety and tolerability data



5.7.2 Safety and tolerability data

Figure 5.3 illustrates the comparison adjusted funnel plot for safety and tolerability data. All studies appeared to be equally symmetrical about the zero line suggesting that there is no evidence of publication bias or small-study effects.

5.8 Quality assessment and risk of bias

Internal validity of trials was assessed through evaluation of random sequence generation, allocation concealment, blinding of outcome, blinding of participants, incomplete outcome data, and selective reporting (Higgins et al., 2011). Figure 5.4 illustrates the between-study quality assessment. Within-study quality assessment is given in Table 5.4. Studies identified in this review often did not adequately report random sequence generation and thus there is a large proportion of studies with an unclear level of bias. This is also true of studies failing to adequately

report allocation concealment. Although the risk of bias in the included studies was mostly low, a large number (48.3%) of trials were considered to have a high risk of bias for selective reporting. In total 84 studies evaluating OAB failed to report at least one relevant outcome measure of interest.

FIGURE 5.4: Between-study quality assessment

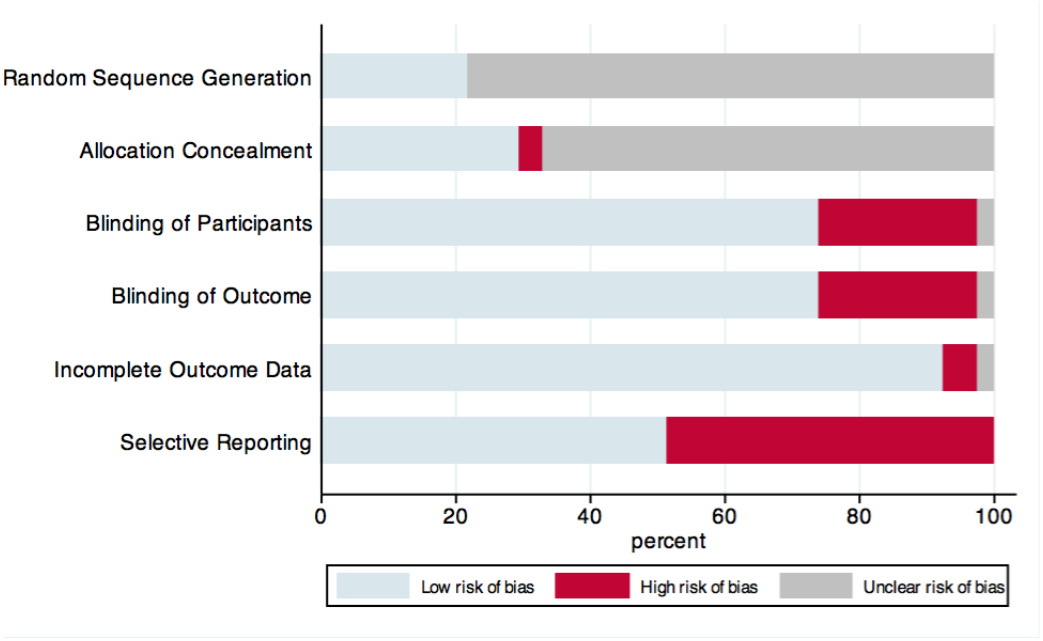


TABLE 5.4: Within-study quality assessment

Reference	First author	Year	Random sequence generation	Allocation concealment	Blinding of participants	Blinding of outcome	Incomplete outcome data	Selective Reporting
A1	Abdelbary	2015	Low	Unclear	High	High	Low	Low
A2	Abrams	1998	Unclear	Unclear	Low	Low	Low	High
A3	Anderson	1993	Unclear	Unclear	Low	Low	Low	High
A4	Anderson	1999	Unclear	Unclear	Low	Low	Low	Low
A5	Appell	2001	Unclear	Unclear	Low	Low	Low	High
A6	Barkin	2004	Unclear	Unclear	Low	Low	Low	High
A7	Batista	2015	Low	Low	Low	Low	High	High
A8	Bent	2008	Low	Low	Low	Low	Low	Low
A9	Burgio	2008	Low	Low	High	High	Low	High
A10	Burgio	2010	Unclear	Unclear	High	High	Low	High
A11	Burgio	2011	Unclear	Low	High	High	Low	High
A12	But	2012	Low	Unclear	High	High	Low	Low
A13	Cardozo	2004	Unclear	Unclear	Low	Low	Low	Low
A14	Cardozo	2005	Unclear	Unclear	Low	Low	Low	Low
A15	Cardozo	2008	Unclear	Unclear	Low	Low	Low	Low
A16	Cartwright	2011	Low	Low	Low	Low	Low	Low
A17	Chancellor	2000	Unclear	Unclear	Low	Low	Unclear	High
A18	Chancellor	2008	Low	Unclear	High	High	Low	Low
A19	Chapple	2004	Unclear	Unclear	Low	Low	Low	Low
A20	Chapple	2004	Unclear	Unclear	Low	Low	Low	Low
A21	Chapple	2005	Unclear	Low	Low	Low	Low	Low
A22	Chapple	2006	Low	Unclear	Low	Low	Low	Low
A23	Chapple	2007	Unclear	Unclear	Low	Low	Low	Low
A24	Chapple	2013	Unclear	Low	Low	Low	Low	Low
A25	Chapple	2013	Low	Low	Low	Low	Low	Low
A26	Chapple	2013	Unclear	Unclear	Low	Low	Low	Low
A27	Chapple	2013	Low	Unclear	Low	Low	Low	High
A28	Chapple	2014	Low	Low	Low	Low	Low	Low
A29	Chapple	2014	Low	Unclear	Low	Low	Low	Low
A30	Choo	2008	Unclear	Unclear	Low	Low	Low	Low
A31	Chu	2009	Low	Low	Low	Low	Low	Low
A32	Chuang	2014	Unclear	High	Low	Low	Low	Low
A33	Davila	2001	Unclear	Unclear	Low	Low	Low	Low
A34	Digesu	2012	Low	Low	Low	Low	Low	Low
A35	Diokno	2003	Unclear	Unclear	Low	Low	Low	High
A36	Dmochowski	2002	Unclear	Unclear	Low	Low	Low	High
A37	Dmochowski	2003	Unclear	Unclear	Low	Low	Low	High
A38	Dmochowski	2008	Unclear	Unclear	Low	Low	Low	High
A39	Dmochowski	2010	Unclear	Unclear	Low	Low	Low	High
A40	Dmochowski	2010	Low	Low	Low	Low	Low	Low
A41	Dmochowski	2014	Unclear	Unclear	Low	Low	Low	High
A42	Drutz	1999	Unclear	Unclear	Low	Low	High	High
A43	Dubeau	2014	Low	Unclear	Low	Low	Low	Low
A44	Enzelsberger	1991	Unclear	Unclear	Unclear	Unclear	Low	High
A45	Enzelsberger	1991	Unclear	Unclear	Unclear	Unclear	Low	High
A46	Enzelsberger	1995	Low	Unclear	Unclear	Unclear	Low	High
A47	Enzelsberger	1995	Unclear	Low	Low	Low	Low	High
A48	Finazzi-Agro	2010	Low	Unclear	Low	Low	Low	High
A49	FitzGerald	2008	Unclear	Unclear	High	High	Low	Low
A50	Franzen	2010	Low	Low	High	High	Low	High
A51	Frenkl	2010	Unclear	Low	Low	Low	Low	High
A52	Fukuda	2013	Low	Low	High	High	Low	Low
A53	Giannitsas	2004	Low	Unclear	High	High	Low	High
A54	Gittelman	2014	Unclear	Unclear	Low	Low	Low	High
A55	Goldfischer	2015	Low	Low	Low	Low	Low	High
A56	Gotoh	2011	Unclear	Low	Low	Low	Low	Low
A57	Haab	2014	Unclear	Unclear	Low	Low	Low	Low
A58	Halaska	2003	Unclear	Unclear	Low	Low	Low	Low
A59	Hassouna	2000	Unclear	Unclear	High	High	Low	High
A60	Herschorn	2004	Unclear	Unclear	High	High	Low	High
A61	Herschorn	2008	Unclear	Unclear	Low	Low	Low	Low
A62	Herschorn	2010	Low	Unclear	Low	Low	Low	Low
A63	Herschorn	2010	Unclear	Low	Low	Low	Low	Low
A64	Herschorn	2013	Unclear	Unclear	Low	Low	Low	Low
A65	Hill	2006	Unclear	Unclear	Low	Low	Low	Low
A66	Ho	2010	Unclear	Unclear	High	High	Low	Low
A67	Holmes	1989	Low	Unclear	High	High	Low	High
A68	Homma	2003	Unclear	Unclear	Low	Low	Low	High
A69	Homma	2008	Unclear	Unclear	Low	Low	Low	Low
A70	Homma	2009	Unclear	Unclear	Low	Low	Low	Low
A71	Hsiao	2011	Low	High	High	High	Low	Low
A72	Huang	2012	Low	Low	Low	Low	Low	High
A73	Huo	2013	Low	High	High	High	Low	High
A74	Jabs	2013	Unclear	Low	Low	Low	Low	High
A75	Jacquetin	2001	Unclear	Unclear	Low	Low	Low	High
A76	Johnson	2005	Unclear	Unclear	Low	Low	Low	Low
A77	Junemann	2005	Unclear	Unclear	Low	Low	Low	Low
A78	Junemann	2006	Unclear	Unclear	Low	Low	Low	Low
A79	Kafri	2013	Unclear	Low	High	High	Low	High
A80	Kaplan	2011	Unclear	Unclear	Low	Low	Low	Low
A81	Kaplan	2014	Unclear	Low	Low	Low	High	High
A82	Karram	2009	Unclear	Unclear	Low	Low	Low	Low
A83	Kaya	2011	Low	Unclear	High	High	Low	High
A84	Khullar	2004	Low	Low	Low	Low	Low	Low
A85	Khullar	2013	Low	Low	Low	Low	Low	Low

TABLE 5.4: Within study quality assessment (cont.)

A86	Kosilov	2014	Low	Low	Low	Low	High	High
A87	Kuo	2011	Unclear	Low	High	High	Low	Low
A88	Kuo	2014	Unclear	Unclear	Low	Low	Low	Low
A89	Kuo	2014	Low	Unclear	Low	Low	High	High
A90	Kuo	2015	Low	Unclear	Low	Low	Low	Low
A91	Kurz	1993	Unclear	Unclear	Low	Low	Low	High
A92	Lauti	2008	Low	Low	High	High	Low	Low
A93	Lee	2002	Low	Unclear	Low	Low	Low	High
A94	Lee	2010	Unclear	High	Low	Low	Low	High
A95	Lee	2013	Low	Low	Low	Low	High	Low
A96	Lehtoranta	2002	Unclear	Unclear	Low	Low	Low	High
A97	Madersbacher	1999	Unclear	Unclear	Low	Low	Low	High
A98	Mak	2007	Low	High	High	High	Low	Low
A99	Malone-Lee	2001	Unclear	Unclear	Low	Low	Low	High
A100	Malone-Lee	2009	Unclear	Unclear	Low	Low	Low	High
A101	Malone-Lee	2009	Unclear	Low	Low	Low	Low	High
A102	Marencak	2011	Low	Unclear	Low	Low	Low	Low
A103	Martinez-Garcia	2009	Unclear	Unclear	Low	Low	Low	Low
A104	Mattiasson	2003	Low	Unclear	High	High	Low	Low
A105	Mattiasson	2010	Unclear	Unclear	High	High	Low	Low
A106	Mazur	1995	Unclear	Unclear	High	High	Low	High
A107	Meyhoff	1983	Unclear	Unclear	Low	Low	Low	High
A108	Millard	1999	Unclear	Unclear	Low	Low	Low	High
A109	Millard	2004	Unclear	Unclear	High	High	Low	Low
A110	Nitti	2007	Low	Unclear	Low	Low	Low	Low
A111	Nitti	2010	Unclear	Unclear	Low	Low	Low	High
A112	Nitti	2013	Low	Unclear	Low	Low	Low	Low
A113	Nitti	2013	Unclear	Low	Low	Low	Low	Low
A114	Norton	1994	Unclear	Low	Low	Low	Unclear	High
A115	O'Reilly	2008	Unclear	Low	Low	Low	Low	High
A116	Ohlstein	2012	Unclear	Unclear	Low	Low	Low	High
A117	Olmo	2013	Unclear	Unclear	Low	Low	Low	Low
A118	Oreskovic	2012	Unclear	Unclear	Low	Low	Unclear	High
A119	Orri	2014	Unclear	Low	Low	Low	Low	High
A120	Ouslander	1993	Unclear	Unclear	Low	Low	Low	High
A121	Ozdedeli	2010	Low	Low	High	High	Low	High
A122	Park	2014	Low	Unclear	Low	Low	Low	Low
A123	Peters	2009	Unclear	Unclear	High	High	Low	Low
A124	Peters	2010	Unclear	Unclear	Low	Low	Low	Low
A125	Preik	2004	Unclear	Unclear	Low	Low	Low	High
A126	Preyer	2015	Low	Low	High	High	Low	High
A127	Rentzhog	1998	Unclear	Unclear	Low	Low	Low	High
A128	Rios	2007	Unclear	Low	Low	Low	Low	High
A129	Rogers	2008	Unclear	Unclear	Low	Low	Low	High
A130	Rudy	2006	Unclear	Unclear	Low	Low	Low	High
A131	Rufford	2003	Unclear	Low	Low	Low	Low	High
A132	Sahai	2007	Low	Low	Low	Low	Low	Low
A133	Sancaktar	2010	Unclear	High	High	High	Low	Low
A134	Schmidt	1999	Unclear	Unclear	High	High	High	Low
A135	Schreiner	2010	Low	Unclear	High	High	Low	High
A136	Song	2006	Unclear	Unclear	High	High	Low	Low
A137	Song	2015	Unclear	Unclear	Low	Low	Low	Low
A138	Soomro	2001	Unclear	Unclear	High	High	Low	High
A139	Staskin	2007	Unclear	Low	Low	Low	Low	High
A140	Staskin	2009	Unclear	Unclear	Low	Low	Low	High
A141	Steers	2005	Unclear	Unclear	Low	Low	Low	Low
A142	Steers	2007	Unclear	Unclear	Low	Low	High	High
A143	Subak	2002	Low	Low	High	High	High	High
A144	Swift	2003	Unclear	Unclear	Low	Low	Low	High
A145	Tang	2014	Unclear	Unclear	High	High	Low	High
A146	Tapp	1989	Unclear	Unclear	Low	Low	Unclear	High
A147	Tincello	2000	Low	Low	High	High	Low	Low
A148	Tincello	2012	Low	Low	Low	Low	Low	Low
A149	Tseng	2009	Low	Unclear	High	High	Low	Low
A150	Ulshofer	2001	Unclear	Unclear	Low	Low	Low	High
A151	Van Kerrebroeck	2001	Unclear	Low	Low	Low	Low	High
A152	Van Kerrebroeck	2009	Unclear	Unclear	Low	Low	Low	Low
A153	Vardy	2009	Unclear	Low	Low	Low	Low	Low
A154	Versi	2000	Low	Low	Low	Low	Low	Low
A155	Visco	2012	Unclear	Unclear	Low	Low	Low	Low
A156	Wagg	2013	Unclear	Low	Low	Low	Low	Low
A157	Wang	2006	Unclear	Low	High	High	Low	Low
A158	Wang	2009	Low	Unclear	High	High	Low	Low
A159	Weiss	2013	Unclear	Low	Low	Low	Low	Low
A160	Yamaguchi	2007	Unclear	Unclear	Low	Low	Low	Low
A161	Yamaguchi	2011	Unclear	Unclear	Low	Low	Low	Low
A162	Yamaguchi	2014	Unclear	Unclear	Low	Low	Low	Low
A163	Yamaguchi	2014	Unclear	Unclear	Low	Low	Low	Low
A164	Yamaguchi	2014	Unclear	Low	Low	Low	Low	Low
A165	Yamanishi	2000	Unclear	Unclear	Low	Low	Low	High
A166	Yokoyama	2013	Unclear	Unclear	Unclear	Unclear	Low	High
A167	Yokoyama	2014	Unclear	Unclear	Low	Low	Low	Low
A168	Yoon	2003	Low	Unclear	High	High	Low	High
A169	Zat'ura	2010	Low	Low	Low	Low	Low	High
A170	Zellner	2009	Unclear	Low	Low	Low	Low	High
A171	Zimmern	2010	Unclear	Unclear	High	High	Low	High
A172	Zinner	2002	Low	Unclear	Low	Low	Low	High
A173	Zinner	2005	Unclear	Unclear	Low	Low	Low	Low
A174	Zinner	2006	Unclear	Unclear	Low	Low	Low	Low

Low refers to a low risk of bias, High refers to a high risk of bias, Unclear refers to an unclear risk of bias

5.9 Networks of evidence

A network diagram can be made up of different interventions, trials, and treatment comparisons. In this analysis, the numbered nodes represent individual interventions with corresponding intervention codes given in Table 5.5. The interconnecting lines illustrate a direct comparison between interventions - that is, one or more of the trials have evaluated both interventions on a head-to-head basis. The density of the interconnecting line represents the number of trials that directly compared the adjoining interventions. In situations where interventions are not compared to a common comparator or to an intervention in the network, the network of evidence becomes disconnected as previously described in Section 3.4.5. This can also be true for trials that evaluate interventions within the network of evidence but fail to report the outcome of interest. For interventions disconnected from the network it is not possible to obtain relative treatment effects and thus these interventions were excluded from the analysis for general NMA methods.

5.9.1 Efficacy data

Figures 5.5, 5.6, 5.7 and 5.8 illustrate the networks of evidence for urinary incontinence, voiding, urgency, and nocturia episodes, respectively. For incontinence, a total of 101 interventions were identified in the systematic review, 4 interventions were disconnected from the network of evidence: darifenacin 7.5-15mg once daily + BT [88], tolterodine + BT [94], darifenacin ER 7.5-15mg once daily [104], and tolterodine [105], and thus, 97 interventions (Figure 5.5) were included in the evidence synthesis and treatment comparisons described in Chapters 6 and 7.

108 different interventions were analysed for voiding episodes in the systematic review, 8 interventions were disconnected from the network of evidence: darifenacin 7.5-15mg once daily + BT [88], tolterodine + BT [94], darifenacin ER 7.5-15mg once daily [104], tolterodine [105], lidocaine gel 2x6ml [129], emepronium bromide

immediate release 200mg three times days [130], sacral nerve stimulation + tolterodine extended release 2mg once daily [136], and tolterodine extended release 2mg once daily [137], and therefore, 100 interventions (Figure 5.6) were included in the evidence synthesis using general NMA frameworks described in Chapters 6 and 7.

In total 58 interventions were analysed for urgency in the systematic review. Overall 54 interventions were included in the evidence synthesis in Chapters 6 and 7, and 4 interventions were disconnected from the network of evidence (Figure 5.7): control intervention [2], reflexology [71], darifenacin 7.5-15mg once daily + BT [88], and darifenacin ER 7.5-15mg once daily [104].

For nocturia a total of 55 different interventions reported intervention effects in the systematic review, 15 interventions were disconnected from the network of evidence: control intervention [2], oxybutynin immediate release 5mg three times daily [7], trospium chloride immediate release 15mg three times daily [46], reflexology [71], PFMT/Physiotherapy [84], PFMT + BT [89], darifenacin 7.5-15mg once daily + BT [88], trospium chloride immediate release 15mg three times daily + physiotherapy [91], tolterodine + BT [94], electrostimulation + PFMT + BT [97], darifenacin ER 7.5-15mg once daily [104], tolterodine [105], propantheline bromide 15mg three times daily [113], lidocaine gel 2x6ml [129], and emepronium bromide immediate release 200mg three times days [130], and thus, 40 interventions (Figure 5.8) were included in the evidence synthesis and clinical effectiveness comparisons described in Chapters 6 and 7.

TABLE 5.5: Intervention codes

Treatment pathway	Code	Intervention	Number of studies	Number of patients	Treatment pathway	Code	Intervention	Number of studies	Number of patients
	[1]	Placebo	108	19415		[71]	Reflexology	1	57
	[2]	Control	5	208		[72]	OnaBoNT-A 100u trigone sparing	6	781
	[3]	Sham therapy	4	188		[73]	OnaBoNT-A 200U trigone sparing	4	203
	[4]	Tolterodine ER 4mg q.d	34	9750		[74]	OnaBoNTA 50u trigone sparing	1	57
	[5]	Tolterodine IR 2mg b.i.d	23	3812		[75]	OnaBoNTA 150u trigone sparing	1	49
	[6]	Tolterodine IR 1mg b.i.d	4	297		[76]	OnaBoNTA 300u trigone sparing	1	56
	[7]	Oxybutynin IR 5mg t.i.d	7	523		[77]	Solifenacin ER 5 - 10mg q.d + BT	1	321
	[8]	Oxybutynin ER 10mg q.d	2	576		[78]	OnaBoNT-A 100u bladder body + trigone	1	35
	[9]	Oxybutynin ER 15mg q.d	1	65		[79]	OnaBoNT-A 100u bladder base + trigone	1	33
	[10]	Oxybutynin trandermal 3.9mg/day	3	294		[80]	Electrostimulation	7	278
	[11]	Oxybutynin transdermal 1.3mg/day	2	168		[81]	Sacral nerve stimulation	2	77
	[12]	Oxybutynin transdermal 2.6mg/day	1	133		[82]	Tarafenacin 0.4mg q.d	1	76
	[13]	Oxybutynin chloride topical gel 1g/day	1	389		[83]	Percutaneous tibial nerve stimulation	5	207
	[14]	Oxybutynin intravesically 5mg t.i.d	1	9		[84]	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	4	83
	[15]	Oxbutynin patch 73.5mg	1	573		[85]	Bladder Training (BT)/Behaviour Therapy	7	304
	[16]	Oxybutynin vaginal ring 4mg q.d	1	115		[86]	Oxybutynin ER 5-30mg q.d + BT	1	32
	[17]	Oxybutynin vaginal ring 6mg q.d	1	96		[87]	Tolterodine ER 4mg q.d + BT	2	288
	[18]	Oxybutynin IR 5mg b.i.d	4	415		[88]	Darifenacin 7.5 - 15mg q.d + BT	1	205
	[19]	Oxybutynin IR 3mg t.i.d	1	244		[89]	PFMT + BT	2	67
	[20]	Oxybutynin ER 2.5mg q.d	2	62		[90]	Tarafenacin 0.2mg q.d	1	77
	[21]	Oxybutynin IR 2.5mg t.i.d	2	47		[91]	Trospium chloride IR 15mg t.i.d + Physiotherapy	2	32
	[22]	Oxybutynin ER 5 - 30mg/day	5	319		[92]	Oxybutynin ER 2.5mg q.d + BT	1	19
	[23]	Oxybutynin IR 5 - 20mg	3	219		[93]	Tolterodine IR 2mg b.i.d + BT	2	275
	[24]	Oxybutynin IR 2.5 - 5mg b.i.d	5	1136		[94]	Tolterodine + BT	2	193
	[25]	Fesoterodine ER 4mg q.d	8	3298		[95]	Tolterodine IR 2mg b.i.d + PFMT	1	223
	[26]	Fesoterodine ER 8mg q.d	7	2847		[96]	Tolterodine ER 4mg q.d + Neurostimulation	1	20
	[27]	Fesoterodine ER 4 - 8mg q.d	4	1517		[97]	Electrostimulation + PFMT + BT	1	25
	[28]	Terodiline 25mg b.i.d	4	173		[98]	Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	1	30
	[29]	Solifenacin ER 5mg q.d	14	2382		[99]	Tolterodine ER 2mg b.i.d + oestrogen 0.625mg 2xwk	1	40
	[30]	Solifenacin ER 10mg q.d	6	1411		[100]	Solifenacin/trospium + placebo injection	1	127
	[31]	Solifenacin ER 5 - 10mg q.d	4	1763		[101]	Tolterodine 2mg + Pilocarpine 9mg b.i.d	1	130
	[32]	Solifenacin ER 2.5mg q.d	1	41		[102]	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	1	188
	[33]	Solifenacin ER 20mg q.d	1	37		[103]	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	1	188
	[34]	Solifenacin ER 5 - 15mg q.d	1	377		[104]	Darifenacin ER 7.5 - 15mg q.d	1	190
	[35]	Imidafenacin 0.05mg b.i.d	1	91		[105]	Tolterodine	2	198
	[36]	Imidafenacin 0.1mg b.i.d	5	651		[106]	Oxybutynin 20mg intravesically q.d	2	44
	[37]	Imidafenacin 0.25mg b.i.d	1	76		[107]	Serlopitant 0.25mg q.d	1	110
	[38]	Darifenacin ER 30mg q.d	3	227		[108]	Serlopitant 1mg q.d	1	110
	[39]	Darifenacin ER 7.5mg q.d	3	413		[109]	Serlopitant 4mg q.d	1	114
	[40]	Darifenacin ER 15mg q.d.	3	397		[110]	Netupitant 50mg q.d	1	62
	[41]	Propiverine ER 20mg q.d	7	1846		[111]	Netupitant 100mg q.d	1	61
	[42]	Propiverine ER 30mg q.d	1	391		[112]	Netupitant 200mg q.d	1	61
	[43]	Propiverine IR 15mg b.i.d	3	541		[113]	Propantheline Bromide 15mg t.i.d	1	23
	[44]	Trospium chloride ER 60mg q.d	2	578		[114]	Naftopidil 25mg q.d	1	22
	[45]	Trospium chloride IR 20mg b.i.d	2	595		[115]	Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	1	21
	[46]	Trospium chloride IR 15mg t.i.d	3	52		[116]	Propiverine IR 15mg t.i.d	1	149
	[47]	Trospium chloride IR 45mg t.i.d	1	828		[117]	Propiverine IR 30mg b.i.d	1	47
	[48]	Mirabegron IR 100mg b.i.d	1	65		[118]	Propiverine IR 45mg t.i.d	1	49
	[49]	Mirabegron IR 150mg b.i.d	1	65		[119]	Propiverine ER 60mg q.d	1	43
	[50]	Mirabegron ER 25mg q.d	3	810		[120]	Fesoterodine IR 4mg b.i.d	1	43
	[51]	Mirabegron ER 50mg q.d	10	4334		[121]	Fesoterodine IR 8mg b.id	1	47
	[52]	Mirabegron ER 100mg q.d	5	2128		[122]	Fesoterodine IR 12mg b.i.d	1	38
	[53]	Mirabegron ER 200mg q.d	1	167		[123]	Trospium 30mg/day + Solifenacin 10mg/day (cyclic)	1	55
	[54]	Solabegron IR 50mg b.i.d	1	88		[124]	Tolterodine IR 0.5mg b.i.d	1	21
	[55]	Solabegron IR 125mg b.i.d	1	85		[125]	Electromagnetic stimulation	1	33
	[56]	Cizolirtine citrate 200mg b.i.d	1	25		[126]	Trospium 30mg/day + Solifenacin 10mg/day (continuous)	1	62
	[57]	Cizolirtine citrate 400mg b.i.d	2	81		[127]	Estradiol 1mg intravaginally	1	15
	[58]	ZD0947IL 25mg/day	1	92		[128]	Estradiol 3mg intravaginally	1	15
	[59]	ONO-8539 30mg b.i.d	1	87		[129]	Lidocaine gel 2x6ml	1	15
	[60]	ONO-8539 100mg b.i.d	1	83		[130]	Emepronium Bromide IR 200mg t.i.d	1	15
	[61]	ONO-8539 300mg b.i.d	1	82		[131]	Estriol 1mg intravesically	1	21
	[62]	Pregabalin 150mg b.i.d	1	188		[132]	Vaginal oestrogen cream 1.25mg/day	1	105
	[63]	Emepronium bromide ER 200mg q.d	1	19		[133]	Electrostimulation + vaginal oestrogen cream 1.25mg/day	1	105
	[64]	Flavoxate chloride 200mg q.d	1	19		[134]	Oxybutynin gel 84mg/day	1	214
	[65]	Duloxetine 40mg b.i.d	1	81		[135]	Oxybutynin gel 56mg/day	1	210
	[66]	Duloxetine 60mg b.i.d	1	153		[136]	Sacral Nerve Stimulation + Tolterodine ER 2mg q.d	1	120
	[67]	Resiniferatoxin 50nM	1	34		[137]	Tolterodine ER 2mg q.d	1	120
	[68]	Estradiol 25mg	1	20		[138]	Lipo-BoNTA 200U	1	31
	[69]	Elocalcitol 150mg	1	87		[139]	Trospium 60mg/day + Solifenacin 20mg/day (cyclic)	1	58
	[70]	Elocalcitol 75mg	1	84		[140]	Tolterodine IR 4mg b.i.d	1	16

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

5.9.2 Safety and tolerability data

Figures 5.9, 5.10, and 5.11 illustrate the networks of evidence for the number of patients experiencing adverse events, discontinuation due to adverse events, and discontinuation due to a lack of efficacy. For the number of participants with adverse events, 88 interventions were adequately reported for participants experiencing adverse events. In total, 7 interventions were disconnected from the network: oxybutynin extended release 5-30mg/day [22], oxybutynin immediate release 5-20

mg/day [23], PFMT + BT [89], electrostimulation + PFMT + BT [97], trospium 30mg/day + solifenacin 10mg/day (cyclic) [123], trospium 30mg/day + solifenacin 10mg/day (continuous) [126], trospium 60mg/day + solifenacin 20mg/day (cyclic) [139], and therefore 81 interventions (Figure 5.9) were included in evidence synthesis methods described in Chapters 6 and 7.

A total of 88 and 62 different interventions were evaluated in the systematic literature review for discontinuation due to adverse events, and discontinuation due to a lack of efficacy, respectively. Seven interventions were disconnected from the network: oxybutynin extended release 5-30mg/day [22], oxybutynin immediate release 5-20 mg/day [23], trospium chloride immediate release 15mg three times daily [46], PFMT/Physiotherapy [84], darifenacin 7.5-15mg once daily + BT [88], trospium chloride immediate release 15mg three times daily + physiotherapy [91], darifenacin ER 7.5-15mg once daily [104], and thus, 81 and 55 interventions were included in the evidence synthesis methods, respectively, based on the general NMA framework described in Chapters 6 and 7.

5.9.3 Interventions not included in the evidence synthesis

It is clear that there is an element of selective reporting in the studies, and as a result, different interventions are evaluated for different outcomes. Disconnected networks can have severe implications for healthcare decision makers as emerging or alternative treatments cannot always be compared to standard of care and thus, it is difficult to estimate patient benefit. In this analysis, darifenacin 7.5-15mg q.d + BT [88], tolterodine (dose not specified) + BT [94], darifenacin ER 7.5-15mg q.d [104], tolterodine (dose not specified) [105], lidocaine gel 6ml [129], emepronium bromide IR 200mg [130], SNS + tolterodine ER 2mg q.d [136], and tolterodine ER 2mg once daily [137] were disconnected from the networks of evidence for all outcomes and were therefore excluded from all analyses.

FIGURE 5.5: Network of evidence (Incontinence)

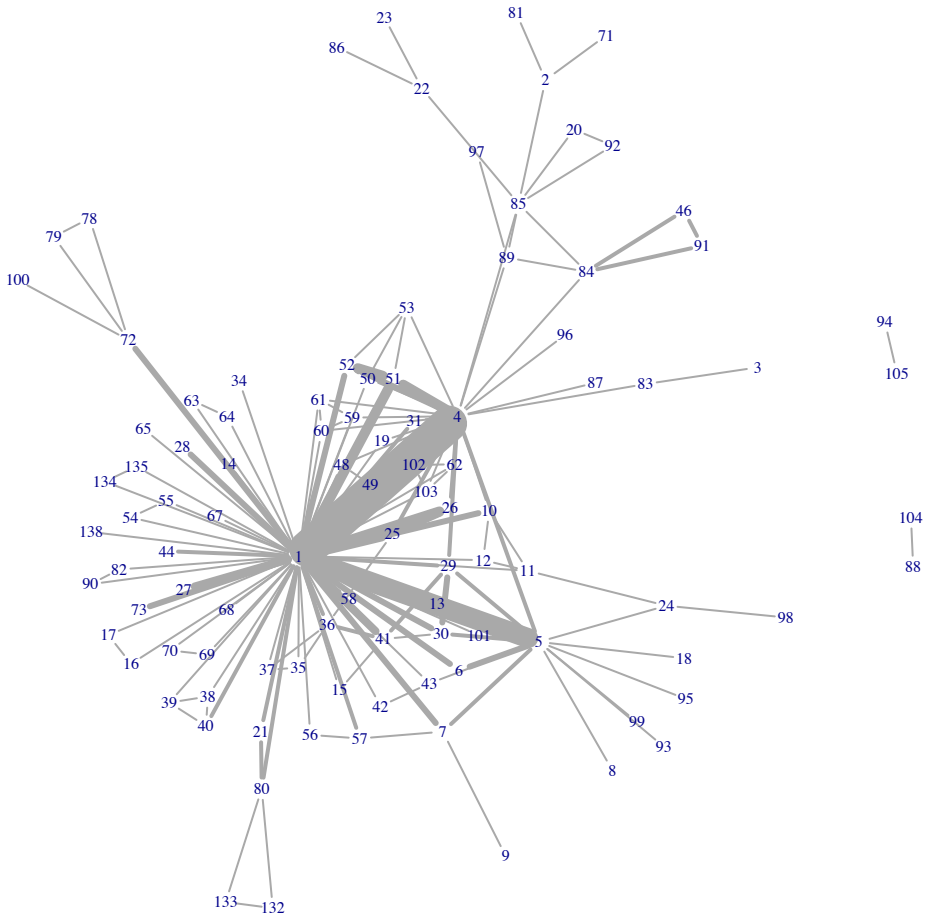


FIGURE 5.6: Network of evidence (Voiding)

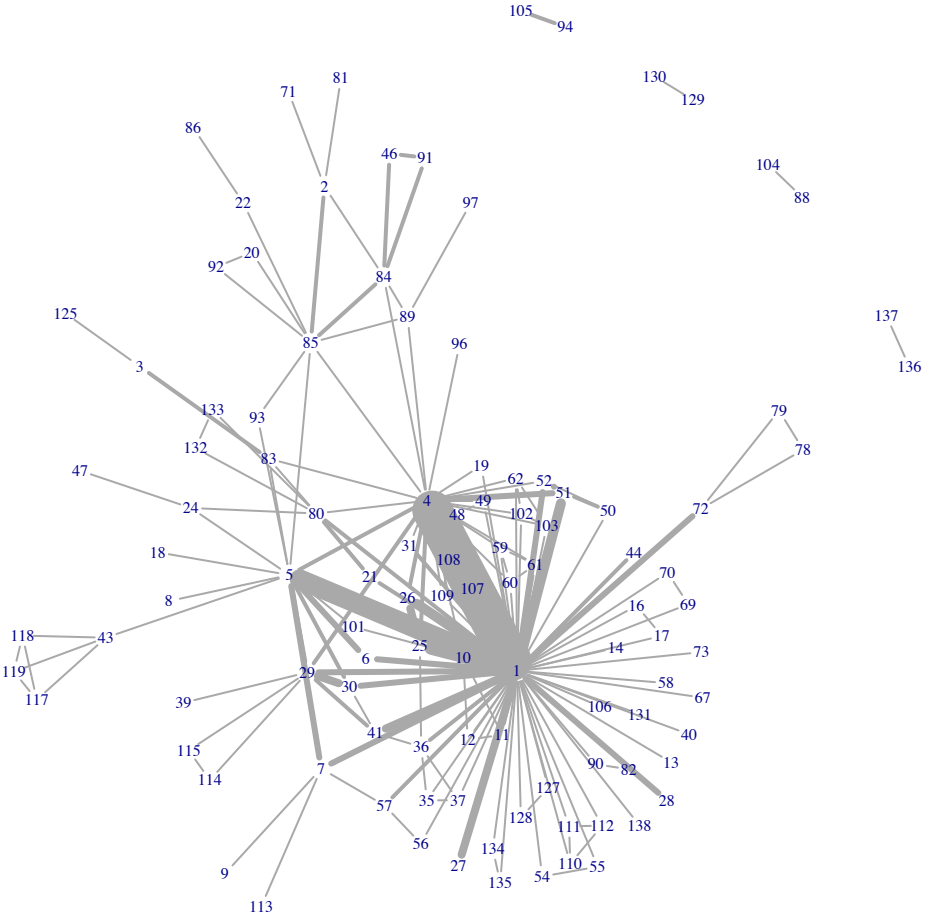


FIGURE 5.7: Network of evidence (Urgency)

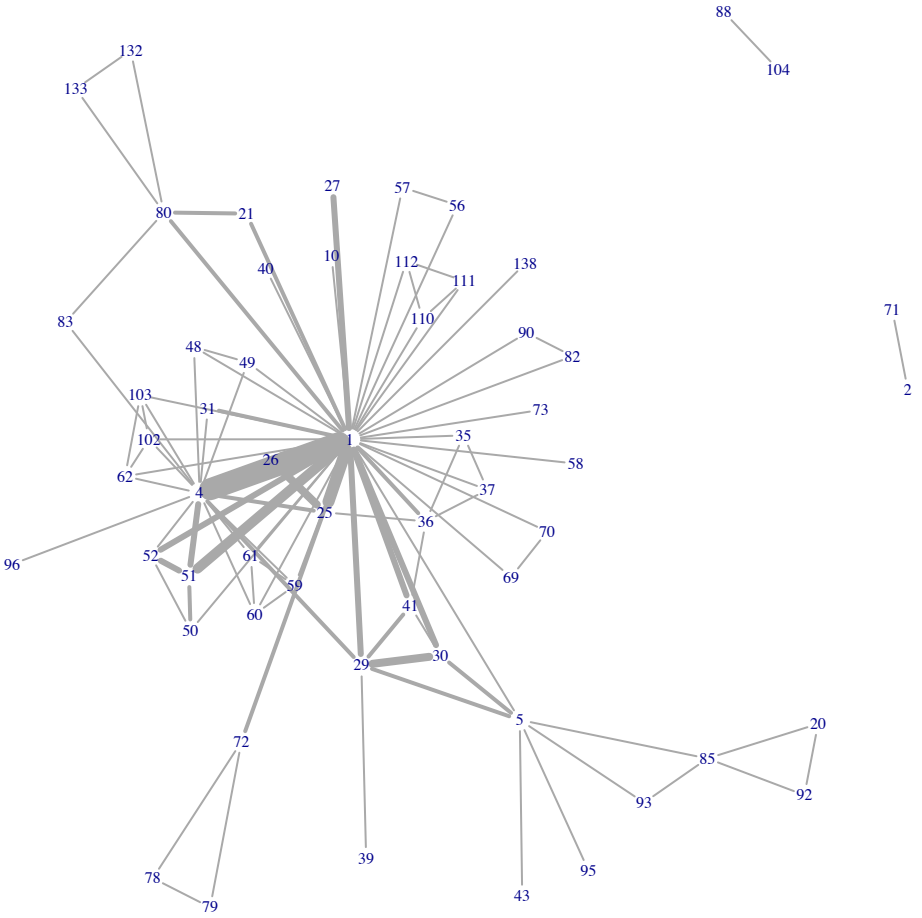


FIGURE 5.8: Network of evidence (Nocturia)

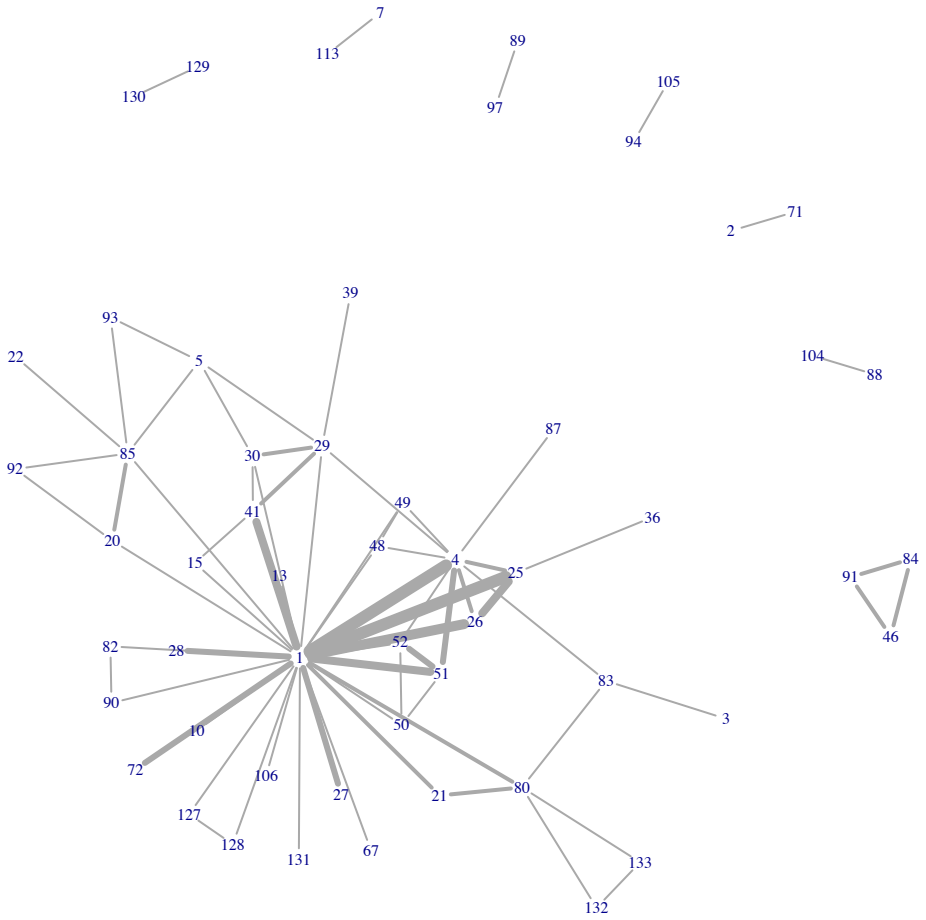


FIGURE 5.9: Network of evidence (Number of patients with adverse events)

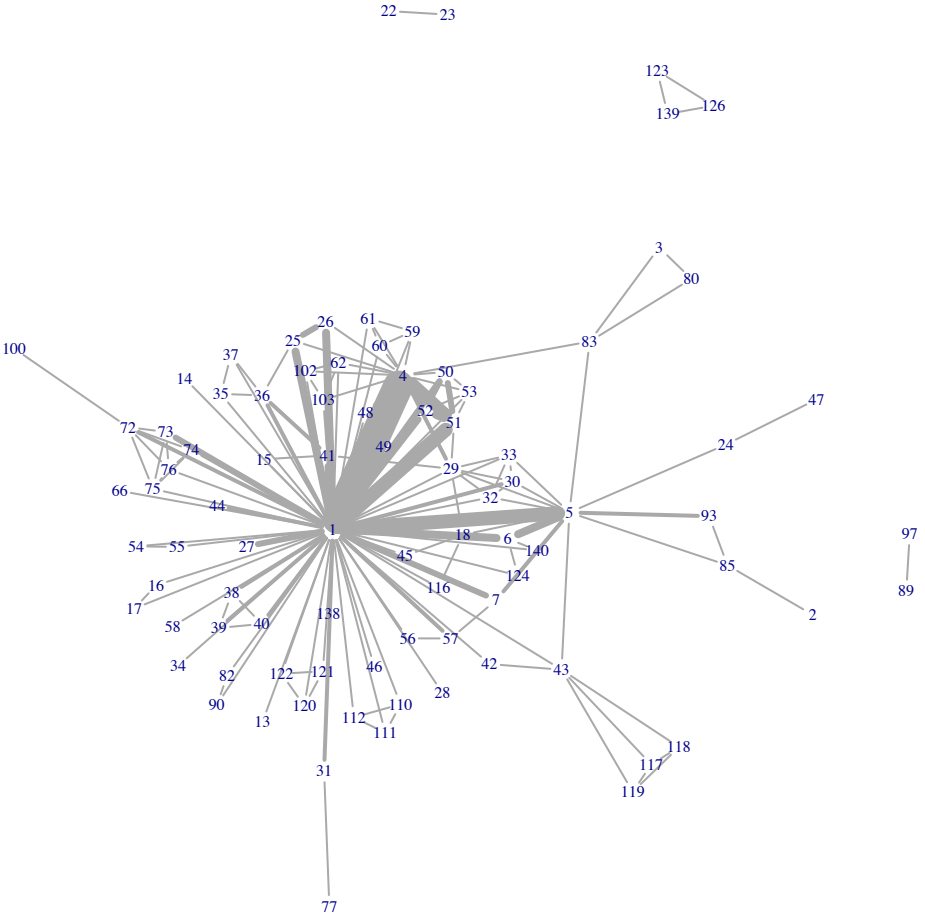


FIGURE 5.10: Network of evidence (Discontinuation due to adverse events)

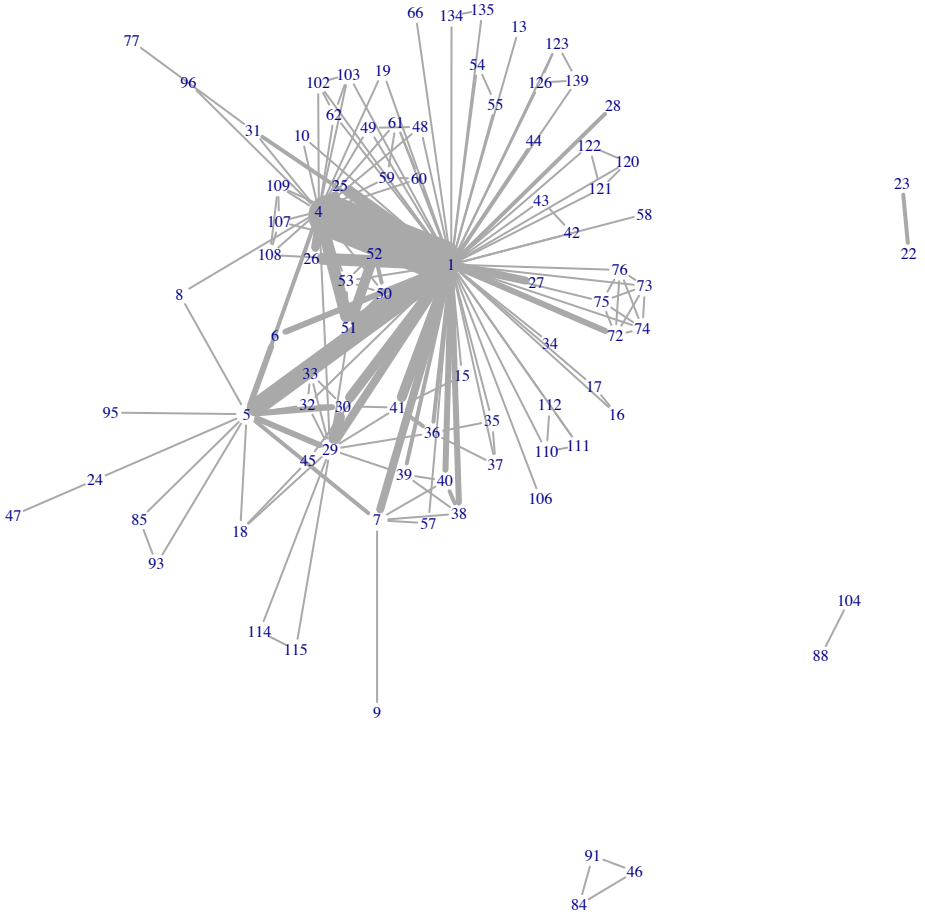
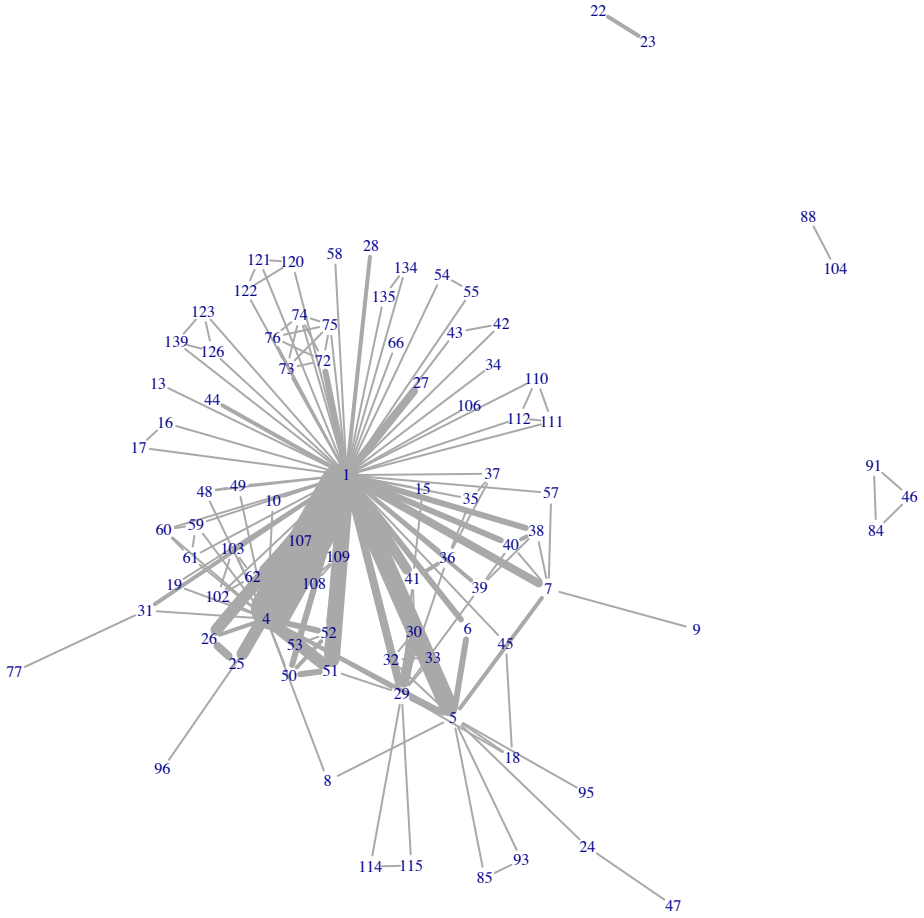


FIGURE 5.11: Network of evidence (Discontinuation due to lack of efficacy)



5.10 Discussion

Medline, EMBASE and clinicaltrials.gov were used to identify RCTs evaluating minimally invasive interventions for the management of OAB and DO. In this comprehensive systematic literature review 174 trials were identified for inclusion in network meta-analysis methods. One-hundred and forty different interventions were evaluated, and one third of studies primarily compared interventions solely with a placebo or control intervention. In total 75,355 participants with OAB and/or DO were included. Participants at baseline experienced, on average, 2.3 incontinence, 11.4 voiding, 9.2 urgency, and 2 nocturia, episodes per 24 hours. The evidence base for trials comparing interventions for patients with OAB and DO is extensive both in terms of the number of published studies, and the number of interventions evaluated. Due to vast numbers of assorted treatment comparisons, there were relatively few trials informing many of the vertices in the networks of evidence, which led to very large networks of treatment comparisons with reasonably sparse data.

In total, 117, 124, 62 and 57 studies reported treatment effects for urinary incontinence, voiding, urgency and nocturia episodes. Only 36% of papers included urgency as an outcome, despite this being regarded as the "cardinal symptom" of OAB (Cardozo et al., 2009). There is a clear element of selective reporting in the original studies with only 38% of efficacy studies reporting treatment effects for all 3 cardinal symptoms. As a result, many of the networks of evidence became disconnected. It is apparent that not all studies report treatment effects for all outcomes and therefore, in a decision making context, different interventions have to be considered in assessing different outcomes. Overall, 97, 100, 54 and 40 interventions were evaluated for incontinence, voiding, urgency and nocturia episodes, respectively. This can have significant implications for decision makers of syndromic conditions, as there is often insufficient evidence to inform complete treatment profiles across the entire symptom syndrome.

One of the most notable findings from this systematic literature review is the poor reporting of RCTs in patients with OAB and DO. Nearly one third of studies failed to report measures of variability in the observed treatment effects. As previously described in Section 3.4, in order to quantify an average effect size from multiple trial results, evidence synthesis methods require both the treatment effects and their relative uncertainty to be specified in the original studies. Trials that fail to report either measure will often be excluded from evidence synthesis methods, which in turn could have a large impact on the evidence base used for decision making. In Chapter 6, missing uncertainties will be estimated within the NMA framework using the correlation between uncertainty in baseline and follow-up estimates as described in Abrams et al. (2005). For incontinence, voiding, urgency and nocturia episodes, estimates of variability were required for 37%, 34%, 31% and 38% of included studies, respectively.

Generally, risk of bias in the eligible studies overall was mostly low, although a large number (48.3%) of trials were considered to have a high risk of bias for selective reporting. There did not appear to be any evidence of publication biases for the primary or secondary outcomes. However, two studies by Wang et al. (2006) (study id: A157) and Wang et al. (2009) (study id: A158) appeared to have unexpectedly small standard errors given the size of the sample population for urgency outcomes, though this was not apparent for all other efficacy measures. As a result, caution must be taken when synthesising evidence for urgency episodes and investigation into the impact of the two studies on the overall result should be evaluated through sensitivity analyses.

To my knowledge, this is the largest and most comprehensive systematic literature review undertaken in participants with OAB and DO to date. Previous systematic reviews have evaluated interventions of a similar nature or interventions belonging to the same class of interventions (as described in Section 2.5) (Henriet and Roumeguère, 2014; Olivera et al., 2016; Sun et al., 2015; Reynolds et al., 2015). Collating information on all conservative and minimally invasive interventions for

the management of OAB and DO allows comparison between the diverse range of interventions possible for decision making.

5.11 Chapter summary

This chapter has described a comprehensive systematic literature review of RCTs evaluating minimally invasive interventions for the management of OAB and DO in adults. As mentioned in Section 5.1, the first step in any HTA process is the acquisition of all relevant data as described in this chapter; the second step is the synthesis of this data in one coherent analysis. Many of the data limitations identified in this chapter have motivated methodological developments of evidence synthesis methods described in Chapters 6, 7 and 8. Chapter 6 addresses the issue of poor reporting - in terms of studies that fail to report uncertainty in the observed treatment effects - by incorporating all relevant information in the evidence synthesis framework using methods described in Abrams et al. (2005). Chapter 6 also explores the impact of potential confounding factors outlined in Section 5.3.3 and accounts for these factors in the analyses. Chapter 7 describes a hierarchical NMA to make use of large networks of evidence with relatively sparse data informing many of the treatment comparisons. Chapter 8 addresses the issue of selective reporting and disconnected networks of evidence by building on the work of Chapters 6 and 7 and extending the NMA framework using a multivariate approach.

Chapter 6

Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder

6.1 Chapter overview

Chapter 5 previously described the HTA process as a two stage approach; the first of which involves identifying all relevant sources of information through a systematic literature review, and the second involves quantifying this information by synthesising all of the available data. This chapter describes the second stage of this process using data identified and discussed in Chapter 5. For each outcome of interest, NMAs are used to synthesise data in a single coherent analysis. Many of the data limitations identified in Chapter 5 are explored and addressed in this chapter using various NMA techniques. The methodologies described in this chapter will be developed and extended in the remaining chapters of this thesis.

6.2 Introduction

As shown in Section 5.4, almost all published articles reporting data on interventions for OAB syndrome compare the intervention solely with placebo, which makes comparison across active interventions difficult without using indirect comparisons or NMA. This is particularly evident for trials evaluating antimuscarinic drugs. There are several published meta-analyses that have been undertaken in the field of OAB syndrome (Buser et al., 2012; Novara et al., 2008a; Luo et al., 2012; Chapple et al., 2008a; Cui et al., 2013, 2014, 2015; Nitti et al., 2013; Wu et al., 2014; Liu et al., 2014; Reynolds et al., 2015; Davis et al., 2015; Huang et al., 2015; Sun et al., 2015; Cao et al., 2016; Sweeney et al., 2016; Freemantle et al., 2016). With the exception of Buser et al. (2012) and Freemantle et al. (2016), the remaining studies compare interventions in a series of head-to-head comparisons, where studies directly evaluating the same pair of interventions, were pooled in a pairwise meta-analysis. Buser et al. (2012) compared the efficacy and adverse events of different antimuscarinic drugs using a NMA framework. Freemantle et al. (2016) assessed the relative efficacy of onaBoNT-A compared with mirabegron, but did not compare these interventions with all other treatment modalilites; and thus, in the current literature there is no coherent comparison between the diverse range of interventions for OAB and consequently, there is little information of a superior treatment for decision making.

Network meta-analyses are commonly used to evaluate multiple interventions in a single coherent analysis. An advantage of this methodology is that where direct evidence between active treatments is not available, such as treatment A and D in Figure 3.3 of Chapter 3, direct comparisons of AB and BD can be used to indirectly infer the efficacy of treatment A relative to treatment D as described in Section 3.4.5. An important assumption that underpins NMA, is that the treatment effects estimated by the AD trials, would be the same as the treatment effects estimated from the AB and BD trials had they included AD arms, and thus, the model assumes that both the direct and indirect estimates are consistent. In the case of

OAB, the networks of treatment comparisons (Figures 5.5 - 5.11 of Chapter 5) are particularly large, but there are relatively few studies to inform the vast amount of treatment comparisons. Consequently, there is little information informing many of the direct treatment comparisons. Network meta-analysis models are particularly useful in this context, where both the direct and indirect estimates can be used to obtain treatment effect estimates for every pair of interventions under consideration, even if they have not been directly compared otherwise. The NICE Guide to Methods for Technology Appraisal recommends this approach if it adds information which is not currently available from head-to-head comparisons (Dias et al., 2011b).

6.3 Methods

6.3.1 Network meta-analysis

For both primary and secondary outcomes, fixed and random effects models were estimated. Random effects NMAs assumed a common heterogeneity parameter between studies across all treatment contrasts. Efficacy outcomes can be interpreted as median reduction of episodes from baseline per 24 hours. Safety and tolerability outcomes were presented as odds ratios (OR) with 95% credible intervals (CrI). Interventions were ranked using the method described in Section 3.4.5.2. For primary outcomes, interventions were ranked from best to worst, and a rank of 1 represents the most effective intervention. For secondary outcomes, interventions were ranked from worst to best, and a rank of 1 indicates the most hazardous or least tolerable intervention.

Section 3.4.5 of Chapter 3 described the general framework for continuous outcomes using random effects NMA models accounting for multi-arm trials as described by Dias et al. (2011b) and repeated in Equation (6.1) for clarity:

$$
\begin{aligned}
y_{ij} &\sim \text{Normal}(\theta_{ij}, se_{ij}^2) \\
\theta_{ij} &= \mu_i + \delta_{i,bk} I_{\{j=k\}} \\
\text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{6.1}
$$

The study-specific relative treatment effect, $\delta_{i,bk}$, for arm $k = 2$ is expressed as:

$$\delta_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{i2}}, \tau^2)$$

For multi-arm trials of $k > 2$, the study-specific relative treatment effect, $\delta_{i,bk}$, for the k^{th} conditional distribution, $k = 3 \ldots na_i$, is defined by:

$$\delta_{i,bk}| \begin{pmatrix} \delta_{i,b2} \\ \vdots \\ \delta_{i,b(k-1)} \end{pmatrix} \sim \text{N} \left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\delta_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right)$$

The relative effect of the study-specific reference treatment in arm 1 (the control arm) relative to itself, $\delta_{i,b1}$, is set to 0 and as such the set of conditional univariate distributions begin with the relative effect of the intervention in arm 2 relative to the control arm, $\delta_{i,b2}$. The baseline intervention means, $\mu_i \in (-\infty, \infty)$, were assumed to follow a Normal$(0, 10^3)$ distribution suggesting that the mean reduction in symptoms from baseline for the reference intervention could plausibly be in the range of 0 ± 62 episodes. Basic parameters, d_{1j}, were assumed to have a Normal$(0, 10^3)$. The between-study standard deviation values of τ were assumed to follow a Uniform(0,5) prior distribution, implying that the between-study standard deviation can take any value between but not including, 0 and 5, and small values of τ are equally likely as large values. A value of 5 for example would indicate that for a random pair of studies, the difference in the mean reduction in symptoms from baseline could be as large as 5.45.

Equation (6.2) illustrates the general model for the binary outcome case as described by Lu and Ades (2004). It is these models that form the foundation for the NMA framework:

$$\begin{aligned} r_{ij} &\sim \text{Binomial}(p_{ij}, n_{ij}) \\ \text{logit}(p_{ij}) &= \mu_i + \phi_{i,bk} I_{\{j=k\}} \\ \text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{6.2}$$

The study-specific relative treatment effect, $\phi_{i,bk}$, for arm $k = 2$ is expressed as:

$$\phi_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{i2}}, \tau^2)$$

For multi-arm trials of $k > 2$, the study-specific relative treatment effect, $\phi_{i,bk}$, for the k^{th} conditional distribution, $k = 3 \ldots na_i$, is defined by:

$$\phi_{i,bk} | \begin{pmatrix} \phi_{i,b2} \\ \vdots \\ \phi_{i,b(k-1)} \end{pmatrix} \sim \text{N} \left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\phi_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right)$$

For binary data, the number of reported events for an intervention $j = 1, 2, ..., n_t$ within the i^{th} study is considered a binomial count, r_{ij}, from a sample number at risk, n_{ij}. This information allows estimation of the probability, p_{ij}, which is associated with the risk of the event in question. Logistic regression models were used where $\phi_{i,bk}$ represents the study-specific log-odds of the intervention in the k^{th} arm, t_{ik}, relative to a study-specific baseline intervention, b, in arm 1, t_{ib}, of study i. The correlation between-arms of the same study is accounted for using conditional distributions (see Section 3.4.5) where $d_{1,t_{ik}}$ represents the pooled log-odds of the intervention in arm k of study i, t_{ik}, relative to the reference intervention

for the entire treatment network, $j = 1$, and $d_{1,t_{ib}}$ denotes the log-odds of the intervention in the baseline arm, t_{ib}, relative to the reference treatment, $j = 1$. The between-study standard deviation is given by τ.

As with the continuous NMA framework described in Section 3.4.5, the baseline intervention means, μ_i, and basic parameters, d_{1j}, require prior specifications. For the binary case, $\mu_i \in (-\infty, \infty)$, were assumed to have a Normal$(0, 10^2)$ prior distribution. The basic parameters, d_{1j} for $j = 2, ...n_t$, were assumed to have plausibly vague prior distributions, such that $d_{1j} \sim$ Normal$(0, 10^2)$, and the functional parameters for treatments A and B, $d_{t_{ib}t_{ik}} = d_{AB}$ for $A > 1$ and $B > 1$, are expressed in terms of Equation (3.21) described in Chapter 3. For the binary case, the log-odds of an event for the reference intervention could plausibly be in the range of 0 ± 20. The between-study standard deviation values of τ were assumed to have a Uniform(0, 2) prior distribution for binary outcomes, respectively. Implying that the between-study standard deviation can take any value between, but not including, 0 and 2 for binary outcomes, and small values of τ are equally likely as large values. A value of 2 on the log-odds scale would suggest that the ratio of maximum to minimum odds ratios could be as large as 8.85. Sensitivity analyses considering two other variance-component prior distributions were considered: 1) gamma(0.001, 0.001) on the precision scale, that is, 1/variance, and 2) half-normal(0,1) on the standard deviation scale (see Section 3.2.4).

6.3.2 Missing data framework

In situations where clinical trial data is heterogeneously reported and/or missing, synthesising such data can be problematic. As further described in Section 3.4, NMA methods require specification of both a measure of effect and a measure of variability in the original trial reports. In the absence of such data, trials are often excluded from the analysis (Chowdhry et al., 2016). In this chapter, interest lies in evaluating the change from baseline for several symptomatic responses in OAB. Whilst symptoms at baseline and follow-up are commonly reported, change

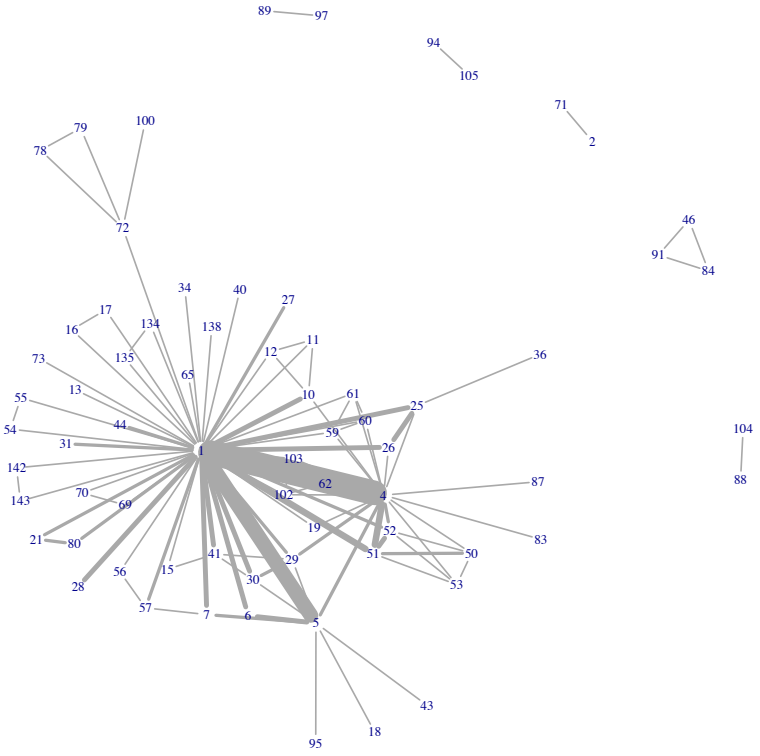
from baseline is often unreported; and thus, whilst it is possible to calculate the mean change from baseline, the variance is often missing. Table 6.1 illustrates the different scenarios for trial reporting in the included studies. In this example, there were no trials that reported follow-up data alone. In a decision making context, missing data can have a detrimental impact on the overall decision, as many of the interventions of interest have to be excluded from the analyses. For OAB data, the magnitude of this affect is illustrated in Figure 6.1, which presents the network of treatment comparisons for urinary incontinence outcomes solely for trials that reported measures of uncertainty in change from baseline (scenarios 1-4 of Table 6.1. In this particular example, it is apparent that many of the interventions became disconnected from the network of evidence, and there were far fewer interventions (59) available for analysis compared to the entire set of interventions (97) illustrated in Figure 5.5.

Table 6.1: Scenarios for trial reporting

	Change from baseline		**Baseline**		**Follow-up**	
Scenario	mean	variance	mean	variance	mean	variance
1	✓	✓	✓	✓	✓	✓
2	✓	✓	✓	✓	NR	NR
3	✓	✓	NR	NR	✓	✓
4	✓	✓	NR	NR	NR	NR
5	✓	NR	✓	✓	✓	✓
6	✓	NR	✓	✓	NR	NR
7	✓	NR	NR	NR	NR	NR

NR denotes not reported

FIGURE 6.1: Example network of treatment comparisons for trials reporting treatment effects and measures of uncertainty for urinary incontinence episodes



For this reason, when meta-analysing data, every effort was made to estimate missing standard errors, based on additional measures of uncertainty (Abrams et al., 2005). In situations where the observed standard errors of treatment effects, se_{ij}, were not reported, but baseline and follow-up variances were (scenario 5 of Table 6.1), the correlation, ρ, between variance at baseline, ${sd_{\text{baseline}_{ij}}}^2$, and follow-up, ${sd_{\text{followup}_{ij}}}^2$, were used to impute estimates of the variance for change from baseline, ${sd_{\text{change}_{ij}}}^2$, (Edwards, 1971). This is calculated as:

$${sd_{\text{change}_{ij}}}^2 = {sd_{\text{baseline}_{ij}}}^2 + {sd_{\text{followup}_{ij}}}^2 - 2\rho(sd_{\text{baseline}_{ij}} \times sd_{\text{followup}_{ij}}) \tag{6.3}$$

$${se_{ij}}^2 = \frac{{sd_{\text{change}_{ij}}}^2}{\sqrt{n_{ij}}}$$

Using external information from trials that report all variance terms (scenario 1 of Table 6.1), an informative prior distribution was placed on the correlation, ρ, using

Fisher's z-transformation (Abrams et al., 2005). For trials that do not report the variability at follow-up (scenario 6 of Table 6.1), a linear predictor with baseline variance as a covariate was included such that:

$$sd_{\text{followup}_{ij}} = \alpha + \beta(sd_{\text{baseline}_{ij}}) \tag{6.4}$$

where α represents a constant term, and β the regression coefficient. Trials that did not provide a measure of variability (scenario 7 of Table 6.1) were excluded from the analyses. All NMA models developed and applied in this thesis incorporate a missing data framework in order to maximise the information from original trial reports, and amelioriate the impact of potential reporting biases.

6.3.3 Network meta-regression and adjustment for baseline risk

As described in Section 3.4.5, combining direct and indirect information using NMA assumes an additive relationship between treatment effects. In the presence of treatment effect modifiers such as differences in patient characteristics at baseline or differences in trial designs, additivity between treatment effects may not hold. Network meta-regression accounts for these potential differences by incorporating potential confounding factors as additional covariates in the model. As previously described in Section 5.3.3, participant age and baseline symptoms were identified as potential treatment effect modifiers. Building on the missing data framework, mean participant age and baseline risk were further explored in this chapter using network meta-regression techniques. Both exchangeable and common regression coefficients were explored for mean participant age, as previously described in Section 3.4.5.1. In this setting, baseline risk represents the average response of a patient in the study-specific control group (e.g. placebo). However, using network meta-regression to incorporate baseline risk leads to structural dependence of regression models and therefore, adoption of a baseline risk model would be more appropriate. Baseline risk models for pairwise meta-analyses are

described in Section 3.4.3.2. Achana et al. (2013) extended this model to the case of mixed treatment comparisons. By accounting for baseline risk in each of the trials it is possible to compare the relative benefit, and risk, of all interventions for patients presenting with a common symptom profile. For studies without a control group, predictive posterior distributions were used to sample estimates of baseline risk had a control intervention been included (Achana et al., 2013).

6.3.4 Model computation and convergence diagnostics

Models were estimated using WinBUGS 1.4.3 (Spiegelhalter et al., 2003). Example WinBUGS code for fixed effect models incorporating a missing data framework and random effects models incorporating a missing data framework together with age-adjusted analyses, and baseline risk models, for continuous data are given in Appendix C.1 - C.4. Example WinBUGS code for fixed and random effects models for binary outcomes are given in Appendix C.5 and C.6. For all models, a 'burn-in' period of 10,000 MCMC iterations were used followed by a 150,000 iteration sampling phase (see Section 3.2.1 for further details). Sensitivity to model 'burn-in' and length of the sampling phase were assessed in sensitivity analyses. Convergence was assessed for key parameters including estimated treatment effects, and between-study standard deviations, and assessed using Brooks-Gelman-Rubin statistics, autocorrelation, history, trace, and density plots as described in Section 3.2.2. For illustration purposes, a random sample of convergence diagnostic plots for key parameters were presented in this thesis (Lunn et al., 2012). Three individual MCMC chains with disparate starting values were used.

6.3.5 Assessing inconsistencies between direct and indirect information

Inconsistencies between direct and indirect information were assessed for all pairs of treatment comparisons that belong to a closed loop in the network (i.e. there exists both direct and indirect information) using a method of node-splitting, which

was first introduced in Section 3.4.5.3. Briefly, node-splitting obtains two relative treatment effect estimates for each treatment comparison or 'node' in a network of evidence. One of these estimates is obtained from information obtained using direct treatment comparisons, and the other is obtained from information obtained using indirect treatment comparisons. Differences between the direct and indirect estimates are quantified using Bayesian p-values in order to identify conflicting evidence, beyond that attributable to chance, in the networks of treatment comparisons. Inconsistencies between direct and indirect information would suggest that the additivity assumption underpinning NMA models may not be satisfied. Sensitivity analyses investigated the impact of studies that potentially contribute to inconsistencies between direct and indirect information by individually removing studies from each of the NMAs.

6.3.6 Goodness of fit and model selection

Model fit was assessed using the posterior mean residual deviance relative to the number of unconstrained data points. If the model is of adequate fit to the data, the posterior mean residual deviance is expected to be approximately equal to the number of unconstrained data points (Spiegelhalter et al., 2002). The DIC was used for model comparison, in order to identify the best fitting model. The DIC is equivalent to the mean posterior residual deviance penalised by the effective number of parameters to be estimated in the model, and thus, the model with the lowest DIC is considered to have the best fit to the data.

6.3.7 Sensitivity analysis

Sensitivity analyses are commonly used to assess the susceptibility of results to alternative assumptions, or violation of existing assumptions. Two sets of sensitivity analyses were conducted to assess the robustness of the overall model results. The first set of sensitivity analyses evaluated the impact of the choice of prior

specification, especially on variance parameters, where two alternative distributions were considered: 1) Gamma(0.001, 0.001) on the precision scale, and 2) Half-normal(0,1) on the standard deviation scale. The second set of sensitivity analyses assessed the influence of incorporating studies with potentially biased treatment effects, as identified by Section 5.7.1, on the overall results.

6.4 Results

Table 6.2 displays the model fit and DIC statistics for all models under consideration. The statistics presented here were used to select the best fitting models from fixed and random effects analyses for each of the outcomes of interest. For efficacy outcomes, the DIC and residual deviance statistics were further used to choose between unadjusted, age adjusted, and baseline risk adjusted models.

TABLE 6.2: Model fit statistics

		N	Residual deviance	DIC	SD (95%CrI)
Efficacy outcomes					
Incontinence episodes	**FE model**	297	341.3	3360.79	
	RE model	297	290.2	3336.39	0.16 (0.10,0.23)
	Age adjusted†	285	285.3	3271.26*	0.16 (0.10,0.23)
	Age adjusted††	285	288.1	3291.80*	0.17 (0.09,0.25)
	Baseline risk	297	332.6	3334.14	0.16 (0.11,0.23)
Voiding episodes	**FE model**	307	368.4	3541.22	
	RE model	307	339.4	3530.21	0.14 (0.06,0.22)
	Age adjusted†	307	295.6	3427.23	0.13 (0.06,0.21)
	Age adjusted††	307	289.1	3444.77	0.09 (0.01,0.19)
	Baseline risk	307	380.6	3527.05	0.14 (0.07,0.22)
Urgency episodes	**FE model**	161	304.9	1876.15	
	RE model	161	180.6	1785.21	0.61 (0.41,0.86)
	Age adjusted†	158	157	1727.14*	0.30 (0.20,0.43)
	Age adjusted††	158	156.1	1728.80*	0.30 (0.19,0.44)
	Baseline risk	161	198.7	1687.50	0.60 (0.41,0.86)
Nocturia episodes	**FE model**	125	130.9	1141.63	
	RE model	125	129.9	1143.83	0.03 (0.001,0.07)
	Age adjusted†	122	122.5	1122.47*	0.03 (0.001,0.07)
	Age adjusted††	122	114.4	1129.90*	0.03 (0.002,0.09)
	Baseline risk	125	150.1	1148	0.02 (0.001,0.07)
Safety and tolerability outcomes					
Number of patients with adverse events	**FE model**	243	318.8	1660.64	
	RE model	243	257.0	1628.96	0.23 (0.14,0.33)
Discontinuations due to adverse events	**FE model**	273	266.8	1432.26	
	RE model	273	264.6	1434.04	0.07 (0.002,0.22)
Discontinuations due to a lack of efficacy	**FE model**	158	169.5	690.364	
	RE model	158	156.2	687.39	0.40 (0.06,0.75)

FE, unadjusted fixed effect; RE, unadjusted random effects; N, number of unconstrained datapoints; DIC, deviance information criterion; SD, between-study standard deviation.

†Assuming a common regression coefficient.

††Assuming exchangeable regression coefficients.

*Denotes that the DIC was calculated on fewer studies and are therefore not comparable to other models. The DIC for these models are displayed solely for completeness.

6.4.1 Network meta-analysis of efficacy outcomes

6.4.1.1 Incontinence

For incontinence episodes, the random effects model appeared to provide a better fit to the data than that of the fixed effect model (Table 6.2), with corresponding DIC statistics of 3336.39 compared to 3360.79, respectively. The between-study standard deviation estimated from the random effects model was 0.16 (95% CrI: 0.10, 0.23). Age did not appear to be an important factor in age adjusted analyses as the 95%CrI for the common β coefficient (as further described in Section 3.4.3.1) included the point of no difference (β: 0.07, 95%CrI: -0.04,0.20). There is little difference between the DIC statistics for the unadjusted random effects model (DIC: 3336.39), compared to the model adjusted for baseline risk (DIC: 3334.14). Residual deviance statistics would suggest that the unadjusted model is of a slightly better fit to the data than that of the baseline risk analysis. The total residual deviance for the unadjusted random effects model was 290.2, which was closer to the number of unconstrained data points (N=297), compared to that of the baseline risk analysis with a residual deviance of 332.6. Therefore, treatment effect estimates presented in this chapter were based on the results of the unadjusted random effects NMA.

Table 6.3 displays the treatment effect estimates, estimated treatment rankings, and the probability that each of the interventions evaluating incontinence episodes was the best. Sacral nerve stimulation appeared to be the most effective intervention for reducing incontinence episodes with a posterior median change from baseline of -8.72 (95%CrI: -11.33, -6.09) episodes, relative to placebo. Sacral nerve stimulation was considerably more effective in reducing incontinence episodes than all other interventions under consideration, with an estimated mean ranking of 1 (95%CrI: 1,1) and 100% of model iterations ranking SNS in first place. Though it is worth noting that SNS was evaluated in a single study of only 34 patients whose mean baseline incontinence episodes were noticeably larger (mean: 9.7, SD:

not reported) than that of the entire cohort (mean: 2.8, SD: 1.5) and thus these results should be interpreted with caution. OnaBoNT-A 200U sparing the trigone ranked in second place (mean rank: 3, 95%CrI: 2,9) with a median difference of -2.3 (95% CrI: -3.16,-1.42) incontinence episodes relative to placebo. OnaBoNT-A 200U was evaluated by 3 studies including a total of 114 participants.

TABLE 6.3: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from unadjusted random effects network meta-analysis

Treatment Pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95% CrI)	Rank (95%CrI)	p(Best)
	Sacral nerve stimulation	[81]	1	34	-8.72 (-11.33,-6.09)	1 (1,1)	1
	OnaBoNT-A 200U trigone sparing	[73]	3	114	-2.3 (-3.16,-1.42)	3 (2,9)	0
	Oxybutynin IR 2.5mg b.i.d + salivary pastilles	[98]	1	8	-2.2 (-4.06,-0.36)	4 (2,52)	0
	Solifenacin/trospium + placebo injection	[100]	1	118	-1.97 (-2.96,-1.01)	5 (2,14)	0
	Electrostimulation + PFMT + BT	[97]	1	25	-1.93 (-2.94,-0.91)	5 (2,17)	0
	OnaBoNT-A 100u trigone sparing	[72]	5	716	-1.88 (-2.31,-1.45)	6 (3,9)	0
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-1.63 (-2.73,-0.54)	7 (2,38)	0
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-1.39 (-4,1.06)	9 (2,95)	0
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-1.29 (-1.7,-0.89)	10 (6,19)	0
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	-1.19 (-2.49,0.1)	11 (3,77)	0
	Mirabegron IR 100mg b.i.d	[48]	1	37	-1.12 (-1.98,-0.25)	12 (5,59)	0
	Oxybutynin ER 10mg q.d	[8]	1	185	-0.98 (-1.55,-0.42)	15 (7,48)	0
	Oxybutynin IR 3mg t.i.d	[19]	1	244	-0.9 (-1.28,-0.52)	18 (9,41)	0
	Solifenacin ER 10mg q.d	[30]	4	870	-0.88 (-1.14,-0.63)	18 (11,33)	0
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-0.82 (-1.43,-0.21)	21 (8,61)	0
	Propiverine ER 30mg q.d	[42]	1	391	-0.79 (-1.31,-0.28)	23 (9,58)	0
	Tolterodine ER 4mg q.d + BT	[87]	1	154	-0.78 (-1.45,-0.12)	23 (8,67)	0
	Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	2	32	-0.78 (-1.71,0.16)	23 (7,77)	0
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	3	150	-0.75 (-1.45,-0.01)	25 (8,72)	0
	Darifenacin ER 30mg q.d	[38]	1	115	-0.74 (-1.36,-0.12)	25 (9,67)	0
	Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg twice/week	[99]	1	40	-0.75 (-1.24,-0.25)	25 (10,60)	0
	Solifenacin ER 5 - 10mg q.d	[31]	3	1312	-0.73 (-1,-0.45)	26 (14,46)	0
	Solifenacin ER 5mg q.d	[29]	6	725	-0.71 (-0.96,-0.46)	27 (15,46)	0
	Oxybutynin IR 5mg t.i.d	[7]	4	194	-0.71 (-1.18,-0.25)	27 (11,60)	0
	Fesoterodine ER 8mg q.d	[26]	5	2266	-0.69 (-0.89,-0.5)	28 (17,43)	0
	Oxybutynin gel 56mg/day	[135]	1	198	-0.69 (-1.59,0.18)	28 (7,80)	0
	Mirabegron ER 25mg q.d	[50]	3	555	-0.67 (-0.99,-0.35)	29 (15,53)	0
	Terodiline 25mg b.i.d	[28]	3	126	-0.66 (-1.21,-0.1)	30 (11,68)	0
	Trospium chloride ER 60mg q.d	[44]	2	565	-0.66 (-1.09,-0.24)	30 (12,61)	0
	Cizolirtine citrate 400mg b.i.d	[57]	2	68	-0.63 (-1.17,-0.08)	32 (11,68)	0
	Mirabegron ER 100mg q.d	[52]	5	1515	-0.62 (-0.83,-0.41)	33 (20,50)	0
	Solifenacin ER 5 - 15mg q.d	[34]	1	377	-0.61 (-1.11,-0.1)	34 (12,68)	0
	Solabegron IR 125mg b.i.d	[55]	1	85	-0.6 (-0.94,-0.26)	34 (16,59)	0
	Elocalcitol 75mg	[70]	1	84	-0.6 (-1.2,0.01)	34 (11,73)	0
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	37	-0.6 (-1.66,0.46)	34 (7,88)	0
	Propiverine IR 15mg b.i.d	[43]	2	495	-0.57 (-1,-0.16)	36 (15,65)	0
	Mirabegron ER 200mg q.d	[53]	1	166	-0.57 (-1.13,-0.01)	36 (12,72)	0
	Mirabegron ER 50mg q.d	[51]	8	2081	-0.57 (-0.76,-0.38)	37 (24,52)	0
	Darifenacin ER 15mg q.d.	[40]	2	319	-0.51 (-0.97,-0.05)	41 (16,71)	0
	Mirabegron IR 150mg b.i.d	[49]	1	41	-0.51 (-1.57,0.54)	41 (8,90)	0
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.5 (-1.02,0.02)	42 (14,74)	0
	Tolterodine ER 4mg q.d	[4]	27	7521	-0.51 (-0.61,-0.4)	42 (32,52)	0
	Fesoterodine ER 4mg q.d	[25]	6	2597	-0.47 (-0.66,-0.29)	44 (30,59)	0
	Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	1	130	-0.48 (-0.79,-0.17)	44 (22,65)	0
	Tolterodine IR 2mg b.i.d	[5]	18	3312	-0.45 (-0.6,-0.3)	46 (34,58)	0
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	-0.45 (-1.07,0.18)	46 (13,81)	0
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	-0.44 (-1.11,0.23)	47 (12,82)	0
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	-0.43 (-1.08,0.22)	48 (13,82)	0
	Propiverine ER 20mg q.d	[41]	6	1381	-0.41 (-0.61,-0.2)	49 (33,64)	0
	Imidafenacin 0.1mg b.i.d	[36]	4	563	-0.4 (-0.7,-0.11)	50 (27,68)	0
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.4 (-1.83,1.03)	50 (6,95)	0
	Elocalcitol 150mg	[69]	1	87	-0.4 (-1.03,0.22)	50 (14,82)	0
	Oxybutynin trandermal 3.9mg/day	[10]	3	292	-0.33 (-0.67,0)	54 (29,73)	0
	Tolterodine IR 1mg b.i.d	[6]	3	250	-0.33 (-0.69,0.02)	54 (28,75)	0
	Oxbutynin patch 73.5mg	[15]	1	391	-0.31 (-0.67,0.05)	56 (29,76)	0
	Fesoterodine ER 4 - 8mg q.d	[27]	4	1485	-0.28 (-0.52,-0.05)	57 (39,72)	0
	Oxybutynin gel 84mg/day	[134]	1	211	-0.29 (-1.13,0.52)	57 (13,90)	0
	Darifenacin ER 7.5mg q.d	[39]	1	108	-0.27 (-0.84,0.3)	58 (21,85)	0
	Oxybutynin ER 15mg q.d	[9]	1	53	-0.28 (-1.41,0.86)	58 (9,94)	0
	Duloxetine 40mg b.i.d	[65]	1	81	-0.26 (-0.75,0.25)	59 (24,83)	0
	PFMT + BT	[89]	2	67	-0.26 (-0.87,0.37)	59 (20,83)	0
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.24 (-0.77,0.27)	60 (24,84)	0
	Solabegron IR 50mg b.i.d	[54]	1	88	-0.2 (-0.54,0.14)	62 (39,80)	0
	Tolterodine IR 2mg b.i.d + BT	[93]	1	141	-0.14 (-0.81,0.52)	65 (22,90)	0
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.13 (-0.85,0.61)	66 (20,91)	0
	ONO-8539 100mg b.i.d	[60]	1	83	-0.11 (-0.85,0.63)	67 (20,91)	0
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-0.08 (-0.39,0.23)	68 (50,83)	0
	ZD0947IL 25mg/day	[58]	1	92	-0.1 (-0.94,0.76)	68 (17,93)	0
	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	3	70	-0.08 (-0.72,0.55)	68 (28,88)	0
	Lipo-BoNTA 200U	[138]	1	29	-0.06 (-0.96,0.83)	69 (16,94)	0
	Oxybutynin transdermal 1.3mg/day	[11]	2	166	-0.03 (-0.66,0.61)	71 (30,91)	0
	Placebo	[1]	77	14282	NA	72 (65,81)	0
	Estradiol 25mg	[68]	1	20	0 (-0.38,0.38)	72 (51,87)	0
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	47	0 (-0.57,0.56)	72 (37,91)	0
	Pregabalin 150mg b.i.d	[62]	1	41	0 (-0.88,0.9)	73 (19,95)	0
	Bladder Training (BT)/Behaviour Therapy	[85]	4	147	0.02 (-0.59,0.62)	73 (37,88)	0
	Oxybutynin ER 5 - 30mg/day	[22]	3	109	0.09 (-0.62,0.78)	76 (35,91)	0
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	0.07 (-0.45,0.59)	76 (46,91)	0
	Electrostimulation	[80]	3	150	0.08 (-0.28,0.44)	76 (58,88)	0

TABLE 6.3: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from unadjusted random effects network meta-analysis (cont.)

	Tarafenacin 0.2mg q.d	[90]	1	77	0.11 (-0.63,0.86)	77 (33,94)	0
	Percutaneous tibial nerve stimulation	[83]	2	59	0.18 (-1.16,1.54)	80 (12,96)	0
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	0.22 (-0.94,1.4)	81 (17,96)	0
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	0.29 (-0.37,0.95)	84 (53,95)	0
	Oxybutynin IR 5mg b.i.d	[18]	1	116	0.35 (-0.3,0.99)	85 (57,96)	0
	Control	[2]	3	142	0.33 (-0.67,1.31)	85 (31,96)	0
	Emepronium bromide ER 200mg q.d	[63]	1	19	0.34 (-0.33,1.01)	85 (54,96)	0
	Flavoxate chloride 200mg q.d	[64]	1	19	0.33 (-0.35,1.01)	85 (53,96)	0
	Reflexology	[71]	1	54	0.33 (-0.76,1.41)	85 (25,96)	0
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	0.38 (-0.15,0.91)	86 (65,95)	0
	Oxybutynin IR 5 - 20mg	[23]	1	52	0.46 (-0.9,1.84)	88 (18,97)	0
	Oxybutynin ER 2.5mg q.d	[20]	1	16	0.52 (-0.5,1.55)	89 (44,97)	0
	Trospium chloride IR 15mg t.i.d	[46]	2	30	0.52 (-0.41,1.41)	89 (51,97)	0
	ONO-8539 300mg b.i.d	[61]	1	82	0.51 (-0.23,1.25)	89 (61,97)	0
	ONO-8539 30mg b.i.d	[59]	1	87	0.58 (-0.18,1.33)	90 (63,97)	0
	Resiniferatoxin 50nM	[67]	1	34	0.58 (-1.08,2.28)	90 (13,97)	0
	Sham therapy	[3]	1	17	0.83 (-0.62,2.27)	94 (34,97)	0
	Oxybutynin ER 5-30mg q.d + BT	[86]	1	32	0.92 (-0.31,2.12)	94 (58,97)	0

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

6.4.1.2 Voiding

DIC statistics presented in Table 6.2 suggest that the random effects model appeared to have a more appropriate fit to the data than that of fixed effect model, with DIC values of 3530.21 compared to 3541.22, respectively. A baseline risk model appeared to have no better fit to the data (DIC: 3527.05) than that of the unadjusted random effects model. Mirroring the factors associated with response to treatment identified from the RELAX trial, described briefly in Section 5.3.3 and more thoroughly in the published paper in Appendix F.2, age appeared to be an important factor in expected treatment benefit for voiding episodes. On average, with every yearly increase in age, participants were expected to have a further -0.03 (β coefficient: -0.03, 95%CrI:-0.06,-0.01) reduction in voiding episodes from baseline. Assuming a common regression coefficient for age appeared to have a better fit to the data (DIC: 3427.23) compared to that of exchangeable regression coefficients (DIC: 3444.27). Overall, age adjusted analyses assuming a common treatment by covariate interaction appeared to have the best fit to the data and therefore, results presented in this chapter were obtained from this model.

Table 6.4 illustrates the expected change from baseline in voiding episodes for patients of average age (57.5 years) in the OAB cohort, relative to a placebo intervention. Treatment effects are ranked in order of clinical efficacy, and presented with relative treatment rankings and probability that each of the interventions was the best overall. Sacral nerve stimulation appeared to be the most effective intervention for voiding episodes with an average reduction of -7.89 (95%CrI: -12.03, -3.76) episodes, relative to placebo. Sacral nerve stimulation appeared to be largely effective for the management of voiding episodes, ranking in first place for 96% of model iterations. Though it is worth noting that SNS was evaluated in a small study of only 25 patients whose mean baseline voiding episodes were noticeably larger (mean: 16.9, SD: not reported) than that of the entire cohort (mean: 11.4, SD: 1.77). Electrostimulation in combination with PFMT and BT ranked in second place (mean rank: 3, 95%CrI:2,20) with a median reduction of

-3.37 (95%CrI: -5.44,-1.23) voiding episodes relative to placebo. Though this intervention was also evaluated in a single study of 25 participants, and thus, these results should be interpreted with caution.

TABLE 6.4: Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age adjusted random effects network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95% CrI)	Rank (95% CrI)	p(Best)
	Sacral nerve stimulation	[81]	1	25	-7.89 (-12.03,-3.76)	1 (1,2)	0.96
	Electrostimulation + PFMT + BT	[97]	1	25	-3.37 (-5.44,-1.23)	3 (2,20)	0.02
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-3.12 (-4.25,-2.01)	4 (2,9)	0.01
	OnaBoNT-A 200U trigone sparing	[73]	1	86	-2.93 (-4.23,-1.61)	4 (2,14)	0.00
	PFMT + BT	[89]	2	67	-2.16 (-4.1,-0.22)	8 (3,71)	0.00
	Estradiol 3mg intravaginally	[128]	1	15	-1.87 (-3.8,-0.01)	10 (2,80)	0.00
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-1.93 (-2.7,-1.15)	10 (4,24)	0.00
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-1.79 (-3.45,-0.05)	11 (3,79)	0.00
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	-1.73 (-4.63,1.12)	12 (2,95)	0.01
	Cizolirtine citrate 400mg b.i.d	[57]	2	68	-1.6 (-2.97,-0.3)	14 (4,70)	0.00
	Propiverine IR 30mg b.i.d	[117]	1	45	-1.54 (-3.78,0.64)	15 (2,91)	0.00
	Estriol 1mg intravesically	[131]	1	21	-1.51 (-2.74,-0.27)	15 (5,71)	0.00
	OnaBoNT-A 100U trigone sparing	[72]	4	603	-1.5 (-1.9,-1.11)	15 (9,26)	0.00
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-1.48 (-2.99,0)	16 (4,80)	0.00
	Lipo-BoNTA 200U	[138]	1	29	-1.44 (-3.09,0.1)	17 (4,83)	0.00
	Reflexology	[71]	1	54	-1.4 (-3.31,0.49)	17 (3,89)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	-1.41 (-2.45,-0.42)	17 (6,63)	0.00
	Oxybutynin ER 10mg q.d	[8]	1	185	-1.33 (-1.99,-0.69)	19 (8,48)	0.00
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-1.25 (-2.06,-0.43)	21 (8,65)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	63	-1.2 (-2.13,-0.29)	22 (8,71)	0.00
	Solifenacin ER 10mg q.d	[30]	5	1376	-1.16 (-1.41,-0.93)	23 (15,34)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	-1.18 (-1.98,-0.39)	23 (9,67)	0.00
	Electrostimulation	[80]	6	236	-1.15 (-1.86,-0.47)	24 (10,61)	0.00
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-1.15 (-2.02,-0.31)	24 (9,69)	0.00
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	96	-1.11 (-1.81,-0.42)	25 (10,65)	0.00
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-1.07 (-3.04,0.81)	26 (4,93)	0.00
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-1.05 (-2.67,0.58)	27 (6,91)	0.00
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	-1 (-1.74,-0.26)	29 (11,72)	0.00
	Fesoterodine ER 8mg q.d	[26]	5	2319	-1.01 (-1.22,-0.8)	29 (19,42)	0.00
	Solabegron IR 125mg b.i.d	[55]	1	85	-0.9 (-1.19,-0.61)	34 (20,55)	0.00
	Mirabegron ER 50mg q.d	[51]	6	2014	-0.8 (-1.01,-0.6)	39 (26,56)	0.00
	Pregabalin 150mg b.i.d	[62]	1	94	-0.81 (-1.35,-0.27)	39 (17,72)	0.00
	Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	1	130	-0.8 (-1.29,-0.32)	39 (19,70)	0.00
	Oxybutynin IR 3mg t.i.d	[19]	1	244	-0.8 (-1.33,-0.26)	40 (18,73)	0.00
	Solifenacin ER 5-10mg q.d	[31]	3	1431	-0.79 (-1.08,-0.51)	40 (24,62)	0.00
	Trospium chloride ER 60mg q.d	[44]	2	565	-0.77 (-1.17,-0.38)	41 (22,68)	0.00
	Mirabegron ER 25mg q.d	[50]	2	619	-0.76 (-1.1,-0.43)	42 (24,66)	0.00
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	3	842	-0.73 (-1.39,-0.1)	44 (18,77)	0.00
	Fesoterodine ER 4mg q.d	[25]	7	3019	-0.73 (-0.92,-0.54)	44 (31,60)	0.00
	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	4	83	-0.71 (-2.14,0.77)	46 (9,93)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.7 (-1.21,-0.19)	47 (21,75)	0.00
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	-0.7 (-1.4,0)	47 (17,81)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	4	1485	-0.69 (-0.91,-0.48)	47 (31,64)	0.00
	Solifenacin ER 5mg q.d	[29]	10	1270	-0.69 (-0.92,-0.5)	47 (31,63)	0.00
	Propiverine ER 20mg q.d	[41]	5	1144	-0.69 (-0.92,-0.47)	47 (31,64)	0.00
	Mirabegron ER 100mg q.d	[52]	3	1136	-0.7 (-0.94,-0.46)	47 (30,65)	0.00
	Oxybutynin gel 84mg/day	[134]	1	211	-0.7 (-1.35,-0.05)	47 (18,79)	0.00
	Tolterodine IR 2mg b.i.d	[5]	18	3173	-0.67 (-0.86,-0.5)	49 (35,62)	0.00
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.66 (-1.55,0.23)	50 (14,86)	0.00
	Oxybutynin IR 5mg t.i.d	[7]	6	355	-0.65 (-1.09,-0.2)	51 (25,74)	0.00
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	-0.65 (-1.5,0.17)	51 (16,86)	0.00
	Tolterodine IR 1mg b.i.d	[6]	3	279	-0.63 (-1.11,-0.16)	52 (24,76)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.62 (-1.26,0.02)	53 (20,82)	0.00
	Tolterodine ER 4mg q.d	[4]	27	7085	-0.62 (-0.74,-0.51)	53 (42,64)	0.00
	Darifenacin ER 7.5mg q.d	[39]	1	29	-0.6 (-2.15,0.95)	54 (8,95)	0.00
	Trospium chloride IR 45mg t.i.d	[47]	1	615	-0.6 (-1.35,0.12)	54 (19,84)	0.00
	Oxybutynin trandermal 3.9mg/day	[10]	3	292	-0.57 (-0.96,-0.18)	56 (30,76)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	97	-0.51 (-1.05,0.03)	60 (27,82)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	3	507	-0.5 (-0.82,-0.18)	61 (38,76)	0.00
	Solabegron IR 50mg b.i.d	[54]	1	88	-0.5 (-0.79,-0.21)	61 (39,75)	0.00
	Propiverine IR 45mg t.i.d	[118]	1	48	-0.46 (-2.59,1.61)	63 (6,96)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	6	234	-0.43 (-1.42,0.49)	64 (20,89)	0.00
	Darifenacin ER 15mg q.d	[40]	1	212	-0.4 (-1.25,0.45)	65 (20,91)	0.00
	Serlopitant 4mg q.d	[109]	1	114	-0.41 (-0.9,0.08)	65 (33,84)	0.00
	Serlopitant 0.25mg q.d	[107]	1	110	-0.41 (-0.9,0.08)	65 (33,84)	0.00
	Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	2	32	-0.4 (-3.71,2.64)	66 (3,98)	0.00
	Terodiline 25mg b.i.d	[28]	3	126	-0.34 (-0.81,0.11)	68 (38,85)	0.00
	Propiverine IR 15mg b.i.d	[43]	2	145	-0.35 (-1.23,0.56)	68 (22,91)	0.00
	Netupitant 200mg q.d	[112]	1	55	-0.32 (-1.47,0.85)	69 (16,94)	0.00
	Oxybutynin 20mg intravesically q.d	[106]	1	21	-0.31 (-1.64,1.02)	70 (13,95)	0.00
	Elocalcitol 150mg	[69]	1	87	-0.29 (-1.07,0.49)	70 (26,91)	0.00
	Oxybutynin gel 56mg/day	[135]	1	198	-0.3 (-0.98,0.37)	70 (30,90)	0.00
	Control	[2]	4	138	-0.27 (-1.74,1.22)	71 (14,96)	0.00
	Oxybutynin ER 5-30mg q.d + BT	[86]	1	32	-0.22 (-2.32,1.87)	73 (8,97)	0.00
	Oxybutynin ER 5 - 30mg/day	[22]	2	92	-0.23 (-1.47,1.04)	73 (18,95)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.21 (-1.7,1.26)	73 (12,96)	0.00
	Netupitant 100mg q.d	[111]	1	59	-0.19 (-1.27,0.93)	74 (20,95)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	77	-0.15 (-1.07,0.76)	75 (27,94)	0.00
	Serlopitant 1mg q.d	[108]	1	110	-0.11 (-0.61,0.38)	77 (54,90)	0.00

TABLE 6.4: Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age adjusted random effects network meta-analysis (cont.)

	Elocalcitol 75mg	[70]	1	84	-0.1 (-0.85,0.64)	77 (36,93)	0.00
	Oxybutynin transdermal 1.3mg/day	[11]	1	128	-0.08 (-0.73,0.58)	78 (45,92)	0.00
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	-0.08 (-0.7,0.55)	78 (47,92)	0.00
	Netupitant 50mg q.d	[110]	1	60	-0.09 (-1.19,0.99)	78 (23,95)	0.00
	Percutaneous tibial nerve stimulation	[83]	5	198	-0.08 (-1.41,1.13)	78 (18,95)	0.00
	Electromagnetic stimulation	[125]	1	33	-0.07 (-2.09,1.94)	78 (9,97)	0.00
	ONO-8539 100mg b.i.d	[60]	1	83	-0.05 (-0.8,0.71)	79 (40,93)	0.00
	Oxybutynin ER 15mg q.d	[9]	1	53	-0.04 (-1.22,1.16)	79 (21,96)	0.00
	Placebo	[1]	77	14550	NA	81 (73,87)	0.00
	Resiniferatoxin 50nM	[67]	1	34	0.06 (-1.22,1.41)	82 (22,97)	0.00
	Oxybutynin IR 5mg b.i.d	[18]	1	116	0.12 (-0.88,1.13)	84 (35,96)	0.00
	ONO-8539 300mg b.i.d	[61]	1	82	0.17 (-0.59,0.93)	85 (55,95)	0.00
	Propantheline Bromide 15mg t.i.d	[113]	1	23	0.33 (-0.84,1.45)	88 (38,97)	0.00
	Propiverine ER 60mg q.d	[119]	1	39	0.45 (-1.78,2.69)	90 (13,98)	0.00
	ONO-8539 30mg b.i.d	[59]	1	87	0.47 (-0.27,1.19)	90 (71,96)	0.00
	Estradiol 1mg intravaginally	[127]	1	15	0.49 (-1.18,2.17)	91 (23,97)	0.00
	ZD0947IL 25mg/day	[58]	1	92	0.71 (-0.58,2.03)	93 (56,97)	0.00
	Sham therapy	[3]	3	157	0.94 (-0.53,2.35)	95 (61,98)	0.00
	Naftopidil 25mg q.d	[114]	1	22	3.24 (0.7,5.57)	98 (93,100)	0.00
	Trospium chloride IR 15mg t.i.d	[46]	2	30	4.59 (1.8,7.32)	99 (98,100)	0.00
	Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	1	21	5.19 (2.57,7.7)	100 (98,100)	0.00

† median relative to a placebo intervention. Results are based on patients of average age.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

6.4.1.3 Urgency

For urgency episodes, the random effects model was a substantially better fit to the data than that of the fixed effect model (Table 6.2), with corresponding DIC statistics of 1785.21 and 1876.15, respectively. This is further illustrated through the estimate of between-study standard deviation (0.61, 95%CrI: 0.41, 0.86), suggesting that for any random pair of studies, the mean change from baseline in urgency episodes can differ by up to 0.65 of an episode. Age did not appear to be an important factor in age adjusted analyses as the 95% credible interval for the common β coefficient contained the point of no difference (β=-0.03, 95%CrI: -0.02, 0.09). The baseline risk model appeared to be the best fitting model with a considerably lower DIC statistic of 1687.50 compared to the random effects model; and thus, the results presented here were based on data obtained from the baseline risk model.

Table 6.5 displays the estimated posterior median difference in urgency episodes from baseline relative to a placebo intervention. The results displayed in this table represent the estimated posterior median treatment effect in patients presenting with average baseline symptoms relative to all patients in the OAB cohort. Interventions were ranked based on clinical efficacy from best to worst. The relative ranking together with the probability that the intervention was the best is given in Table 6.5. Electrostimulation with vaginal oestrogen cream 1.25mg daily, as a combination therapy, appeared to be the most effective intervention for reducing urgency episodes (-6.94, 95%CrI: -8.65, -5.23). This combination therapy dominated the analysis, ranking in first place (rank 1, 95% CrI: 1,1) for 99% of model iterations. Though electrostimulation with vaginal oestrogen cream was evaluated in a single study of 102 participants. Electrostimulation ranked in second place (rank 2, 95% CrI: 2,4) with estimated median reduction of -4.84 (95% CrI: -5.94, -3.74) urgency episodes relative to placebo. Electrostimulation as an independent therapy was evaluated in 4 studies with a total of 161 participants.

TABLE 6.5: Estimated posterior median difference (and 95% credible interval) in change from baseline for urgency episodes relative to placebo obtained from random effects network meta-analysis adjusted for baseline risk

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95%CrI)	Rank (95%CrI)	p(Best)
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-6.94 (-8.65,-5.23)	1 (1,1)	0.99
	Electrostimulation	[80]	4	161	-4.84 (-5.94,-3.74)	2 (2,4)	0.00
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	-3.34 (-5.06,-1.62)	4 (3,15)	0.00
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-2.69 (-4.83,-0.54)	6 (2,36)	0.00
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-2.56 (-5.14,0.04)	7 (2,46)	0.00
	Cizolirtine citrate 400mg b.i.d	[57]	1	16	-2.62 (-4.66,-0.59)	7 (2,35)	0.00
	Percutaneous tibial nerve stimulation	[83]	2	52	-2.5 (-4.37,-0.64)	7 (3,34)	0.00
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-2.38 (-3.82,-0.94)	8 (3,27)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	1	86	-2.25 (-3.98,-0.51)	9 (3,37)	0.00
	OnaBoNT-A 100u trigone sparing	[72]	3	592	-2.07 (-3.06,-1.08)	10 (5,25)	0.00
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-1.94 (-3.02,-0.82)	11 (5,31)	0.00
	Darifenacin ER 7.5mg q.d	[39]	1	29	-1.68 (-3.34,-0.02)	14 (4,45)	0.00
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-1.48 (-3.08,0.12)	16 (5,47)	0.00
	Fesoterodine ER 8mg q.d	[26]	5	2319	-1.47 (-2.03,-0.9)	17 (10,29)	0.00
	Lipo-BoNTA 200U	[138]	1	29	-1.33 (-3.59,0.94)	19 (4,53)	0.00
	Solifenacin ER 10mg q.d	[30]	5	1341	-1.34 (-2.02,-0.69)	19 (10,33)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1182	-1.23 (-2.01,-0.46)	21 (10,39)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	63	-1.15 (-2.8,0.5)	22 (6,51)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	-1.14 (-2.64,0.37)	23 (7,50)	0.00
	Solifenacin ER 5mg q.d	[29]	9	1246	-1.07 (-1.7,-0.49)	24 (14,38)	0.00
	Oxybutynin trandermal 3.9mg/day	[10]	1	48	-1.02 (-2.38,0.34)	25 (8,50)	0.00
	Fesoterodine ER 4mg q.d	[25]	7	3037	-1.05 (-1.55,-0.53)	25 (15,38)	0.00
	Solifenacin ER 5 - 10mg q.d	[31]	3	1431	-0.9 (-1.68,-0.13)	28 (14,44)	0.00
	Propiverine ER 20mg q.d	[41]	4	1000	-0.9 (-1.58,-0.23)	28 (15, 43)	0.00
	Propiverine IR 15mg b.i.d	[43]	1	100	-0.88 (-2.79,0.99)	29 (6,53)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	-0.77 (-2.42,0.85)	31 (8,51)	0.00
	Pregabalin 150mg b.i.d	[62]	1	94	-0.78 (-2.02,0.45)	31 (11,50)	0.00
	Tolterodine ER 4mg q.d	[4]	16	4782	-0.77 (-1.16,-0.39)	31 (22,41)	0.00
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	96	-0.79 (-2.2,0.62)	31 (9,51)	0.00
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.74 (-2.53,1.02)	32 (7,53)	0.00
	Elocalcitol 75mg	[70]	1	84	-0.7 (-2.17,0.77)	33 (9,52)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.69 (-2.11,0.72)	33 (10,52)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	3	488	-0.72 (-1.58,0.13)	33 (15,48)	0.00
	Mirabegron ER 100mg q.d	[52]	3	1136	-0.69 (-1.4,0.01)	33 (18,46)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.65 (-2.48,1.16)	34 (8,53)	0.00
	Mirabegron ER 50mg q.d	[51]	6	2014	-0.67 (-1.24,-0.1)	34 (20,45)	0.00
	Tolterodine IR 2mg b.i.d	[5]	5	746	-0.66 (-1.64,0.29)	34 (16,47)	0.00
	Netupitant 100mg q.d	[111]	1	59	-0.54 (-1.93,0.86)	37 (11,52)	0.00
	Elocalcitol 150mg	[69]	1	87	-0.4 (-1.86,1.07)	39 (12,53)	0.00
	Mirabegron ER 25mg q.d	[50]	2	618	-0.41 (-1.37,0.54)	39 (18,51)	0.00
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.38 (-3.06,2.27)	40 (5,54)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	77	-0.39 (-2.14,1.34)	40 (10,54)	0.00
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	-0.36 (-2.08,1.33)	40 (11,54)	0.00
	Darifenacin ER 15mg q.d.	[40]	1	212	-0.37 (-1.82,1.09)	40 (12,53)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	97	-0.38 (-1.69,0.93)	40 (14,53)	0.00
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-0.19 (-2.82,2.46)	43 (6,54)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	2	44	-0.17 (-1.84,1.5)	43 (14,53)	0.00
	ONO-8539 300mg b.i.d	[61]	1	82	-0.19 (-1.65,1.25)	43 (15,53)	0.00
	Netupitant 50mg q.d	[110]	1	60	-0.17 (-1.55,1.22)	43 (16,54)	0.00
	ZD0947IL 25mg/day	[58]	1	92	-0.09 (-2.04,1.87)	44 (11,54)	0.00
	Netupitant 200mg q.d	[112]	1	55	-0.08 (-1.47,1.31)	45 (17,54)	0.00
	Placebo	[1]	42	9362	NA	46 (40,51)	0.00
	ONO-8539 100mg b.i.d	[60]	1	83	0.06 (-1.39,1.51)	47 (18,54)	0.00
	ONO-8539 30mg b.i.d	[59]	1	87	0.87 (-0.58,2.3)	53 (36,54)	0.00

† median relative to a placebo intervention. Results are based on patients of average baseline risk.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

6.4.1.4 Nocturia

Values of DIC were comparable for both fixed and random effects NMAs (DIC: 1141.63, and DIC: 1143.83, respectively), which is further explained through the minimal between-study standard deviation (0.03, 95%CrI: 0.001,0.07) estimated from the random effects model (Table 6.2). Based on the DIC statistics alone, there is little information to choose between the models. Typically, results from random effects models are preferred for decision making as they incorporate any between-study variability (Whitehead, 2002), which appears to exist in this example, and they tend to produce more conservative estimates of treatment effects than that of fixed effect models (Spiegelhalter et al., 2004). Measures of model fit would suggest that the random effects model is of a slightly better fit to the data than that of the fixed effect equivalent. The total residual deviance for the random effects model was 129.9, relative to 125 unconstrained data points, compared to a total residual deviance of 130.9 for the fixed effect model. Age did not appear to be an important factor in age adjusted analyses as the 95%CrI for the common β coefficient (β: -0.46, 95%CrI: -0.76,0.12) contained the point of no difference. The baseline risk model did not appear to be of any better fit to the data with an increased DIC value of 1148 compared to a DIC of 1141.63 for the unadjusted random effects NMA. The results presented in this chapter were based on the best fitting model for nocturia - the unadjusted random effects model.

Table 6.6 illustrates the estimated posterior median difference relative to placebo for all interventions evaluating change from baseline in nocturia episodes, ranked in order of clinical efficacy. The relative rankings together with the probability that the interventions were the 'best' overall are presented. Once daily oxybutynin 20mg intravesically appeared to be the most effective intervention for reducing nocturia episodes with a probability best of 0.6 and an average rank of 1 (95%CrI: 1,3). The overall median difference in change from baseline was -2.61 (95%CrI: -3.84, -1.39) nocturia episodes relative to placebo. However, oxybutynin 20mg intravesically was only evaluated in one study of 21 participants and these results

should be interpreted with caution. Estriol 1mg intravesically had a probability of being the best intervention overall of 0.3, and ranked in second place (mean rank 2, 95%CrI: 1,3), and estradiol 3mg intravaginally ranked in third place (mean rank 3, 95%CrI: 1,4). However both of these interventions were also evaluated in single studies with very few participants (n=21 and n=15, respectively) and thus should be interpreted with caution.

TABLE 6.6: Estimated posterior median difference (and 95% credible interval) in change from baseline for nocturia episodes relative to placebo obtained from unadjusted random effects network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95%CrI)	Rank	p(Best)
	Oxybutynin 20mg intravesically q.d	[106]	1	21	-2.61 (-3.84,-1.39)	1 (1,3)	0.6
	Estriol 1mg intravesically	[131]	1	21	-2.29 (-3.41,-1.32)	2 (1,3)	0.3
	Estradiol 3mg intravaginally	[128]	1	15	-1.8 (-2.98,-0.74)	3 (1,4)	0.1
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	-0.64 (-1.25,0)	5 (3,32)	0
	Percutaneous tibial nerve stimulation	[83]	3	162	-0.44 (-0.99,0.17)	8 (4,36)	0
	Bladder Training (BT)/Behaviour Therapy	[85]	4	155	-0.41 (-0.69,-0.15)	9 (4,20)	0
	Electrostimulation	[80]	4	161	-0.41 (-0.75,-0.06)	9 (4,29)	0
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-0.41 (-0.85,0.04)	9 (4,33)	0
	Mirabegron IR 100mg b.i.d	[48]	1	58	-0.36 (-0.7,-0.06)	10 (4,30)	0
	Tolterodine ER 4mg q.d + BT	[87]	1	134	-0.35 (-0.64,-0.09)	10 (5,27)	0
	Imidafenacin 0.1mg b.i.d	[36]	1	77	-0.32 (-0.66,0.01)	11 (4,33)	0
	Estradiol 1mg intravaginally	[127]	1	15	-0.32 (-1.77,1.04)	12 (4,40)	0
	OnaBoNT-A 100u trigone sparing	[72]	3	566	-0.26 (-0.41,-0.1)	14 (7,26)	0
	Oxybutynin ER 2.5mg q.d	[20]	2	62	-0.26 (-0.51,0.01)	14 (7,32)	0
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1047	-0.24 (-0.35,-0.13)	15 (9,24)	0
	Mirabegron ER 25mg q.d	[50]	1	179	-0.25 (-0.43,-0.06)	15 (7,30)	0
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.22 (-0.72,0.38)	17 (4,39)	0
	Mirabegron ER 50mg q.d	[51]	5	1803	-0.15 (-0.23,-0.07)	21 (14,30)	0
	Solifenacin ER 10mg q.d	[30]	2	445	-0.15 (-0.28,-0.02)	21 (12,32)	0
	Mirabegron IR 150mg b.i.d	[49]	1	54	-0.14 (-0.48,0.17)	23 (7,37)	0
	Solifenacin ER 5mg q.d	[29]	5	530	-0.14 (-0.28,0)	23 (13,33)	0
	Tolterodine ER 4mg q.d	[4]	10	3363	-0.13 (-0.2,-0.07)	23 (17,30)	0
	Mirabegron ER 100mg q.d	[52]	3	1415	-0.13 (-0.22,-0.02)	24 (16,32)	0
	Fesoterodine ER 4mg q.d	[25]	7	2369	-0.13 (-0.2,-0.06)	24 (16,31)	0
	Fesoterodine ER 8mg q.d	[26]	5	1692	-0.13 (-0.22,-0.05)	24 (15,31)	0
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.13 (-0.45,0.22)	24 (7,38)	0
	Tolterodine IR 2mg b.i.d	[5]	2	143	-0.13 (-0.42,0.1)	24 (9,36)	0
	Oxybutynin trandermal 3.9mg/day	[10]	1	48	-0.12 (-0.4,0.16)	25 (9,37)	0
	Oxbutynin patch 73.5mg	[15]	1	481	-0.1 (-0.22,0.01)	26 (15,34)	0
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.1 (-0.3,0.1)	27 (12,36)	0
	Propiverine ER 20mg q.d	[41]	5	1290	-0.07 (-0.15,0)	29 (21,34)	0
	Resiniferatoxin 50nM	[67]	1	34	-0.06 (-0.62,0.58)	30 (6,40)	0
	Oxybutynin ER 5 - 30mg/day	[22]	1	60	-0.02 (-0.48,0.47)	31 (8,40)	0
	Sham therapy	[3]	1	110	-0.04 (-0.74,0.62)	31 (5,40)	0
	Placebo	[1]	35	6665	NA	33 (29,36)	0
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	0.13 (-0.15,0.4)	36 (22,39)	0
	Terodiline 25mg b.i.d	[28]	3	126	0.14 (-0.04,0.31)	36 (31,39)	0
	Tarafenacin 0.2mg q.d	[90]	1	77	0.14 (-0.16,0.46)	36 (21,40)	0
	Darifenacin ER 7.5mg q.d	[39]	1	29	0.29 (-0.34,0.91)	39 (11,40)	0
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	0.3 (-0.12,0.73)	39 (24,40)	0

† median relative to a placebo intervention

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

6.4.2 Network meta-analysis of safety and tolerability outcomes

For the number of patients experiencing adverse events, the random effects model was of a better fit to the data than that of the fixed model, with a lower DIC of 1628.96 relative to 1660.64, respectively. There were little differences in DIC between fixed and random effects models for discontinuation due to adverse events and discontinuation due to a lack of efficacy. In practice, random effects models are preferred for inference as they capture between-study variability and produce more conservative estimates of treatment effects (Spiegelhalter et al., 2004); and therefore, the results for safety and tolerability outcomes, presented below, are based on random effects models.

6.4.2.1 Number of patients experiencing adverse events

Table C.1 in Appendix C displays the estimated odds ratio and 95% credible intervals for the number of patients experiencing adverse events. Interventions were ranked from the most hazardous (or 'worst' intervention for adverse events) to the least hazardous intervention, with corresponding probability that the interventions were the most hazardous (or 'worst') overall (p(Worst)). Propiverine ER 60mg q.d appeared to be the most hazardous intervention associated with adverse events relative to placebo with an estimated odds ratio of 8.06 (95%CrI: 2.64,25.57). This intervention was comparable to an alternative formulation of propiverine (propiverine IR 45mg t.i.d) with an estimated odds ratio of 7.8 (95% CrI: 2.61,24) relative to placebo. Both interventions had an average rank of 3 (95%CrI: 1,19). Propiverine ER 60mg q.d had a slightly higher probability of being the worst of 0.20 compared to a probability of being the worst of 0.17 for propiverine IR 45mg t.i.d.

6.4.2.2 Discontinuation due to adverse events

Table C.3 in Appendix C displays the estimated odds ratios, rankings, and probabilities that each intervention was the least tolerable in terms of adverse events. No single intervention dominated the analyses for discontinuation due to adverse events, which is shown through the very small probabilities that each intervention was the least tolerable overall (p(Worst)). Imidafenacin 0.25mg b.i.d had the highest estimated mean rank of 6 (95%CrI: 1,20) and corresponding odds ratio relative to placebo of 15.05 (95%CrI: 4.69,61.35). As shown through the wide 95% credible intervals there appeared to be a large degreee of uncertainty in both the estimated odds ratios and estimated subsequent rankings. Trospium 30mg/day in combination with solifenacin 10mg/day on a continuous cycle had the highest probability of discontinuing treatment due to adverse events (p(Worst)=0.20). This was closely followed by oxybutynin 20mg intravesically q.d (p(Worst)=0.19).

6.4.2.3 Discontinuation due to a lack of efficacy

There is a considerable amount of between-study heterogeneity for discontinuation due to a lack of efficacy. The estimated between-study standard deviation (Table 6.2) from the random effects model was 0.4 (95%CrI: 0.06, 0.75). This result suggests that the median ratio of the maximum to minimum odds ratio for a random pair of studies was 1.55. In other words, one study may show no effect whilst the other may have an odds ratio of 1.55 (Spiegelhalter et al., 2004). Table C.3 in Appendix C displays the estimated odds ratios and 95% credible intervals for discontinuation due to a lack of efficacy. Interventions were ranked from least tolerable to most tolerable, with corresponding probabilities. Similarly to discontinuation due to adverse events, no single intervention dominated the analyses as the least tolerable intervention. ONO-8539 30mg b.i.d and serlopitant 4mg q.d - a neurokinin 1 receptor antagonist - both had an average rank of 8 (95%CrI: (1, 42) and 95%CrI: (1, 37), respectively). ONO-8539 30mg b.i.d had a slightly increased probability of being the worst (p(Worst)=0.08) compared to serlopitant 4mg q.d (p(Worst)=0.06), but generally, the probabilities across all interventions

were very small. Cizolirtine citrate 400mg b.i.d had the highest probability of being the worst of 0.13 for discontinuation due to a lack of efficacy.

6.4.3 Treatment profiles

For syndromic conditions such as OAB, interest will often lie in the performance of interventions across the entire symptom profile. Figure 6.2 illustrates the efficacy, safety and tolerability profiles of each treatment in ranked order across each of the outcome measures. Treatments are ranked in order of efficacy for each outcome from left to right i.e. treatments are first ranked by their effective management of incontinence, then by voiding, urgency, and so on so forth, across all outcome measures. Dark green indicates better performing interventions (i.e. the most effective, tolerable or safest intervention) and red indicates the least effective, tolerable and most hazardous interventions. Where blank, data were not available i.e. the interventions were not analysed or were disconnected from the networks of evidence. Sacral nerve stimulation appeared to be the most effective intervention for reducing both urinary incontinence and voiding episodes, but data were not available for all other efficacy and safety measures (Figure 6.2), and these results were based on studies with few participants. Similarly, electrostimulation with PFMT and BT appeared to be highly effective for reducing urinary incontinence and voiding episodes, and oxybutynin IR 2.5mg b.i.d in combination with salivary pastilles were effective for reducing incontinence episodes, but efficacy data for remaining outcomes were unavailable for both combination therapies. Across all evaluable outcomes, onaBoNT-A had the best treatment profile, ranking amongst the top interventions for efficacy, safety and tolerability measures. With regard to efficacy, onaBoNT-A ranked among the top 5 interventions for reducing the three cardinal symptoms of OAB: urinary incontinence, increased voiding and urgency.

Electrostimulation plus vaginal oestrogen cream 1.25mg daily as a combination therapy was most effective at reducing urgency episodes, although, in combination and individually, electrostimulation and vaginal oestrogen cream appeared to be less effective for treating urinary incontinence (Figure 6.2). In terms of safety, electrostimulation as an individual treatment ranked amongst the top interventions for minimising the number of patients experiencing adverse events, but in combination, little is known regarding safety or tolerability. However, it is worth noting that results regarding urgency should be interpreted with caution for electrostimulation interventions due to the inclusion of 2 potentially biased and small scale studies (study id: A157 and A158).

Of the drug therapies, mirabegron appeared to have equal efficacy to all antimuscarinic drugs including solifenacin, oxybutynin, imidafenacin, fesoterodine, darifenacin, and tolterodine, but with an improved tolerability profile (Figure 6.2). Mirabgeron also had a better safety profile but equal efficacy to other β-3-adrenoceptor agonists such as solabegron. It is apparent that antimuscarinics and β-3 adrenoceptor agonists in combination with alternative strategies such as neuromodulation appeared amongst the top interventions for efficacy outcomes.

FIGURE 6.2: Heatmap of treatment profiles across all outcomes

*Darifenacin 7.5-15mg q.d + BT [88], Tolterodine (dose not specified) + BT [94], Darifenacin ER 7.5-15mg q.d [104], Tolterodine (dose not specified) [105], Lidocaine gel 6ml [129], Emepronium bromide IR 200mg [130], Sacral nerve stimulation + Tolterodine ER 2mg q.d [136], Tolterodine ER 2mg q.d [137] were disconnected from the network

6.4.4 Model assessment

6.4.4.1 Assessing inconsistency between direct and indirect information

Inconsistencies between direct and indirect information for efficacy outcomes are given in Tables C.4 - C.7 of Appendix C. Inconsistencies in networks of evidence were first introduced in Section 3.4.5.3. Conflicts between direct and indirect evidence were also assessed for safety and tolerability outcomes, but the results of which are not presented in this thesis. Overall, there were very few inconsistencies between direct and indirect evidence. For incontinence episodes, treatment comparisons between tolterodine ER 4mg q.d versus solifenacin ER 5-10mg q.d, and fesoterodine ER 4mg q.d versus fesoterodine ER 8mg q.d appeared to have conflicting direct and indirect evidence. For voiding episodes, placebo versus electrostimulation, and control versus PFMT appeared to have inconsistent estimates of treatment effects from direct and indirect sources of information. Urgency did not appear to have any inconsistencies between closed loops of evidence. Nocturia appeared to have inconsistencies between direct and indirect information for treatment comparisons including placebo versus electrostimulation, placebo versus BT, and tolterodine IR 2mg b.i.d versus solifenacin ER 5mg q.d. It is worth noting that studies A157 (Wang et al., 2006) and A158 (Wang et al., 2009) appeared to have inconsistent estimates of treatment effects compared to all other trials comparing the same interventions for voiding and nocturia. These studies were previously identified in Section 5.7 as potentially biased studies in terms of small-study effects for urgency episodes. Discrepancies between direct and indirect evidence can only be assessed on closed loops of evidence with more than one trial - by definition, inconsistencies within single trials can not exist. For this reason, treatment comparisons assessed by Wang et al. (2006) and Wang et al. (2009) (placebo versus electrostimulation) were not investigated for urgency episodes as no other trials in the network of evidence compared these interventions of interest. However, the impact of studies A157 and A158 for urgency episodes were further explored in Section 6.4.5 below. For all other outcomes, individually removing

studies that potentially contributed to inconsistent estimates of direct and indirect information had very little impact on the overall treatment effect estimates, rankings, and therefore clinical conclusions.

6.4.4.2 Convergence diagnostics

Convergence diagnostic plots for the basic parameters, d_{1j}, the pooled effect estimate of intervention j relative to placebo, and τ, the between-study standard deviation, for urinary incontinence episodes are given in Appendix C. Similar plots were obtained and evaluated for all other models described in this chapter, however, for illustrative purposes, a small sample were presented. A description of all convergence diagnostic plots are given in Section 3.2.2. All models appeared to achieve a reasonable degree of convergence as illustrated by Figure C.1 - C.4. Figure C.1 illustrates the Brooks-Gelman-Rubin plots for parameters of interest. Both the between-, and within-chain variability, appeared to reach stability, and the ratio of these, appeared to converge to 1. Autocorrelation plots given in Figure C.2 suggested adequate mixing of the chains with all plots appearing to show reduced autocorrelation with increased lag. History and trace plots are presented in Figure C.3, and took the appearance of random noise about an overall average, thus suggesting that there was no evidence to suggest non-convergence of the MCMC chains. Furthermore, there did not appear to be any obvious differences between multiple chains using very different starting values. Density plots are given in Figure C.4. All parameters appeared to have a characteristically bell-shaped appearance reflecting a normal distribution. Overall, the diagnostic plots appeared to suggest that both d_{1j} and τ were estimated from samples which did not appear to have inadequate convergence or mixing of the chains.

6.4.4.3 Model fit

Random effects models had a better fit to the data for most outcomes as can be seen by Table 6.2. Table 6.2 displayed estimates of the between-study standard

deviations for all models. Heterogeneity between studies was generally low, with the between-study standard deviations of 0.16 for urinary incontinence, 0.13 for voiding, 0.6 for urgency, and 0.03 for nocturia. This means that at most, the difference in the range of OAB episodes that the studies varied by, in terms of change from baseline, was approximately 2.4 episodes. For binary outcomes on the log-odds scale, the between-study standard deviations were 0.23 for number of patients with adverse events, 0.07 for discontinuation due to adverse events, and 0.40 for discontinuation due to a lack of efficacy. On an odds-ratio scale this means that at most, the difference in the ratio of odds ratios was 1.55. Overall, models were of a good fit to the data where many of the posterior mean residual deviances, given in Table 6.2, closely reflected the number of unconstrained data points.

6.4.5 Sensitivity analysis

Sensitivity analyses assessing the choice of prior distributions on variance parameters for primary outcomes are given in Tables C.8 - C.11 of Appendix C. Sensitivity analyses assessing the choice of variance parameters were also assessed for safety and tolerability outcomes, and all other prior distributions, but these results are not presented in this thesis. Overall the choice of prior distribution for variance parameters had very little impact on the treatment effect estimates. Table C.12 in Appendix C displays the treatment effect estimates for urgency episodes after the exclusion of potentially biased studies identified in Section 5.7. With the exclusion of studies A157 (Wang et al., 2006) and A158 (Wang et al., 2009), treatment effects for oxybutynin IR 2.5mg t.i.d were removed from the network of evidence. Therefore 53 interventions were evaluated in total. Contrary to the evidence presented in Section 6.4.1.3, results from the sensitivity analysis indicated that cizolirtine citrate 400mg b.i.d and onaBoNT-A 100u injected in to the bladder base and trigone were the most effective interventions, each with a 25% probability of being the best

intervention overall. Uncertainty in treatment effect estimates for electrostimulation interventions increased, which is shown through the wider credible intervals. Electrostimulation in combination with vaginal oestrogen cream 1.25mg/day had an estimated reduction of -0.92 (95%CrI: -4.13,2.44) urgency episodes relative to placebo, compared to an estimated reduction of -6.94 (95%CrI: -8.65, -5.23) from the analysis including all trials. This increase in uncertainty resulted in just a 7% probability that electrostimulation plus vaginal oestrogen cream 1.25mg/day was the best intervention overall. This result is vastly different to the results including studies A157 and A158 and thus, caution must be taken with the interpretation of results for urgency outcomes including these estimates. A sensitivity analysis assessing the impact of outlying studies with large treatment effects, namely trials evaluating SNS, are given in Table C.13 of Appendix C.14. Removing SNS from the network of treatment comparisons had very little impact on the remaining treatment effect estimates of alternative therapies.

6.5 Discussion

This thesis reports the largest and most comprehensive comparison of interventions for OAB to date. To my knowledge, this is the first study to compare all conservative and minimally invasive treatments for the management of OAB in a single coherent analysis of the most salient outcomes. Clinical-effectiveness, safety, and tolerability of interventions were considered. In these analyses, SNS appeared to be the most effective intervention for treating urinary incontinence and voiding episodes, with effect sizes much larger than the next ranked interventions, but data on urgency and nocturia, as well as safety and tolerability outcomes, were missing. It is also worth noting that efficacy data for incontinence and voiding episodes were both based on single studies evaluating the efficacy of SNS in a small number of participants. Electrostimulation and vaginal oestrogen, as a combination therapy, appeared to be the most effective for treating urgency episodes, and each intervention alone ranked second and third for this outcome. However, these results should be interpreted with caution due to the inclusion of potentially

biased and small-scale studies. Further, it is worth noting that electrostimulation in combination with BT and PFMT ranked highly for both urinary incontinence and voiding frequency. Across three of the cardinal symptoms of OAB (urinary incontinence, voiding frequency and urgency), different doses and injection regimes of SNS, electrostimulation and onaBoNT-A featured in the top 5 most effective treatments.

Antimuscarinic drugs barely feature in the most effective interventions for efficacy and safety analyses, and when they do, they appear in the lower rankings in combination with neuromodulation. The heatmap displayed in Figure 6.2 illustrates all of the analysed interventions and demonstrates the same pattern, with very little difference in efficacy or adverse event profiles between any of the antimuscarinic drugs. The most important observation, in terms of oral drug therapies, is that mirabegron, a β_3-agonist, demonstrates equivalent efficacy to antimuscarinic drugs, but with a better adverse event profile.

There may be additional benefit in administering SNS in combination with antimuscarinic drugs for voiding episodes, as reported in a study by Tang et al. (2014) (study id: A145) who evaluated SNS in combination with tolterodine ER 2mg q.d. However, as previously discussed in Section 5.9.3, treatment comparisons including SNS in combination with tolterodine ER 2mg q.d were disconnected from the networks of evidence, and thus could not be compared in the NMAs. Tang et al. (2014) evaluated tolterodine ER 2mg q.d alone, and in combination with SNS, in 240 participants equally randomised to two treatment groups. In this study, patients experienced an average reduction of -5 and -12.5 voiding episodes (measures of uncertainty not reported), respectively. Empirical evidence (Tang et al., 2014; Abrams et al., 2015; Kosilov et al., 2015; Burgio et al., 2008, 2010; Chancellor et al., 2008; Fitzgerald et al., 2008; Herschorn et al., 2004; Kafri et al., 2013; Kaya et al., 2011; Millard, 2004; Ozdedeli et al., 2010; Zimmern et al., 2010) suggests that there may be a benefit of administering combination therapies of

multi-modalities such as SNS and tolterodine for the management of OAB. Future RCTs comparing SNS in combination with antimuscarinics, compared to SNS alone (or standard of care), are required in order to comprehensively assess additional benefit of combination therapies.

Previous studies have performed a series of pairwise meta-analyses comparing treatments for OAB on a head-to-head basis (Novara et al., 2008a; Luo et al., 2012; Chapple et al., 2008a; Cui et al., 2013, 2014, 2015; Nitti et al., 2013; Wu et al., 2014; Liu et al., 2014; Reynolds et al., 2015; Davis et al., 2015; Huang et al., 2015; Sun et al., 2015; Cao et al., 2016; Sweeney et al., 2016). A study by Buser et al. (2012) evaluated multiple treatments of antimuscarinic drugs in a single, comprehensive analysis and found that solifenacin, fesoterodine, oxybutynin, and trospium chloride were most effective for treating urinary incontinence, voiding, urgency and nocturia, respectively. For incontinence and urgency, the treatments ranked in first place mirror that of the results presented here, but due to minor differences in treatment effects between all antimuscarinics, the rankings described in this chapter differ slightly to Buser et al. (2012) for voiding and nocturia. Freemantle et al. (2016) compared onaBoNT-A with mirabegron 25mg q.d and 50mg q.d, and found comparable estimates for the difference in median reduction of incontinence, urgency, voiding and nocturia episodes, for onaBoNT-A relative to both formulations of mirabegron. Overall, the findings of this chapter are consistent with those in the literature and estimate similar effect sizes for comparable treatments. The National Institute for Health and Care Excellence (NICE) in the UK currently recommends bladder training, and oxybutynin IR as first line therapy for urge urinary incontinence, primarily on the basis of low cost (National Institute for Health and Care Excellence, 2013b). Second-line therapies include other antimuscarinic drugs (trospium, propiverine, or tolterodine) and then more invasive procedures such as onaBoNT-A injection and SNS. Mirabegron is currently recommended as an alternative second-line medication. The American Urological Association (AUA) recommends a broader set of interventions for the management of OAB (Gormley et al., 2012). These include behavioural therapies

as first-line, antimuscarinics or oral β-3-adrenoceptor agonists as second-line, and onaBoNT-A (100U) as third-line therapies. Taking in to account the increasing degree of invasive procedures, the different classes of treatments recommended by NICE and AUA broadly agree with the findings of this chapter.

This analysis goes beyond that of pairwise comparisons and allows comparison between all conservative and minimally invasive treatments. In doing so, mirabegron appeared to be equally effective to oxybutynin IR (and better than other antimuscarinics) but with an improved safety and tolerability profile, which agrees with recently published data (Maman et al., 2014). Despite concerns about cardiovascular risk (Caremel et al., 2014), mirabegron appears safe (Rosa et al., 2016; Wagg et al., 2015), and the findings of this chapter support a growing body of expert opinion (Duckett and Balachandran, 2016; Wagg et al., 2016) that mirabegron could be offered as a first-line drug treatment instead of antimuscarinic drugs after more conservative measures have failed.

As highlighted above, the data described here can add value to clinical decisions about choice of intervention, and the order in which different interventions can be delivered. Whilst the common perception is that less severe symptoms require less invasive interventions, for some patients with moderate to severe symptoms, the greater relative efficacy afforded by more invasive therapy, as indicated here, could be offset by the level of invasiveness, regardless of the traditionally perceived position in the treatment pathway. It is anticipated that clinicians together with patients will use this data to arrive at better decisions for treatment hierarchies. Of course, clinical decision making takes into account patient factors and choices, cost, and expected gain in quality of life. Whilst these outcomes are not evaluated here, it is anticipated that the analyses described in this chapter will be used to inform the decision making process, and treatment effect estimates could be used to populate appropriate cost-effectiveness models.

The traditionally perceived treatment pathway suggests that patients with more severe symptoms are more likely to be assessed in clinical trials of more invasive interventions, and therefore baseline severity of the patient population may act as a potential treatment effect modifier in NMAs. This potential covariate effect was considered, and assessed through the use of baseline risk-adjusted models. However, it is well documented that network meta-regression models may lack power when there are few studies (Jansen et al., 2008), and are susceptible to ecological biases when dealing with aggregate data (Lambert et al., 2002; Berlin et al., 2002). It is clear for urinary incontinence episodes and voiding frequency that participants receiving SNS had an increased rate of incontinence and voiding episodes at baseline compared to the rest of the cohort, and thus these results should be interpreted with caution. However, excluding SNS from the NMA in a sensitivity analysis given in Appendix C.14 had very little impact on the results of alternative therapies for urinary incontinence. This is largely because SNS does not form a connected loop in the network (Figures 5.5 and 5.6) and therefore has very little impact on the remaining treatment effect estimates. (Mills et al., 2013a,b).

A fundamental assumption underpinning all network meta-analyses is the assumption of additivity (see Section 3.4.5). Additivity of treatment effects was assessed, where possible, using node-splitting methods. However, this assumption can only be assessed on closed loops of treatment comparisons, and thus, given the geometry of the networks of evidence, this assumption could not be tested for all sets of treatment comparisons. This is of most concern if there exists outlying studies with particularly large treatment effects (Mills et al., 2013b). Therefore, sensitivity analyses assessing the impact of outlying studies are crucial. In this example, removing outlying studies evaluating SNS in a sensitivity analysis had very little impact on the remaining treatment effect estimates. The analyses presented in this chapter make a further assumption that treatment effects are exchangeable. Assuming exchangeability across very different treatments such as SNS and behavioural therapies may not be reasonable. Chapter 7 relaxes this assumption and

assumes partial exchangeability of treatment effects.

One of the limitations of this analysis is that it does not account for longer-term outcomes. For example, it is known that many patients do not continue oral medication into the long term (Kelleher et al., 1997b). As illustrated in Chapter 4, repeated treatment of onaBoNT-A has a prolonged duration of effect, but also some potentially important adverse events (Mangera et al., 2014). Whilst SNS can provide long lasting relief (Brazzelli et al., 2006), it is expensive and prone to failure. However, onaBoNT-A appears to be equally effective and well tolerated in treating symptoms of OAB and may be considered a safer alternative for the treatment of urinary incontinence, although there is a risk of medium term urinary retention. For reasons that are unclear, electrostimulation in its different forms has not achieved widespread use despite this analysis demonstrating it to be (alone or in combination with PFMT and BT) a promising intervention in terms of safety and tolerability compared to many other interventions. A more complete interpretation of these findings requires further information in terms of clinical efficacy in a wider patient population, compliance to treatment, and long-term cost-effectiveness.

This chapter investigated fixed and random effects NMAs, as well as age-adjusted, and baseline risk-adjusted models for efficacy outcomes. Models with the best fit to the data were chosen for presentation; however, model choice had little impact on the overall treatment effect estimates and relative treatment rankings. Sensitivity analyses investigating the choice of prior distribution had little impact on the treatment effect estimates, and there were few inconsistencies between direct and indirect estimates for each of the outcomes of interest. Removing potentially inconsistent studies from the analyses had no impact on the overall rankings. For urgency episodes, sensitivity analyses investigating the inclusion of potentially biased studies (study id: A157 and A158) showed that both RCTs had an important impact on the uncertainty in treatment effect estimates, and consequently, the overall clinical decision. Thus, inferences regarding treatment effect estimates

for urgency episodes should be interpreted with caution, and further information regarding the efficacy of oxybutynin IR 2.5mg t.i.d and electrostimulation in the form of a large RCT is required to support the existing evidence.

However, one concerning factor highlighted in this chapter was that many of the published papers omitted key symptoms. Eighty-four studies evaluating OAB failed to report at least one relevant outcome - 117 (67%) papers reported urinary incontinence, 124 (71%) included voiding frequency and only 62 (36%) papers included urgency as an outcome, despite this being regarded as the "cardinal symptom" (Cardozo et al., 2009). This finding demonstrates the need for developing core outcome sets for OAB (Gargon et al., 2015). The absence of data for one or more outcomes for many of the interventions had a substantial effect on the analyses as many efficacy profiles were incomplete, which is further illustrated in Figure 6.2. To overcome this, and to limit the risk of bias from selective reporting (Kirkham et al., 2012), a multivariate approach that simultaneously models the outcomes of interest is needed (Achana et al., 2014; Efthimiou et al., 2014); and so, for studies that fail to report any given outcome, a multivariate NMA can be used to predict and impute a missing value using the correlation between outcomes. Chapter 8 further explores the use and development of multivariate NMAs in order to address this issue in the context of OAB.

6.6 Chapter summary

This comprehensive systematic review and NMA has enabled all treatment modalities, of varying levels of invasiveness, for DO and OAB to be compared with one another in terms of efficacy, safety and tolerability. It enables patients to make informed personal decisions regarding the management of their condition, whilst also providing appropriate evidence for future net-benefit decisions or cost-effectiveness analyses in order to inform health policy decision making for this debilitating condition. Sacral nerve stimulation appeared to dominate for incontinence and voiding

outcomes, but the invasiveness and cost may make it less attractive as a primary intervention. OnaBoNT-A appeared to have important advantages over other treatments, both in terms of efficacy and safety. The new beta-3-adrenoreceptor agonists showed comparable efficacy to established antimuscarinics but with much better tolerability.

There is, however, considerable uncertainty in the treatment effect estimates for many of the interventions assessing efficacy, safety, and tolerability outcomes. This is further shown through the 95% credible intervals of relative treatment rankings, in which the majority of credible intervals span a large range of plausible values. Chapter 7 aims to develop and extend the methodology described in this chapter, in order to increase precision in treatment effect estimates for large networks of treatment comparisons. This approach incorporates a hierarchical structure by making use of similarities between the same interventions with different formulations and/or methods of administration.

Chapter 7

Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder

7.1 Chapter overview

From a decision-makers perspective, it is imperative that decisions regarding health-care policy are formulated using all of the relevant evidence, and with appropriate uncertainty - the foundations of which are largely rooted with the size of the evidence base, and the level of uncertainty in the treatment effect estimates obtained from empirical evidence. As illustrated in Chapter 6, in the case of OAB, synthesising data from a large network of treatment comparisons using a general NMA framework produced considerable uncertainty in the treatment effect estimates and relative treatment rankings, for both efficacy and safety outcomes. The method described in this chapter aims to explore whether an increase in precision in the treatment effect estimates is possible by accounting for similarities between the same interventions, but with different formulations, and treatment regimes,

using a natural hierarchical structure. Classifying interventions in this way makes use of a common 'class effect' which will allow strength to be 'borrowed' across interventions that share a common classification. In situations where the networks of treatment comparisons are particularly extensive and the evidence base is somewhat sparse, use of hierarchical models could be of particular value as they have the potential to increase precision in treatment effect estimates, and permit more certain inferences regarding interventions effectiveness.

In this chapter, a hierarchical approach previously proposed by Dakin et al. (2011) and Warren et al. (2014) is extended to incorporate dose-response constraints. The methodological developments described here were published in *Value in Health* (Owen et al., 2015), with the full manuscript given in Appendix F.3. The chapter begins with an outline of the motivation for developing and fitting hierarchical NMAs in the context of OAB, before illustrating the implications of adopting a class-based approach. Hierarchical NMAs, with and without dose response constraints, were then discussed for both continuous and binary outcomes. For illustration purposes, one continuous outcome - mean change from baseline in urinary incontinence episodes, and one binary outcome - number of patients experiencing adverse events, were presented here, though the methodologies described in this chapter were applied to all outcomes of interest, and are revisited in Chapter 8.

7.2 Introduction

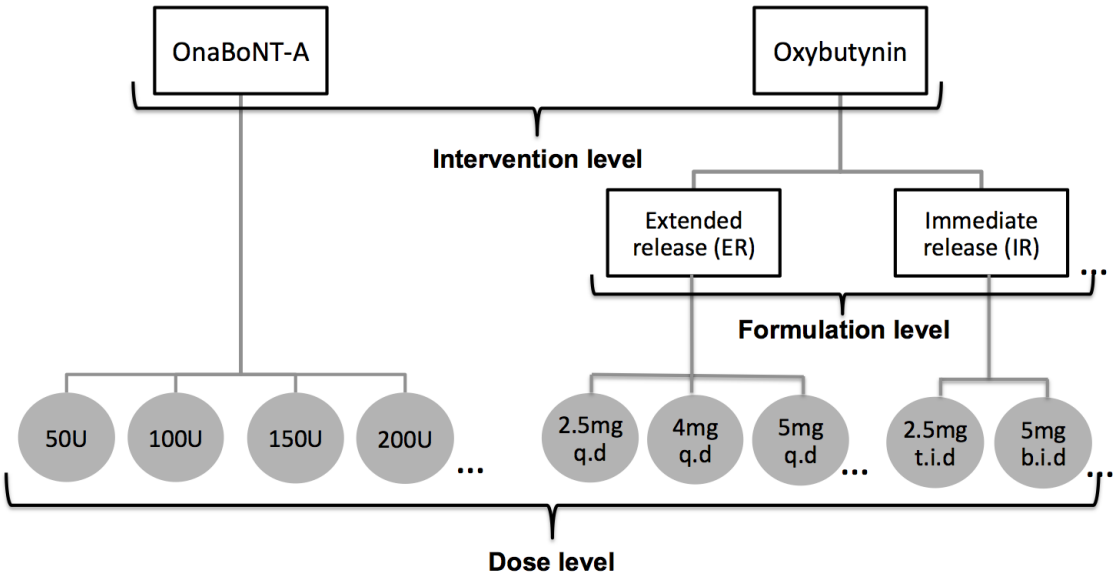
Network meta-analyses are widely used in an evidence synthesis setting due to the attractive nature of utilizing all relevant information from both direct and indirect evidence. Nevertheless, in situations where there are a large number of interventions of interest and relatively few trials informing each treatment comparison, there is a potential issue with the sparsity of data in the treatment networks, which can lead to substantial parameter uncertainty. Collapsing the intervention

arms into their respective treatment classes, also known as "lumping" interventions, increases the evidence base and precision in the effect estimates, but with such a class-based approach, the direct interpretation of individual intervention effects (especially those of dose or formulation effects) are lost, which makes decision making difficult. To overcome this issue, a three-level hierarchical NMA can be applied. In the current literature, use of the term "hierarchical NMA" is intermittently used to describe what is commonly known as a "random effects NMA" described in detail in Section 3.4.2 of Chapter 3. Briefly, a random effects model has variance components at two levels in the model, one at the within-study level and one at the between-study (within intervention) level. In this chapter, a third level was incorporated in to the model to account for an additional variance component between interventions that share a common classification. This approach incorporates the exchangeability between interventions of the same class to estimate a treatment effect for each of the individual interventions. In doing so, this approach allows information to be borrowed within the classes of interventions, strengthening inferences, and potentially reducing the uncertainty around the individual intervention effects; and thus, increases the ability to rank interventions and inform healthcare decision-making. To further increase the precision in the effect estimates, constraints can be applied to increasing doses of an intervention, such that higher doses have a greater or equal treatment effect to that of lower doses, or vice versa, for safety and tolerability outcomes. An example of the treatment hierarchies for the management of OAB are illustrated in Figure 7.1.

As previously described in Chapter 2, given the availability of numerous interventions for the management of OAB, and emerging treatments such as β-3-adrenoceptor agonists, onaBoNTA, and SNS, there is an increasing need to identify the most beneficial intervention from a diverse range of treatment modalities. A comprehensive systematic literature review described in Chapter 5 identified a large number of RCTs evaluating a wide variety of interventions, and treatment regimes, but with relatively few direct treatment comparisons between active interventions. Synthesising these data using a general NMA framework as described

in Chapter 6 resulted in a considerable level of uncertainty associated with many of the individual treatment effect estimates and relative treatment rankings, which can hinder decision making. However, difficultly in drawing overall conclusions is also partly determined by the extent of heterogeneity in a meta-analysis (Higgins et al., 2009). It is known that in situations where there are a limited number of trials in a meta-analysis, estimating heterogeneity between studies may be problematic (Hardy and Thompson, 1998). One approach to overcome this issue, and increase precision in the treatment effects, involves incorporating external information of between-study variability from similar studies relevant to the treatment comparison of interest. However, such external information may be limited. The aim of this chapter was to develop and apply hierarchical NMAs to evaluate interventions for the management of OAB by borrowing information between interventions of the same class and applying ordering constraints on increasing doses. This approach has the potential to increase precision in effect estimates but maintain the interpretability of results at the individual intervention level. Adding an additional level, i.e. variance component, in the model allows for sources of heterogeneity to be more easily identified as it would be possible to quantify heterogeneity between treatments as well as between classes of interventions.

FIGURE 7.1: Example of intervention hierarchies for the management of OAB

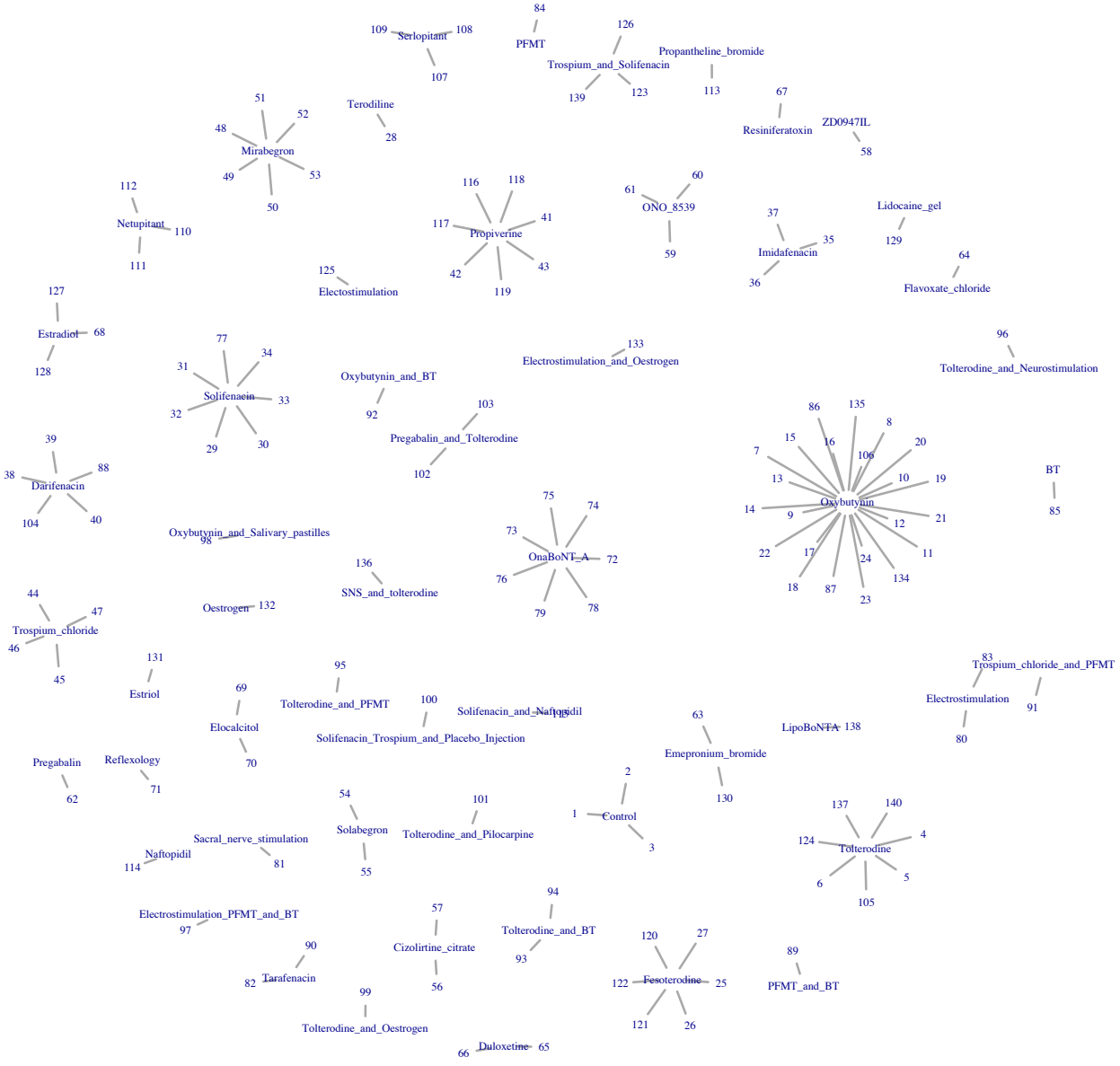


7.3 Methods

For analyses associated with the classes of interventions, expert clinical opinion was sought from Professor Douglas Tincello to group interventions in to clinically meaningful hierarchies. Figure 7.2 demonstrates the classification of each of the individual interventions, where the central node represents the classes of treatments and the linked arms represent each of the individual interventions within those classes. In this example, antimuscarinics and β_3-adrenoceptor agonists were grouped at the intervention levels, and therefore, all formulations and doses of the same intervention were considered to be of the same class. The onaBoNT-A treatment class contained all onaBoNT-A interventions regardless of the site of administration. The electrostimulation group contained electrical magnetic therapy and all types of electrostimulation therapies as described in Section 2.5.2.3. Control, sham therapies and placebo interventions were grouped in a single class of control interventions. Combination therapies were grouped only if the combined therapies evaluated the same set of classes of interventions.

Chapter 5 illustrated the networks of direct treatment comparisons evaluating the mean reduction in incontinence episodes from baseline (Figure 5.5), and the number of patients experiencing adverse events (Figure 5.9). Figures 7.3 and 7.4 illustrate the networks of direct treatment comparisons for classes of interventions evaluating mean reduction in incontinence episodes from baseline, and the number of patients reporting adverse events, respectively. The nodes represent the classes of interventions and the interconnecting lines demonstrate a direct treatment comparison between classes of interventions. The density of the interconnecting lines reflect the number of trials that directly compare adjoining classes of interventions. As previously described in Section 5.9, interventions which were disconnected from the network were excluded from the analysis.

FIGURE 7.2: Classification of interventions



The central nodes represent classes of interventions. The linked arms represent each individual intervention within those classes, with corresponding treatment codes given in Table 5.5 of Chapter 5.

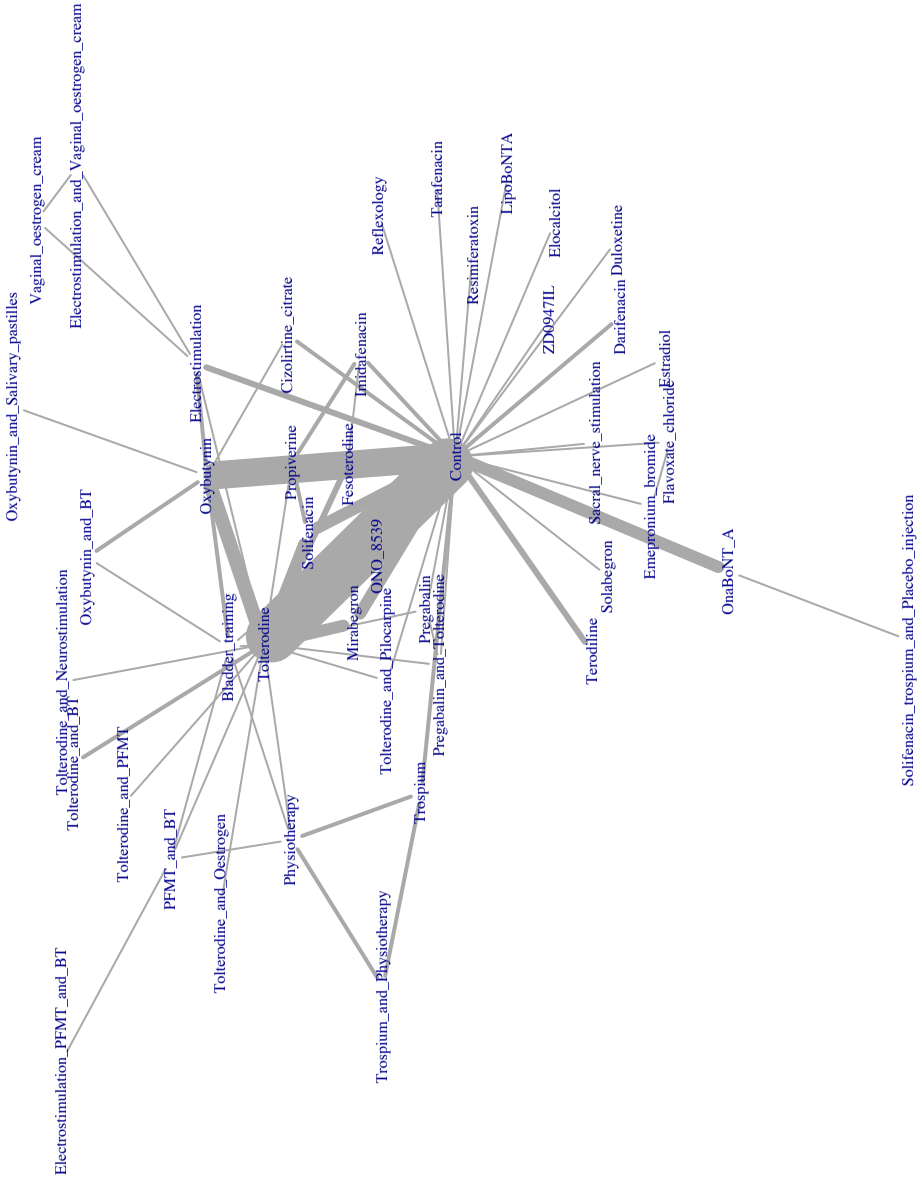


FIGURE 7.3: Network of evidence for classes of interventions (Incontinence)

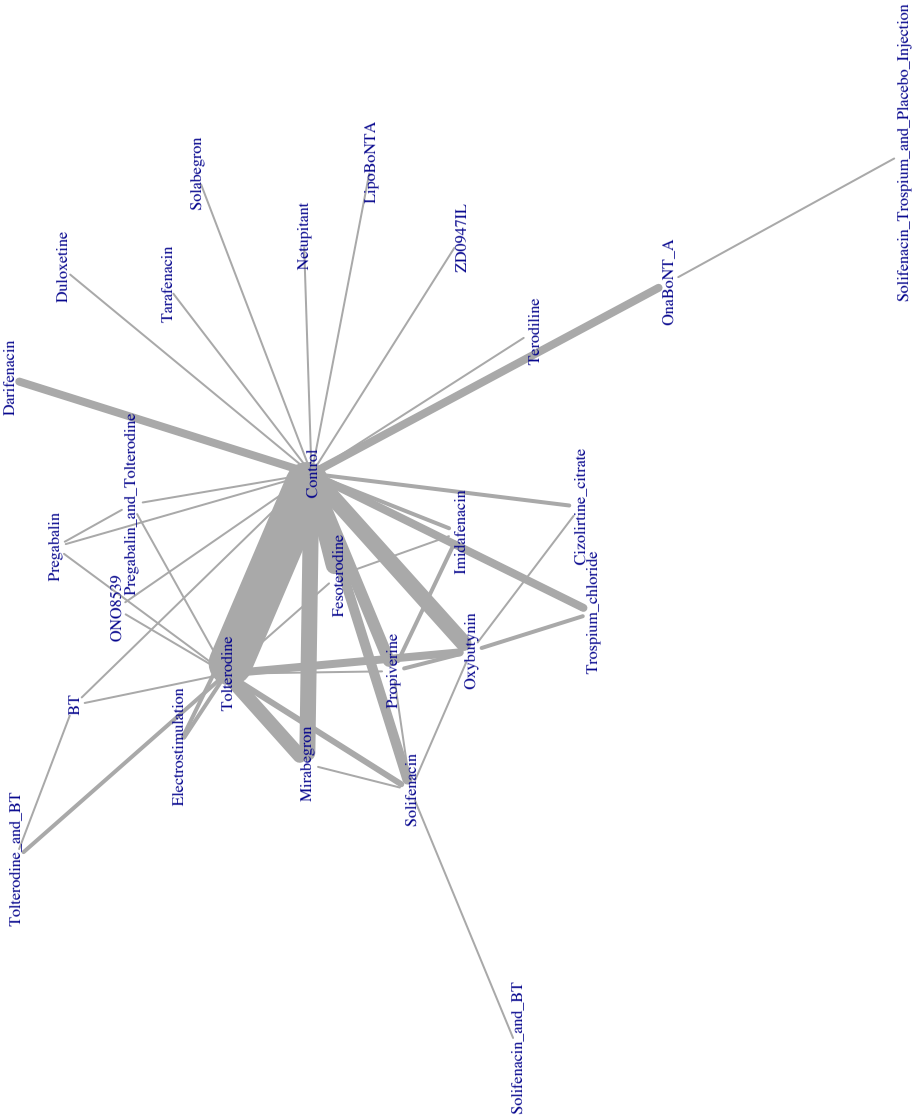


FIGURE 7.4: Network of evidence for classes of interventions (Adverse events)

7.3.1 Class-based network meta-analysis

For class-based NMAs, studies reporting multiple interventions that belong to the same class of interventions were combined to obtain a pooled average and standard deviation for the overall class effect, using Equation (5.1) described in Section 5.3.1. Studies that solely compared interventions within the same class of interventions were excluded from the analyses as there was no longer a comparator arm with which to model the class effects whilst maintaining randomisation.

Random effects NMAs were used to estimate the clinical effectiveness, and safety of classes of interventions, using the general NMA frameworks presented in Equation (6.1) and (6.2) of Chapter 6. In this instance interest lies in evaluating classes of interventions, m, rather than individual interventions j, and so these frameworks have a slightly different interpretation to that of Chapter 6. As an example, for continuous outcomes, such as mean change from baseline in incontinence episodes, y_{im} are estimates of true class effects, θ_{im}, with observed standard error, se_{im}, for a class of interventions, $m = 1, 2, ..., n_m$, in study i, and assumed to follow a normal distribution. μ_i represents the mean change from baseline for a study-specific reference class, t_{ib}, in arm 1 of the i^{th} study. The mean difference in class effects between treatment classes in the k^{th} arm, t_{ik}, relative to t_{ib} was given by $\delta_{i,bk}$, which for 2-arm trials was drawn from a normal distribution with mean equal to the relative difference in class effect of the treatment class in arm k, compared to the class effect in arm 1, of study i, $d_{t_{ib}t_{ik}}$, with between-study variance, τ^2. The correlation between-arms of the same study for multi-arm trials was accounted for using conditional distributions (see Section 3.4.5). Here, the relative effect of the study-specific reference treatment in arm 1 (the control arm) relative to itself, $\delta_{i,b1}$, is set to 0 and as such the set of conditional univariate distributions begin with the relative effect of the intervention in arm 2 relative to the control arm, $\delta_{i,b2}$.

Thus the class-based NMA model is given by:

$$\begin{aligned} y_{im} &\sim \text{Normal}(\theta_{im}, se_{im}^2) \\ \theta_{im} &= \mu_i + \delta_{i,bk} I_{\{m=k\}} \\ \text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{7.1}$$

The study-specific relative treatment effect, $\delta_{i,bk}$, for arm $k = 2$ is expressed as:

$$\delta_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{ik}}, \tau^2)$$

For multi-arm trials of $k > 2$, the study-specific relative treatment effect, $\delta_{i,bk}$, for the k^{th} conditional distribution, $k = 3 \ldots na_i$, is defined by:

$$\delta_{i,bk} | \begin{pmatrix} \delta_{i,b2} \\ \vdots \\ \delta_{i,b(k-1)} \end{pmatrix} \sim \text{N}\left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\delta_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right)$$

For binary outcomes, the number of events, r_{im}, in each study i of class m was assumed to follow a Binomial distribution with probability of an event, p_{im}, and denominator n_{im} . A logit link was used to relate the probability of event to the treatment effect of the class:

$$\begin{aligned} r_{im} &\sim \text{Binomial}(p_{im}, n_{im}) \\ \text{logit}(p_{im}) &= \mu_i + \phi_{i,bk} I_{\{m=k\}} \\ \text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{7.2}$$

The log-odds of the class in arm $k = 2$ relative to the baseline class in arm 1, $\phi_{i,bk}$, is expressed as:

$$\phi_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{i2}}, \tau^2)$$

For multi-arm trials of $k > 2$, $\phi_{i,bk}$ for the k^{th} conditional distribution, $k = 3 \ldots na_i$, is defined by:

$$\phi_{i,bk}| \begin{pmatrix} \phi_{i,b2} \\ \vdots \\ \phi_{i,b(k-1)} \end{pmatrix} \sim \text{N} \left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\phi_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right)$$

In keeping with the methodology described in Chapter 6, the study-specific baseline intervention means, μ_i, were assumed to have a Normal$(0, 10^3)$ prior distribution for continuous outcomes, and Normal$(0, 10^2)$ prior distribution for binary outcomes on the log odds scale. For both continuous and binary outcomes, relative class effect for the reference class for the entire network, $m = 1$, which in this example was the control interventions, were set to 0 such that $d_{11} = 0$. For an intervention class, $m = 2, ..., n_m$, basic parameters, d_{1m}, were assumed to have vague prior distributions, such that $d_{1m} \sim \text{Normal}(0, 10^3)$ for continuous outcomes, and $d_{1m} \sim \text{Normal}(0, 10^2)$ for binary outcomes on a log odds scale.

7.3.2 Hierarchical network meta-analysis

Building on the general NMA frameworks presented in Equation (6.1) and (6.2) of Chapter 6, a random effects model was used to estimate the effect of each individual intervention, accounting for the similarity between different formulations (e.g. oral, intravesical, transdermal etc.) of the same class of interventions.

7.3.2.1 Continuous outcomes

For study i evaluating intervention j belonging to class m, a continuous outcome, y_{ij_m}, was assumed to follow a normal distribution with mean equal to the underlying effect, θ_{ij_m}, and observed standard error, se_{ij_m}, given in Equation (7.3). For 2-arm trials, logistic regression models were used to express $\delta^*_{i,bk}$ as normally distributed parameters with mean equal to the relative difference, $d_{t_{ib}t_{ik}}$, of the intervention in the k^{th} arm, t_{ik}, compared to a study-specific reference intervention in arm 1, t_{ib}, of the i^{th} study, with between-study variance τ^2. For multi-arm trials of $k > 2$, the correlation between-arms of the same study was accounted for using conditional distributions as further described in Section 3.4.5. The relative effect of the study-specific reference treatment in arm 1 relative to itself, $\delta^*_{i,b1}$, is set to 0 and as such the set of conditional univariate distributions begin with the relative effect of the intervention in arm 2 relative to the control arm, $\delta^*_{i,b2}$. Therefore the hierarchical NMA model for continuous outcomes is given by:

$$\begin{aligned} y_{ij_m} &\sim \text{Normal}(\theta_{ij_m}, se^2_{ij_m}) \\ \theta_{ij_m} &= \mu_i + \delta^*_{i,bk} I_{\{j_m=k\}} \\ \text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{7.3}$$

The study-specific relative treatment effect, $\delta^*_{i,bk}$, for arm $k = 2$ is expressed as:

$$\delta^*_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{i2}}, \tau^2)$$

For multi-arm trials of $k > 2$, the study-specific relative treatment effect, $\delta^*_{i,bk}$, for the k^{th} conditional distribution, $k = 3 \ldots na_i$, is defined by:

$$\delta^*_{i,bk} | \begin{pmatrix} \delta^*_{i,b2} \\ \vdots \\ \delta^*_{i,b(k-1)} \end{pmatrix} \sim \text{N}\left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\delta^*_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right)$$

7.3.2.2 Binary outcomes

For binary outcomes, the number of reported events, r_{ij_m}, for an intervention j belonging to class m within the i^{th} study was assumed to follow a binomial distribution with denominator, n_{ij_m} and probability p_{ij_m}, given in Equation (7.4). For 2-arm trials, logistic regression models were used to express $\phi^*_{i,bk}$ on the log scale as normally distributed parameters with mean equal to the relative difference, $d_{t_{ib}t_{ik}}$, of the intervention in the k^{th} arm, t_{ik}, compared to a study-specific reference intervention in arm 1, t_{ib}, of the i^{th} study, with between-study variance τ^2. For multi-arm trials, $k > 2$, a multi-arm correction is applied and expressed in terms of an appropriate set of conditional univariate distributions (see Section 3.4.5 for further details). The relative effect of the study-specific reference treatment in arm 1 (the control arm) relative to itself, $\phi^*_{i,b1}$, on the log scale is set to 0 and as such the set of conditional univariate distributions begin with the relative effect of the intervention in arm 2 relative to the control arm, $\phi^*_{i,b2}$. The hierarchical NMA model for binary outcomes is thus given by:

$$
\begin{aligned}
r_{ij_m} &\sim \text{Binomial}(p_{ij_m}, n_{ij_m}) \\
\text{logit}(p_{ij_m}) &= \mu_i + \phi^*_{i,bk} I_{\{j_m=k\}} \\
\text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\tag{7.4}
$$

The log-odds of the treatment in arm $k = 2$ relative to the baseline class in arm 1, $\phi^*_{i,bk}$, is expressed as:

$$\phi^*_{i,b2} \sim \text{Normal}(d_{t_{ib}t_{i2}}, \tau^2)$$

For multi-arm trials of $k > 2$, $\phi^*_{i,bk}$ for the k^{th} conditional distribution, $k = 3 \ldots na_i$, is defined by:

$$\phi^*_{i,bk} \Bigg| \begin{pmatrix} \phi^*_{i,b2} \\ \vdots \\ \phi^*_{i,b(k-1)} \end{pmatrix} \sim \mathrm{N}\left((d_{1,t_{ik}} - d_{1,t_{ib}}) + \frac{1}{k-1} \sum_{q=1}^{k-1} \left[\phi^*_{i,1q} - \left(d_{1,t_{iq}} - d_{1,t_{ib}} \right) \right], \frac{k}{2(k-1)} \tau^2 \right)$$

Similarly to the general model described in Section 3.4.5, for trials that compare interventions A and B belonging to class c, $d_{t_{ib}t_{ik}} = d_{A_cB_c}$. The pooled treatment effect of treatment A_c relative to treatment B_c is given by:

$$d_{A_cB_c} = d_{1_1B_c} - d_{1_1A_c} \tag{7.5}$$

As described in Chapter 3, the intervention effect of the reference treatment for the entire treatment network, $j = 1_1$, usually a placebo or control intervention, was set to 0 such that $d_{1_11_1} = 0$. However, following Dakin et al. (2011) and Warren et al. (2014) the basic parameters for relative treatment effects, $d_{1_1j_m}$, of intervention j within class m, relative to the reference treatment were assumed to follow a normal distribution with a class-specific mean (μ_m) and variance (σ^2_m):

$$d_{1_1j_m} \sim \mathrm{Normal}(\mu_m, \sigma^2_m) \tag{7.6}$$

where μ_m denotes the pooled effect estimate for the m^{th} class of interventions, and assumed to follow a Normal$(0, 10^3)$ distribution for continuous outcomes and Normal$(0, 10^2)$ distribution for binary outcomes. The class-specific between-intervention standard deviation, σ_m, was assumed to follow a Uniform(0,5) distribution for continuous outcomes, and Uniform(0,2) distribution for binary outcomes. Sensitivity to the choice of prior distribution for σ_m was explored and further described in Section 7.3.7. The functional parameters, $d_{A_cB_c}$ (where $A_c > 1$ and $B_c > 1$) are expressed in terms of the basic parameters described in Equation

(7.5).

7.3.3 Incorporating dose constraints

Ordering constraints were placed on multiple doses of interventions, with the assumption that larger doses would have a greater or equal treatment effect compared with its respective lower dose. In this example, continuous outcomes were measured as change from baseline in symptom severity and thus a larger negative intervention effect suggested greater efficacy. For this reason, ordering constraints were placed such that $d_{1\ unit} \geqslant d_{2\ units} \geqslant \cdots \geqslant d_{n\ units}$). These constraints were imposed by assigning an indicator function γ, equal to 1, given by

$$\gamma = \prod {}^{n-1}_{l=1} I(d_l - d_{l+1}) \tag{7.7}$$

$$I(x) = \begin{cases} 1 & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

Using equation (7.7) forces $(d_l - d_{l+1}) \geqslant 0$, and consequently imposes ordering constraints on increasing doses such that $d_{l+1} \leqslant d_l$, remembering that in this example, a larger negative effect suggests greater efficacy. Ordering constraints can be placed in either direction depending on the outcome of interest. For adverse event data, increased doses are expected to have a more severe outcome than its respective lower dose, and therefore, for outcomes including number of patients with adverse events, and discontinuations due to adverse events, intervention constraints were placed in the opposite direction, such that higher doses of an intervention have an increased log-odds of an event (i.e. $d_l \leqslant d_{l+1}$).

7.3.4 Model computation and convergence diagnostics

All models were estimated using WinBUGS 1.4.3 (Spiegelhalter et al., 2003). Example WinBUGS code for the hierarchical NMA models with and without dose constraints for continuous and binary outcomes are given in Section D.1 - D.4 of Appendix D. The results were based on 150,000 MCMC samples, with which the first 10,000 samples were discarded from the analyses in the form of a 'burn-in'. Three individual MCMC chains with disparate starting values were analysed and non-convergence was assessed using Brooks-Gelman-Rubin statistics, autocorrelation, history, trace, and density plots, as described in Section 3.2.2. For illustration purposes, convergence plots were presented for a random sample of basic parameters of interest (Lunn et al., 2012).

7.3.5 Assessing inconsistencies between direct and indirect information

Node-splitting methods described in Section 3.4.5.3 were used to detect inconsistencies between direct and indirect information for all treatment comparisons that form a closed loop in the networks of evidence. Differences between direct and indirect information - more than that attributable to chance alone - were identified using Bayesian p-values, with a threshold of $p < 0.05$ (Dias et al., 2010). Inconsistencies between direct and indirect information would suggest that the additivity assumption underpinning all NMA models may not be satisfied, and the impact of trials that may contribute to inconsistencies should be explored in a series of sensitivity analyses.

7.3.6 Goodness of fit and model selection

DIC statistics were used to compare the model fit of hierarchical NMAs with and without dose constraints. These models were further compared to individual-intervention NMAs, which were previously presented in Chapter 6. DIC statistics

were described in more detail in Section 3.2.3. Briefly, the DIC is a measure of deviance, estimated by the posterior mean of minus twice the log-likelihood, plus the effective number of parameters in the model. The DIC is thus considered as a Bayesian measure of model fit (Spiegelhalter et al., 2002), and is regularly used as a relative measure of model suitability for models that are applied to the same dataset. Collapsing interventions in to their respective treatment classes resulted in the omission of several studies, as further described in Section 7.3.1. Thus, all class-based models were applied to slightly different datasets compared to that of the corresponding individual-intervention and hierarchical NMAs. For this reason, the DIC statistic could not be used as a comparative measure of model suitably for class-based NMAs, and DIC statistics for class-based NMAs were presented in this chapter solely for completeness. The posterior mean residual deviance was used to assess model fit for each of the models described in this chapter. For models with an adequate fit to the data, the posterior mean residual deviance is expected to approximately equal the number of unconstrained data points (Spiegelhalter et al., 2002).

7.3.7 Sensitivity analysis

As mentioned in Section 3.4.2, variance parameters can have a notable impact on the overall effect estimates, and different choices of vague prior distributions should be investigated through a series of sensitivity analyses (Lambert et al., 2005). This is especially true for the between-study variance parameter, τ^2 (Higgins et al., 2009) and class-specific between-intervention variance, σ_m^2 (Owen et al., 2015). Due to the hierarchical nature of the model, both τ^2 and σ^2 may be estimated on fewer data points, and for this reason sensitivity analyses are crucial. Sensitivity analyses evaluated the impact of the choice of prior specification on the between-study standard deviation, τ, and class-specific between-intervention standard deviation, σ, for all hierarchical NMAs. Two alternative distributions were considered: 1) Gamma(0.001,0.001) on the precision scale, and 2) Half-normal(0,1) on the standard deviation scale.

7.4 Results

Table 7.1 displays the goodness-of-fit statistics for each of the models individually. It is worth noting that analyses associated with class-based models were calculated on different datasets which was a consequence of treatment clustering into endonodal treatment classes (Kanters et al., 2014) - that is collapsing interventions in to a single class of interventions (see Section 7.3.1). As a result, model fit statistics for class-based NMAs were solely presented for completeness, and cannot be directly compared with the remaining models. In relation to individual-intervention NMAs, hierarchical models appeared to have a slightly better fit to the data for both efficacy and safety outcomes as illustrated through the reduced DIC statistics. For the number of patients with adverse events, incorporating ordering constraints further improved model fit with respect to both the DIC and the total residual deviance. However, incorporating dose constraints did not improve model fit for mean change from baseline in urinary incontinence episodes. Notably, for both outcomes, class-based models led to increased between-study standard deviation from 0.16 (95%CrI: 0.10,0.23) to 0.27 (95%CrI: 0.21,0.35) for efficacy data, and from 0.23 (95%CrI: 0.14,0.33) to 0.32 (95%CrI: 0.23,0.42) for safety data.

Use of a hierarchical NMA model resulted in a decrease in between-study standard deviation for safety data from 0.23 (95%CrI: 0.14,0.33) for individual-intervention NMA, to 0.20 (95%CrI: 0.13,0.29) for hierarchical models without dose constraints (Table 7.1). For efficacy data, use of a hierarchical NMA did not have a noticeable impact on the between-study standard deviation, though the uncertainty in the between-study standard deviation slightly decreased. Incorporating dose constraints increased the between-study standard deviation for both outcomes, but decreased the uncertainty in the estimate. To a certain extent, a decrease in between-study standard deviation would be expected for hierarchical NMA models as some of the heterogeneity between treatments will be absorbed in the variance for the class effects. Having said that, the between-study standard deviation may

increase if there is substantial heterogeneity within classes of interventions. Reducing this heterogeneity within classes may help to reduce the between-study variability.

TABLE 7.1: Model fit statistics

		N	Residual deviance	DIC	SD (95%CrI)
Efficacy data					
Incontinence episodes	**Individual-intervention NMA**	297	290.2	3336.39	0.16 (0.10,0.23)
	Class-based NMA*	249	249.7	2795.94	0.27 (0.21,0.35)
	Hierarchical NMA	297	286	3315.9	0.16 (0.10,0.22)
	Hierarchical NMA + dose constraints	297	290.6	3318.4	0.19 (0.14,0.25)
Safety data					
Number of patients with adverse events	**Individual-intervention NMA**	243	257	1628.96	0.23 (0.14,0.33)
	Class-based NMA*	196	211.9	1351.36	0.32 (0.23,0.42)
	Hierarchical NMA	243	256.4	1608.74	0.20 (0.13,0.29)
	Hierarchical NMA + dose constraints	243	250.8	1598.68	0.24 (0.14,0.29)

NMA, network meta-analysis; N, number of unconstrained datapoints; DIC, deviance information criterion; SD, between-study standard deviation.

*Denotes that the DIC was calculated on fewer studies and are therefore not comparable to other models. The DIC for these models are displayed solely for completeness.

7.4.1 Efficacy data

Clinical efficacy was assessed based on the mean change from baseline in urinary incontinence, voiding, urgency, and nocturia episodes. For illustration purposes, results from class-based, and hierarchical NMAs for urinary incontinence episodes are presented in this section. Results obtained from age-adjusted hierarchical NMAs for voiding, and unadjusted hierarchical NMAs for urgency, and nocturia are given in Tables D.1, D.2 and D.3 of Appendix D, respectively. A summary of intervention effects across all outcomes are given in Section 7.4.3. For urinary

incontinence episodes, a hierarchical NMA without dose constraints was the best fitting model with the lowest DIC statistic of 3315.9, compared to individual-intervention NMA (DIC: 3336.39) and a hierarchical NMA with dose constraints (DIC: 3318.4).

7.4.1.1 Class-based network meta-analysis

Table 7.2 provides the classes of interventions ranked in order of their estimated efficacy for reducing incontinence episodes. Sacral nerve stimulation appeared to be the most effective class of interventions for the management of urinary incontinence episodes with a posterior median reduction of -9.07 (95%CrI: -11.52,-6.49) episodes from baseline, relative to a control group. Similar to the individual-intervention NMA described in Chapter 6, SNS dominated the analyses as the most effective class of interventions, though this result was based on a single study of only 34 participants and should therefore be interpreted with caution. In this instance, class-based analyses did not increase precision in the overall class effect for SNS. This was largely because SNS did not belong to a broader class of interventions for urinary incontinence episodes and thus, pooling interventions in to classes had no impact on the contributed evidence-base for the class of SNS therapies. However, there was a clear increase in precision for estimated treatment effects for which there were multiple formulations of interventions within the same class, and consequently there were numerous trials that contributed to the overall class effect. For example, in the individual-intervention NMA of Chapter 6, onaBoNT-A 100U injected in to the bladder base and trigone appeared to be amongst the top 10 interventions for the management of urinary incontinence, with an estimated posterior median reduction of -1.39 (95%CrI: -4,1.06), relative to placebo (see Table 6.3). It was clear that in an individual-intervention NMA, onaBoNT-A 100U injected in to the bladder base and trigone appeared to be a promising intervention for the management of urinary incontinence, however, there was considerable uncertainty in the estimated treatment effect as shown through the width of the 95% credible interval. This was largely due to very few studies evaluating this treatment regime

of onaBoNT-A for urinary incontinence; and thus, it can be difficult to make inferences regarding treatment recommendations in a decision making framework. Collapsing all onaBoNT-A treatment regimes in to a class-based NMA (Table 7.2) resulted in a similar treatment effect to that of the individual-intervention NMA but obtained with increased precision in the class effect estimate by 6-fold (estimated posterior median difference relative to placebo: -1.98, 95%CrI: -2.4,-1.56). However, using this approach loses the interpretability of the individual intervention effects, and inferences regarding particular doses and treatment regimes are lost. Consequently, in a decision making context, treatment recommendations can only be made at the class-level.

TABLE 7.2: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to a control group obtained from a class-based network meta-analysis

Treatment pathway	Class of treatments	Number of studies	Number of participants	Median difference† (95%CrI)	Rank (95%CrI)	p(Best)
	Sacral nerve stimulation	1	34	-9.07 (-11.52,-6.49)	1 (1,2)	1
	OnaBoNT-A	7	793	-1.98 (-2.4,-1.56)	4 (2,6)	0
	Solifenacin/trospium + placebo injection	1	118	-2.08 (-3.1,-1.05)	4 (2,8)	0
	Electrostimulation+PFMT+BT	1	25	-2.14 (-3.16,-1.13)	4 (2,7)	0
	Oxybutynin + salivary pastilles	1	8	-1.91 (-3.67,-0.13)	5 (2,31)	0
	Tolterodine + neurostimulation	1	20	-1.25 (-1.76,-0.75)	7 (5,11)	0
	Trospium + physiotherapy	2	32	-1.36 (-2.16,-0.57)	7 (3,15)	0
	Solifenacin	11	3284	-0.76 (-0.96,-0.57)	11 (8,16)	0
	Tolterodine + oestrogen	1	40	-0.77 (-1.34,-0.19)	11 (6,30)	0
	Terodiline	3	126	-0.67 (-1.24,-0.09)	13 (7,33)	0
	Mirabegron	8	3884	-0.6 (-0.81,-0.4)	15 (9,24)	0
	Darifenacin	2	542	-0.5 (-0.95,-0.04)	18 (8,34)	0
	Elocalcitol	1	171	-0.5 (-1.13,0.13)	18 (7,38)	0
	Physiotherapy	3	70	-0.5 (-1,-0.02)	18 (8,34)	0
	Cizolirtine citrate	2	88	-0.51 (-1.13,0.11)	19 (7,37)	0
	Tolterodine + pilocarpine	1	130	-0.49 (-0.89,-0.08)	19 (9,33)	0
	Fesoterodine	11	6348	-0.47 (-0.64,-0.32)	20 (13,28)	0
	Imidafenacin	4	730	-0.46 (-0.79,-0.14)	20 (10,32)	0
	PFMT + BT	2	67	-0.47 (-1.02,0.08)	20 (8,37)	0
	Propiverine	8	2267	-0.48 (-0.69,-0.27)	20 (12,29)	0
	Tolterodine	43	11083	-0.46 (-0.57,-0.36)	20 (15,26)	0
	Tolterodine + PFMT	1	223	-0.46 (-1.15,0.22)	20 (7,39)	0
	Tolterodine + BT	2	295	-0.44 (-0.96,0.07)	21 (8,37)	0
	Trospium	4	595	-0.43 (-0.85,0)	22 (10,35)	0
	Oxybutynin	21	2889	-0.4 (-0.55,-0.24)	23 (16,30)	0
	Solabegron	1	173	-0.4 (-0.85,0.05)	23 (9,36)	0
	Pregabalin + tolterodine	1	84	-0.28 (-0.96,0.39)	27 (9,41)	0
	Duloxetine	1	81	-0.26 (-0.84,0.32)	28 (10,41)	0
	Electrostimulation + vaginal oestrogen cream	1	102	-0.22 (-0.79,0.36)	29 (10,40)	0
	BT	4	147	-0.2 (-0.5,0.12)	30 (19,38)	0
	Electrostimulation	5	209	-0.21 (-0.52,0.1)	30 (18,37)	0
	ZD0947IL	1	92	-0.1 (-1.01,0.82)	33 (8,43)	0
	LipoBoNTA	1	29	-0.06 (-0.99,0.88)	34 (8,43)	0
	Oxybutynin + BT	2	44	-0.03 (-0.74,0.69)	34 (11,43)	0
	Control	81	14441	NA	35 (31,39)	0
	Estradiol	1	20	0 (-0.48,0.48)	35 (19,42)	0
	Pregabalin	1	41	0.02 (-0.91,0.93)	35 (9,44)	0
	Reflexology	1	54	0 (-0.54,0.54)	35 (17,42)	0
	Tarafenacin	1	153	-0.01 (-0.74,0.72)	35 (11,43)	0
	Vaginal oestrogen cream	1	98	0.09 (-0.49,0.67)	37 (20,43)	0
	Emepronium bromide	1	19	0.34 (-0.41,1.08)	40 (23,44)	0
	Flavoxate chloride	1	19	0.33 (-0.43,1.08)	40 (22,44)	0
	ONO-8539	1	252	0.35 (-0.3,1)	40 (27,44)	0
	Resiniferatoxin	1	34	0.57 (-1.11,2.28)	42 (8,44)	0

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average class rank over all iterations and ranked according to the probability that each class of interventions is the best overall. Due to similarities in class effects and uncertainty around the point estimates, classes are ranked in a different order at each iteration. Therefore, classes that share a rank have a similar class effect on average. p(Best) denotes the probability that each of the classes of interventions is the best overall. The probability best is calculated based on the number of iterations for which the classes of interventions are ranked is first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

7.4.1.2 Hierarchical network meta-analysis

Table 7.3 provides the estimated treatment effects and relative treatment rankings obtained from a hierarchical NMA. The interventions were ranked in order of their estimated efficacy for reducing urinary incontinence episodes. Sacral nerve stimulation appeared to be the most effective intervention with an estimated median reduction of -9.08 (95%CrI: -11.76,-6.52) in the number of incontinence episodes, relative to placebo. It was apparent that treatment effect estimates obtained from hierarchical models mirror the treatment effect estimates obtained from individual-intervention NMA as described in Chapter 6; however, there was a substantial increase in the precision surrounding the individual effect estimates obtained from hierarchical analyses for interventions that belong to a broader class of interventions. For example, in comparison to placebo, onaBoNT-A 100 U injected in to the bladder base and trigone had a similar reduction of –1.93 (95%CrI: -3.09,–0.40) urinary incontinence episodes per 24 hours obtained from hierarchical NMA, compared to – 1.39 (95%CrI: –4.0,1.06) obtained from individual-intervention NMA. Though the point estimates were braodly similar, there was a noticeable increase in precision (i.e. 1/variance) of approximately 88% for the treatment effect estimate obtained from hierarchical models.

7.4.1.3 Hierarchical network meta-analysis incorporating dose constraints

Table 7.4 provides the estimated treatment effects and relative treatment rankings obtained from a hierarchical NMA incorporating dose constraints. In this example, imposing ordering constraints on increasing doses had very little impact on the estimated treatment effects and precision compared to that of the hierarchical NMA without dose constraints. For example, for onaBoNT-A 200U, the estimated median reduction in incontinence obtained from hierarchical NMAs with and without dose constraints was -2.19 (95%CrI: -2.91,-1.68) and -2.08 (95%CrI: -2.86,-1.45) episodes relative to placebo, respectively. It was apparent that there was a slight increase in the precision around the treatment effect estimate but this increase was small in magnitude (approximately 15%). For onaBoNT-A 100U injected in to the bladder base and trigone, the estimated treatment effect obtained from the hierarchical NMA incorporating dose constraints was –1.93 (95% CrI –3.05,-0.31), and thus there was a 2% decrease in precision in the estimated treatment effect compared to results obtained from the unconstrained hierarchical NMA. Overall, incorporating dose constraints for urinary incontinence episodes appeared to have little impact on both the point estimate and precision in the treatment effect estimates.

TABLE 7.3: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis

Treatment Pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95% CrI)	Rank (95%CrI)	p(Best)
	Sacral nerve stimulation	[81]	1	34	-9.08 (-11.76,-6.52)	1 (1,1)	1
	OnaBoNT-A 200U trigone sparing	[73]	3	114	-2.08 (-2.86,-1.45)	4 (2,8)	0
	Electrostimulation + PFMT + BT	[97]	1	25	-2.16 (-3.11,-1.2)	4 (2,10)	0
	Oxybutynin IR 2.5mg b.i.d + salivary pastilles	[98]	1	8	-2.01 (-3.81,-0.24)	5 (2,66)	0
	OnaBoNT-A 100u trigone sparing	[72]	5	716	-1.93 (-2.34,-1.52)	5 (3,8)	0
	Solifenacin/trospium + placebo injection	[100]	1	118	-2.03 (-3.01,-1.05)	5 (2,11)	0
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-1.89 (-2.61,-0.98)	6 (2,12)	0
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-1.93 (-3.09,-0.4)	6 (2,53)	0
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-1.29 (-1.69,-0.89)	9 (6,15)	0
	Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	2	32	-1.05 (-1.94,-0.17)	11 (5,69)	0
	Solifenacin ER 10mg q.d	[30]	4	870	-0.81 (-1.06,-0.61)	17 (10,32)	0
	Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg twice/week	[99]	1	40	-0.75 (-1.24,-0.25)	20 (9,66)	0
	Solifenacin ER 5 - 10mg q.d	[31]	3	1312	-0.74 (-0.95,-0.52)	21 (12,41)	0
	Solifenacin ER 5mg q.d	[29]	6	725	-0.73 (-0.93,-0.51)	22 (13,42)	0
	Solifenacin ER 5 - 15mg q.d	[34]	1	377	-0.73 (-1.01,-0.35)	22 (11,59)	0
	Oxybutynin IR 3mg t.i.d	[19]	1	244	-0.71 (-1.07,-0.36)	23 (11,57)	0
	Tolterodine ER 4mg q.d + BT	[87]	1	154	-0.69 (-1.37,-0.07)	24 (9,77)	0
	Terodiline 25mg b.i.d	[28]	3	126	-0.67 (-1.22,-0.11)	26 (10,75)	0
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-0.67 (-1.32,-0.16)	26 (9,72)	0
	Fesoterodine ER 8mg q.d	[26]	5	2266	-0.66 (-0.86,-0.45)	27 (14,48)	0
	Oxybutynin ER 10mg q.d	[8]	1	185	-0.66 (-1.15,-0.26)	27 (10,65)	0
	Mirabegron IR 100mg b.i.d	[48]	1	37	-0.64 (-1.11,-0.38)	28 (11,55)	0
	Darifenacin ER 30mg q.d	[38]	1	115	-0.64 (-1.23,-0.12)	29 (10,74)	0
	Trospium chloride ER 60mg q.d	[44]	2	565	-0.63 (-1.05,-0.2)	29 (11,70)	0
	Mirabegron ER 25mg q.d	[50]	3	555	-0.63 (-0.89,-0.42)	29 (14,51)	0
	Propiverine ER 30mg q.d	[42]	1	391	-0.62 (-1.16,-0.26)	30 (10,66)	0
	Mirabegron ER 100mg q.d	[52]	5	1515	-0.62 (-0.81,-0.44)	30 (17,49)	0
	Mirabegron IR 150mg b.i.d	[49]	1	41	-0.61 (-0.95,-0.26)	31 (13,65)	0
	Mirabegron ER 200mg q.d	[53]	1	166	-0.61 (-0.91,-0.31)	31 (14,61)	0
	Mirabegron ER 50mg q.d	[51]	8	2081	-0.59 (-0.77,-0.42)	33 (20,52)	0
	Solabegron IR 125mg b.i.d	[55]	1	85	-0.59 (-0.92,-0.25)	33 (13,67)	0
	Cizolirtine citrate 400mg b.i.d	[57]	2	68	-0.59 (-1.12,-0.06)	33 (10,77)	0
	Oxybutynin IR 5mg t.i.d	[7]	4	194	-0.57 (-0.95,-0.24)	35 (13,66)	0
	Elocalcitol 75mg	[70]	1	84	-0.57 (-1.17,0.02)	35 (10,80)	0
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	3	150	-0.51 (-1.03,-0.09)	41 (12,75)	0
	Darifenacin ER 15mg q.d.	[40]	2	319	-0.51 (-0.95,-0.09)	41 (13,76)	0
	Propiverine IR 15mg b.i.d	[43]	2	495	-0.51 (-0.91,-0.19)	41 (14,71)	0
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	-0.49 (-1.13,0.01)	43 (11,79)	0
	Tolterodine ER 4mg q.d	[4]	27	7521	-0.5 (-0.6,-0.4)	43 (31,55)	0
	Oxybutynin gel 56mg/day	[135]	1	198	-0.48 (-1.01,-0.03)	44 (12,77)	0
	Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	1	130	-0.48 (-0.79,-0.17)	44 (18,72)	0
	PFMT + BT	[89]	2	67	-0.49 (-1.03,0.05)	44 (12,80)	0
	Fesoterodine ER 4mg q.d	[25]	6	2597	-0.46 (-0.65,-0.29)	46 (27,65)	0
	Tolterodine IR 2mg b.i.d	[5]	18	3312	-0.45 (-0.57,-0.31)	48 (34,63)	0
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	-0.45 (-1.07,0.17)	48 (11,86)	0
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.44 (-0.83,-0.07)	49 (17,76)	0
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.44 (-1.66,0.88)	49 (7,96)	0
	Propiverine ER 20mg q.d	[41]	6	1381	-0.43 (-0.62,-0.23)	50 (29,69)	0
	Tolterodine IR 1mg b.i.d	[6]	3	250	-0.43 (-0.64,-0.09)	50 (28,76)	0
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	37	-0.43 (-1.5,0.48)	50 (8,92)	0
	Imidafenacin 0.1mg b.i.d	[36]	4	563	-0.41 (-0.71,-0.13)	51 (23,74)	0
	Elocalcitol 150mg	[69]	1	87	-0.42 (-1.03,0.19)	51 (12,87)	0
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	-0.4 (-0.83,0.03)	53 (17,80)	0
	Oxybutynin trandermal 3.9mg/day	[10]	3	292	-0.39 (-0.67,-0.11)	54 (26,74)	0
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	-0.39 (-0.81,0.03)	54 (17,80)	0
	Darifenacin ER 7.5mg q.d	[39]	1	108	-0.37 (-0.87,0.19)	55 (15,87)	0
	Oxbutynin patch 73.5mg	[15]	1	391	-0.34 (-0.63,-0.04)	58 (29,78)	0
	Oxybutynin ER 15mg q.d	[9]	1	53	-0.34 (-0.85,0.22)	58 (16,87)	0
	Oxybutynin ER 5 - 30mg/day	[22]	3	109	-0.33 (-0.72,0.14)	59 (24,82)	0
	Fesoterodine ER 4 - 8mg q.d	[27]	4	1485	-0.33 (-0.56,-0.09)	59 (35,77)	0
	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	3	70	-0.34 (-0.89,0.22)	59 (15,86)	0
	Oxybutynin gel 84mg/day	[134]	1	211	-0.32 (-0.77,0.18)	60 (20,86)	0
	Bladder Training (BT)/Behaviour Therapy	[85]	4	147	-0.33 (-0.74,0.11)	60 (23,82)	0
	Oxybutynin IR 5 - 20mg	[23]	1	52	-0.31 (-0.82,0.34)	61 (17,89)	0
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.31 (-0.77,0.19)	61 (20,87)	0
	Oxybutynin transdermal 1.3mg/day	[11]	2	166	-0.29 (-0.69,0.2)	63 (25,86)	0
	Duloxetine 40mg b.i.d	[65]	1	81	-0.26 (-0.76,0.24)	65 (20,88)	0
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.25 (-1.22,0.71)	66 (10,95)	0
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-0.22 (-0.68,0.44)	68 (27,91)	0
	Solabegron IR 50mg b.i.d	[54]	1	88	-0.22 (-0.56,0.12)	68 (36,85)	0

TABLE 7.3: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis (cont.)

	Tolterodine IR 2mg b.i.d + BT	[93]	1	141	-0.21 (-0.84,0.43)	68 (16,92)	0
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-0.17 (-0.46,0.12)	70 (47,84)	0
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	-0.15 (-0.56,0.41)	72 (37,91)	0
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.11 (-0.82,0.58)	74 (17,94)	0
	ZD0947IL 25mg/day	[58]	1	92	-0.11 (-0.96,0.76)	74 (13,96)	0
	Oxybutynin IR 5mg b.i.d	[18]	1	116	-0.1 (-0.52,0.46)	75 (41,92)	0
	Lipo-BoNTA 200U	[138]	1	29	-0.06 (-0.94,0.82)	76 (13,96)	0
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	47	-0.02 (-0.57,0.54)	79 (35,93)	0
	Placebo	[1]	77	14282	NA	80 (72,86)	0
	Pregabalin 150mg b.i.d	[62]	1	41	0 (-0.88,0.9)	80 (15,96)	0
	Estradiol 25mg	[68]	1	20	0 (-0.38,0.38)	80 (54,92)	0
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	0.01 (-0.5,0.51)	80 (42,93)	0
	Electrostimulation	[80]	3	150	0.02 (-0.32,0.36)	80 (60,90)	0
	Percutaneous tibial nerve stimulation	[83]	2	59	0.01 (-1.03,1.15)	80 (12,96)	0
	Control	[2]	3	142	0.04 (-0.81,0.86)	81 (18,95)	0
	Tarafenacin 0.2mg q.d	[90]	1	77	0.07 (-0.65,0.8)	82 (28,96)	0
	ONO-8539 100mg b.i.d	[60]	1	83	0.06 (-0.71,0.76)	82 (23,95)	0
	Reflexology	[71]	1	54	0.05 (-0.91,0.97)	82 (14,96)	0
	Trospium chloride IR 15mg t.i.d	[46]	2	30	0.21 (-0.66,1.07)	87 (29,97)	0
	Flavoxate chloride 200mg q.d	[64]	1	19	0.32 (-0.36,1.01)	89 (56,97)	0
	Emepronium bromide ER 200mg q.d	[63]	1	19	0.34 (-0.35,1.01)	90 (58,97)	0
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	0.32 (-0.19,0.83)	90 (70,96)	0
	Oxybutynin ER 5-30mg q.d + BT	[86]	1	32	0.39 (-0.62,1.47)	91 (32,97)	0
	ONO-8539 30mg b.i.d	[59]	1	87	0.48 (-0.21,1.23)	92 (68,97)	0
	ONO-8539 300mg b.i.d	[61]	1	82	0.45 (-0.24,1.17)	92 (66,97)	0
	Sham therapy	[3]	1	17	0.61 (-0.48,1.86)	94 (46,97)	0
	Resiniferatoxin 50nM	[67]	1	34	0.59 (-1.02,2.17)	94 (12,97)	0

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

TABLE 7.4: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis incorporating dose constraints

Treatment Pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95% CrI)	Rank (95%CrI)	p(Best)
	Sacral nerve stimulation	[81]	1	34	-9.11 (-11.7,-6.47)	1 (1,1)	1
	OnaBoNT-A 200U trigone sparing	[73]	3	114	-2.19 (-2.91,-1.68)	4 (2,6)	0
	Electrostimulation + PFMT + BT	[97]	1	25	-2.17 (-3.16,-1.19)	4 (2,9)	0
	Solifenacin/trospium + placebo injection	[100]	1	118	-1.99 (-2.98,-0.99)	5 (2,12)	0
	OnaBoNT-A 100u trigone sparing	[72]	5	716	-1.89 (-2.3,-1.47)	6 (4,8)	0
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-1.89 (-2.62,-0.92)	6 (2,13)	0
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-1.93 (-3.05,-0.31)	6 (2,65)	0
	Oxybutynin IR 2.5mg b.i.d + salivary pastilles	[98]	1	8	-1.91 (-3.71,-0.11)	6 (2,78)	0
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-1.29 (-1.74,-0.84)	9 (6,16)	0
	Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	2	32	-1.08 (-2,-0.18)	11 (5,74)	0
	Solifenacin ER 10mg q.d	[30]	4	870	-0.86 (-1.09,-0.64)	15 (10,27)	0
	Solifenacin ER 5 - 10mg q.d	[31]	3	1312	-0.76 (-0.95,-0.57)	19 (13,34)	0
	Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg twice/week	[99]	1	40	-0.75 (-1.28,-0.23)	19 (9,72)	0
	Darifenacin ER 30mg q.d	[38]	1	115	-0.74 (-1.3,-0.26)	20 (9,69)	0
	Solifenacin ER 5 - 15mg q.d	[34]	1	377	-0.72 (-1.04,-0.31)	21 (11,65)	0
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-0.72 (-1.32,-0.29)	21 (9,67)	0
	Propiverine ER 30mg q.d	[42]	1	391	-0.69 (-1.22,-0.36)	23 (9,58)	0
	Mirabegron ER 200mg q.d	[53]	1	166	-0.69 (-0.97,-0.48)	23 (12,40)	0
	Tolterodine ER 4mg q.d + BT	[87]	1	154	-0.69 (-1.38,-0.04)	23 (9,81)	0
	Solifenacin ER 5mg q.d	[29]	6	725	-0.68 (-0.87,-0.46)	24 (15,46)	0
	Mirabegron IR 150mg b.i.d	[49]	1	41	-0.67 (-1.08,-0.42)	24 (11,48)	0
	Terodiline 25mg b.i.d	[28]	3	126	-0.67 (-1.23,-0.1)	25 (9,79)	0
	Fesoterodine ER 8mg q.d	[26]	5	2266	-0.64 (-0.85,-0.45)	27 (14,49)	0
	Mirabegron ER 100mg q.d	[52]	5	1515	-0.64 (-0.82,-0.45)	27 (17,45)	0
	Trospium chloride ER 60mg q.d	[44]	2	565	-0.63 (-1.08,-0.19)	28 (11,75)	0
	Cizolirtine citrate 400mg b.i.d	[57]	2	68	-0.61 (-1.17,-0.06)	29 (10,80)	0
	Mirabegron IR 100mg b.i.d	[48]	1	37	-0.6 (-0.89,-0.31)	31 (15,65)	0
	Mirabegron ER 50mg q.d	[51]	8	2081	-0.6 (-0.79,-0.4)	31 (19,51)	0
	Solabegron IR 125mg b.i.d	[55]	1	85	-0.59 (-0.98,-0.21)	31 (12,74)	0
	Elocalcitol 150mg	[69]	1	87	-0.59 (-1.16,-0.02)	31 (10,82)	0
	ONO-8539 300mg b.i.d	[61]	1	82	-0.58 (-0.98,-0.3)	32 (12,61)	0
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	37	-0.56 (-1.54,0.21)	34 (8,88)	0
	Propiverine IR 15mg b.i.d	[43]	2	495	-0.55 (-0.96,-0.21)	35 (13,74)	0
	Mirabegron ER 25mg q.d	[50]	3	555	-0.56 (-0.76,-0.34)	35 (22,62)	0
	Darifenacin ER 15mg q.d.	[40]	2	319	-0.52 (-0.94,-0.08)	39 (13,79)	0
	Oxybutynin IR 5mg t.i.d	[7]	4	194	-0.51 (-0.88,-0.29)	39 (14,59)	0
	PFMT + BT	[89]	2	67	-0.5 (-1.05,0.04)	40 (11,83)	0
	Tolterodine ER 4mg q.d	[4]	27	7521	-0.5 (-0.61,-0.39)	41 (28,57)	0
	ONO-8539 100mg b.i.d	[60]	1	83	-0.48 (-0.81,-0.23)	42 (17,67)	0
	Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	1	130	-0.48 (-0.84,-0.13)	42 (15,78)	0
	Oxybutynin ER 15mg q.d	[9]	1	53	-0.47 (-0.86,-0.27)	43 (15,64)	0
	Fesoterodine ER 4 - 8mg q.d	[27]	4	1485	-0.46 (-0.63,-0.3)	44 (26,68)	0
	Tolterodine IR 2mg b.i.d	[5]	18	3312	-0.45 (-0.59,-0.31)	46 (31,67)	0
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	-0.45 (-1.11,0.2)	46 (11,89)	0
	Oxybutynin trandermal 3.9mg/day	[10]	3	292	-0.43 (-0.66,-0.23)	49 (26,68)	0
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	-0.42 (-0.74,-0.2)	50 (20,71)	0
	Propiverine ER 20mg q.d	[41]	6	1381	-0.42 (-0.63,-0.21)	50 (28,74)	0
	Oxybutynin IR 5mg b.i.d	[18]	1	116	-0.41 (-0.66,-0.24)	51 (27,67)	0
	Imidafenacin 0.1mg b.i.d	[36]	4	563	-0.41 (-0.72,-0.11)	51 (21,79)	0
	Elocalcitol 75mg	[70]	1	84	-0.42 (-0.99,0.15)	51 (13,88)	0
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	3	150	-0.4 (-0.78,-0.14)	53 (18,76)	0
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	-0.39 (-0.8,-0.1)	54 (17,78)	0
	Fesoterodine ER 4mg q.d	[25]	6	2597	-0.4 (-0.56,-0.22)	54 (34,75)	0
	Oxybutynin ER 10mg q.d	[8]	1	185	-0.4 (-0.7,-0.2)	54 (25,72)	0
	Oxybutynin gel 56mg/day	[135]	1	198	-0.38 (-0.62,-0.15)	55 (32,74)	0
	Oxybutynin IR 3mg t.i.d	[19]	1	244	-0.37 (-0.58,-0.19)	58 (35,73)	0
	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	3	70	-0.36 (-0.93,0.21)	58 (14,88)	0
	Oxbutynin patch 73.5mg	[15]	1	391	-0.35 (-0.59,-0.1)	59 (33,78)	0
	Oxybutynin ER 5 - 30mg/day	[22]	3	109	-0.36 (-0.64,-0.05)	59 (29,79)	0
	Tolterodine IR 1mg b.i.d	[6]	3	250	-0.36 (-0.54,-0.03)	59 (37,83)	0
	Oxybutynin IR 5 - 20mg	[23]	1	52	-0.35 (-0.66,0.03)	60 (26,83)	0
	Oxybutynin gel 84mg/day	[134]	1	211	-0.34 (-0.63,0)	60 (29,82)	0
	Bladder Training (BT)/Behaviour Therapy	[85]	4	147	-0.35 (-0.71,0.03)	60 (23,82)	0
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	-0.33 (-0.52,-0.05)	62 (41,79)	0
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	-0.32 (-0.56,-0.03)	64 (37,81)	0
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.3 (-0.51,-0.02)	66 (44,81)	0
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.29 (-1.22,0.62)	67 (10,95)	0
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-0.28 (-0.49,0.12)	68 (46,86)	0
	Darifenacin ER 7.5mg q.d	[39]	1	108	-0.28 (-0.75,0.27)	68 (21,91)	0
	Duloxetine 40mg b.i.d	[65]	1	81	-0.26 (-0.8,0.28)	69 (17,91)	0
	Oxybutynin transdermal 1.3mg/day	[11]	2	166	-0.25 (-0.46,0.1)	70 (52,86)	0
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-0.25 (-0.45,0.03)	70 (54,83)	0
	Tolterodine IR 2mg b.i.d + BT	[93]	1	141	-0.22 (-0.88,0.47)	72 (15,93)	0
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.21 (-0.58,0.26)	73 (32,91)	0
	Solabegron IR 50mg b.i.d	[54]	1	88	-0.21 (-0.6,0.18)	73 (30,89)	0

TABLE 7.4: Estimated posterior median difference (and 95% credible interval) in change from baseline for incontinence episodes relative to placebo obtained from a hierarchical network meta-analysis incorporating dose constraints (cont.)

	Tarafenacin 0.4mg q.d	[82]	1	76	-0.19 (-0.86,0.53)	74 (15,93)	0
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.16 (-0.92,1.02)	76 (14,97)	0
	ZD0947IL 25mg/day	[58]	1	92	-0.11 (-0.99,0.79)	79 (12,96)	0
	Lipo-BoNTA 200U	[138]	1	29	-0.05 (-0.96,0.86)	81 (13,96)	0
	ONO-8539 30mg b.i.d	[59]	1	87	-0.03 (-0.56,0.73)	81 (36,96)	0
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-0.02 (-0.57,0.54)	82 (33,94)	0
	Electrostimulation	[80]	3	150	-0.01 (-0.36,0.35)	82 (59,91)	0
	Placebo	[1]	77	14282	NA	83 (76,89)	0
	Pregabalin 150mg b.i.d	[62]	1	41	0.01 (-0.9,0.91)	83 (14,96)	0
	Estradiol 25mg	[68]	1	20	0 (-0.42,0.43)	83 (48,93)	0
	Percutaneous tibial nerve stimulation	[83]	2	59	0 (-1.06,1.2)	83 (11,96)	0
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	47	0 (-0.57,0.57)	83 (34,95)	0
	Control	[2]	3	142	0.03 (-0.82,0.86)	84 (17,96)	0
	Reflexology	[71]	1	54	0.03 (-0.94,0.99)	84 (13,96)	0
	Tarafenacin 0.2mg q.d	[90]	1	77	0.14 (-0.54,0.87)	87 (37,96)	0
	Trospium chloride IR 15mg t.i.d	[46]	2	30	0.17 (-0.7,1.08)	88 (23,97)	0
	Emepronium bromide ER 200mg q.d	[63]	1	19	0.34 (-0.37,1.05)	91 (56,97)	0
	Flavoxate chloride 200mg q.d	[64]	1	19	0.33 (-0.39,1.05)	91 (55,97)	0
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	0.29 (-0.27,0.85)	91 (70,96)	0
	Oxybutynin ER 5-30mg q.d + BT	[86]	1	32	0.34 (-0.64,1.4)	91 (27,97)	0
	Resiniferatoxin 50nM	[67]	1	34	0.58 (-1.07,2.23)	94 (11,97)	0
	Sham therapy	[3]	1	17	0.6 (-0.52,1.92)	95 (40,97)	0

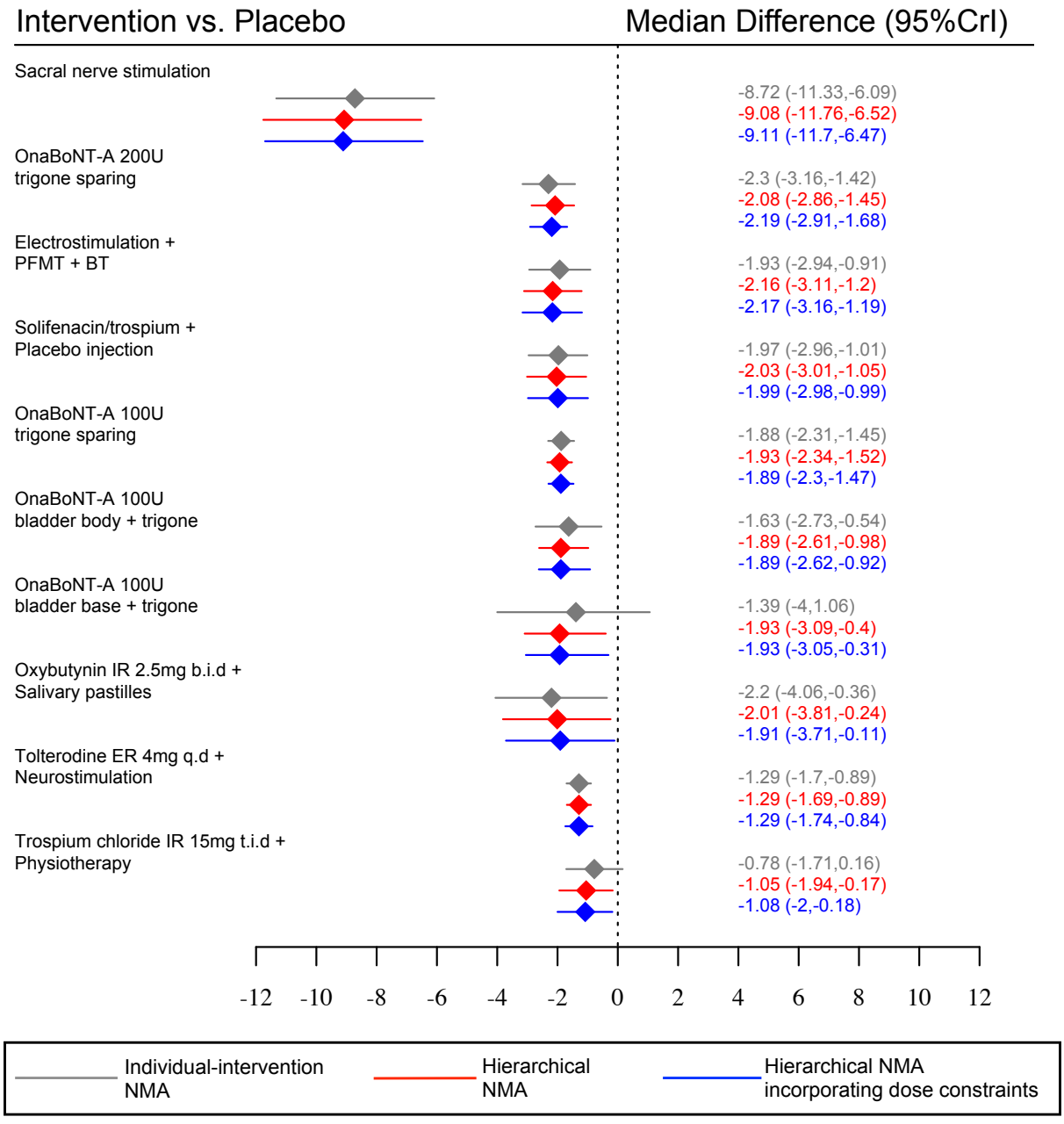
† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

Figure 7.5 illustrates the point estimates and 95% credible intervals obtained from individual-intervention, and hierarchical NMAs. For illustration purposes, results for the top 10 interventions were presented. As shown through the narrower credible intervals, there was an apparent reduction in posterior uncertainty for effect estimates obtained from hierarchical NMAs compared to that of the individual-intervention NMA. This gain in precision was particularly apparent for interventions that belong to a class of interventions, for which there were relatively few trials evaluating the intervention of interest but many studies informing the estimated class effect. For example, in the top 10 interventions for urinary incontinence episodes, onaBoNT-A 100U bladder base and trigone, and onaBoNT-A 100U bladder body and trigone belong to a class of onaBoNT-A interventions, but with only one trial evaluating each of the individual interventions. Using onaBoNT-A 100U bladder base and trigone as an example, it was apparent that there was a large amount of uncertainty associated with the treatment effect estimate obtained from individual-intervention NMA, as shown through the width of the 95% credible intervals which span the point of no difference. Assuming that the treatment effect estimate for this intervention was exchangeable with interventions belonging to the same class of onaBoNT-A interventions, using hierarchical NMAs, sufficiently increased the precision in the treatment effect estimate whilst maintaining the interpretability at the individual intervention level. However, there appeared to be little gain in precision from hierarchical NMAs incorporating dose-response constraints compared to those without dose constraints.

FIGURE 7.5: Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in incontinence episodes relative to placebo between individual-intervention, hierarchical, and dose constraint models for the top 10 interventions



7.4.2 Safety and tolerability data

Regulatory bodies such as NICE in England and Wales would consider both benefits and harms of medical interventions to inform treatment recommendations and clinical guidance. Safety and tolerability of interventions were assessed through the number of patients experiencing adverse events, discontinuations due to adverse events, and discontinuations due to a lack of efficacy. For illustration purposes, results from class-based, and hierarchical models for the number of patients experiencing adverse events are presented in this chapter. A summary of efficacy, safety and tolerability profiles across all outcome measures are presented in Section 7.4.3. For the number of patients with adverse events, a hierarchical NMA with dose constraints was the best fitting model, with the lowest DIC statistic of 1598.08, compared to the hierarchical NMA (DIC: 1608.74) and individual-intervention NMA (DIC: 1628.96).

7.4.2.1 Class-based network meta-analysis

Table 7.5 provides the classes of interventions ranked in order of the estimated posterior odds of a patient reporting an adverse event. Terodiline and duloxetine, as classes of interventions, appeared to be the most hazardous intervention class with estimated rankings of 3 (95%CrI: 1,21) and 3 (95%CrI: 1,17), and estimated probability of being the worst of 0.26 and 0.20, respectively. Though, it is worth noting that there was considerable uncertainty in the treatment rankings, and broadly there was little difference between the top 5 most hazardous classes of interventions. The estimated posterior odds ratio for classes of terodiline and duloxetine were 3.09 (95%CrI: 1.03, 9.66) and 3.05 (95%CrI: 1.35,6.9), relative to placebo, respectively. These effect estimates were comparable to effect estimates obtained from individual-intervention NMAs described in Chapter 6, where terodiline 25mg b.i.d and duloxetine 60mg b.i.d had estimated posterior odds ratios of 3.09 (95%CrI: 1.13,8.8) and 3.06 (95%CrI: 1.55,6.09), relative to a placebo intervention, respectively (Table C.1 of Appendix C). However, the estimated rankings and probability of being the worst intervention overall were vastly

different. For individual-intervention NMA, terodiline 25mg b.i.d and duloxetine 60mg b.i.d had an average rank of 15 (95%CrI: 2,58) and 15 (95%CrI: 5,44), respectively, and propiverine ER 60mg q.d, and propiverine IR 45mg t.i.d appeared to be the most hazardous individual interventions for patients reporting adverse events. Thus, modelling interventions individually or as classes of treatments would lead to vastly different conclusions regarding treatment recommendations. It was apparent that there was a large amount of uncertainty in the estimates obtained from individual-intervention NMA (Table C.1 of Appendix C), as shown through the width of the credible intervals. Pooling interventions in to their respective classes of treatments, increased the precision of the estimated odds ratios but the safety of individual interventions were difficult to ascertain.

TABLE 7.5: Estimated posterior odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from a class-based network meta-analysis

Treatment pathway	Class of treatments	Number of studies	Number of participants	Odds ratio† (95%CrI)	Rank (95%CrI)	p(Best)
	Terodiline	1	46	3.09 (1.03,9.66)	3 (1,21)	0.26
	Duloxetine	1	153	3.05 (1.35,6.9)	3 (1,17)	0.2
	Darifenacin	4	849	2.91 (1.93,4.52)	4 (1,10)	0.09
	Oxybutynin	13	2950	2.64 (2.06,3.4)	5 (2,9)	0.02
	Cizolirtine citrate	2	106	2.44 (1.14,5.24)	6 (1,20)	0.08
	Tarafenacin	1	160	2.19 (0.93,5.12)	7 (1,23)	0.06
	Propiverine	9	2493	2.16 (1.65,2.85)	8 (3,13)	0
	Fesoterodine	10	3730	1.94 (1.54,2.46)	9 (5,15)	0
	Imidafenacin	4	769	1.77 (1.17,2.67)	11 (4,20)	0
	Trospium chloride	6	2024	1.7 (1.21,2.37)	12 (6,19)	0
	ZD0947IL	1	90	1.61 (0.67,3.91)	13 (2,25)	0.02
	OnaBoNT-A	5	533	1.65 (0.88,3.15)	13 (3,23)	0
	Solifenacin	11	3131	1.64 (1.23,2.18)	13 (7,19)	0
	Solifenacin + BT	1	304	1.56 (0.71,3.34)	14 (3,25)	0.01
	Solifenacin/trospium + placebo injection	1	127	1.36 (0.47,3.92)	16 (2,26)	0.02
	Pregabalin	1	105	1.39 (0.65,2.98)	16 (3,25)	0
	Tolterodine	32	6356	1.27 (1.08,1.5)	18 (13,21)	0
	Electrostimulation	4	214	1.17 (0.53,2.63)	19 (5,26)	0
	Pregabalin + tolterodine	1	207	1.19 (0.59,2.41)	19 (6,26)	0
	Mirabegron	10	6689	1.21 (0.96,1.51)	19 (14,23)	0
	Tolterodine + BT	2	275	1.12 (0.58,2.14)	20 (8,26)	0
	Netupitant	1	184	1 (0.37,2.76)	22 (4,26)	0
	ONO-8539	1	263	1 (0.51,1.98)	22 (9,26)	0
	Control	70	12294	NA	22 (19,25)	0
	Solabegron	1	173	0.88 (0.38,2.02)	23 (9,26)	0
	LipoBoNTA	1	29	0.83 (0,151.7)	24 (1,27)	0.23
	Bladder training (BT)	2	103	0.06 (0,0.37)	27 (26,27)	0

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average class rank over all iterations and ranked according to the probability that each class of interventions has the highest incidence of adverse events overall. p(Worst) denotes the probability that the class of interventions had the highest incidence of adverse events. The probability worst is calculated based on the number of iterations for which the classes of interventions were ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

7.4.2.2 Hierarchical network meta-analysis

Table 7.6 provides the estimated posterior odds ratios and relative treatment rankings obtained from a hierarchical NMA. Darifenacin ER 30mg q.d appeared to be the most hazardous intervention with an estimated odds ratio of 6.29 (95%CrI: 3.49,11.44) and relative treatment ranking of 2 (95%CrI: 1,7). Though the effect estimate was comparable to results obtained from individual-intervention NMA (7.08, 95%CrI: 3.92,13.24), use of hierarchical NMA substantially increased the precision in the treatment effect estimate, and consequently increased the precision in the relative treatment rank. In doing so, Darifenacin ER 30mg q.d appeared to be the most hazardous intervention for 40% of MCMC iterations. For individual-intervention NMA, the probability of being the worst intervention overall was spread across the top 10 interventions, with propiverine ER 60mg q.d having the highest probability of being the worst of 20%. Therefore, in this example, borrowing strength between similar interventions using hierarchical NMA allowed more certain inferences with regard to the safety of interventions for decision making.

7.4.2.3 Hierarchical network meta-analysis incorporating dose constraints

Table 7.7 provides the estimated posterior odds ratios and relative treatment rankings obtained from a hierarchical NMA incorporating dose constraints. Imposing ordering constraints slightly increased the precision in the estimated odds ratios. This was particularly true for interventions with which there were many alternative doses evaluated in the network, such as onaBoNT-A sparing the trigone. For example, in hierarchical NMA, onaBoNT-A 150U trigone sparing had an estimated odds ratio of 1.75 (95%CrI: 0.93,3.63) relative to placebo, incorporating ordering constraints on increasing doses increased the precision in the estimated effect by 55% compared to unconstrained hierarchical NMAs with an estimated odds ratio of 1.6 (95%CrI: 0.92,2.66) relative to placebo. Overall, darifenacin ER 30mg q.d appeared to be the most hazardous intervention with the highest estimated odds ratio of 6.24 (95%CrI: 3.51,11.64) and average rank of 2 (95%CrI: 1,7).

TABLE 7.6: Estimated posterior odds ratios (95% credible intervals) for the proportion of patients experiencing adverse events obtained from hierarchical network meta-analysis

Treatment Pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95% CrI)	Rank (95%CrI)	p(Best)
	Darifenacin ER 30mg q.d	[38]	2	151	6.29 (3.49,11.44)	2 (1,7)	0.40
	Imidafenacin 0.25mg b.i.d	[37]	1	76	4.62 (2.27,9.36)	3 (1,22)	0.13
	Tarafenacin 0.4mg q.d	[82]	1	81	3.43 (1.56,7.7)	6 (1,44)	0.05
	Solifenacin ER 20mg q.d	[33]	1	37	3.28 (1.57,7.43)	7 (1,43)	0.04
	Terodiline 25mg b.i.d	[28]	1	46	3.09 (1.15,8.77)	9 (1,59)	0.06
	Duloxetine 60mg b.i.d	[66]	1	153	3.05 (1.6,5.88)	9 (2,43)	0.01
	Oxybutynin IR 5mg b.i.d	[18]	4	415	2.97 (2.22,4.46)	10 (3,23)	0.00
	Oxbutynin patch 73.5mg	[15]	1	572	2.67 (1.97,3.77)	14 (5,30)	0.00
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	2.57 (1.45,4.47)	15 (4,47)	0.00
	Propiverine ER 60mg q.d	[119]	1	43	2.51 (1.62,7.96)	16 (2,41)	0.02
	Propiverine IR 45mg t.i.d	[118]	1	49	2.5 (1.63,7.87)	16 (2,41)	0.02
	Oxybutynin IR 5mg t.i.d	[7]	3	257	2.56 (1.76,3.66)	16 (5,37)	0.00
	Oxybutynin vaginal ring 6mg q.d	[17]	1	147	2.48 (1.57,3.53)	17 (6,43)	0.00
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	2	1018	2.49 (1.7,3.52)	17 (6,38)	0.00
	Cizolirtine citrate 400mg b.i.d	[57]	2	81	2.44 (1.22,4.99)	18 (3,56)	0.00
	Fesoterodine ER 8mg q.d	[26]	4	1187	2.45 (1.87,3.18)	18 (7,35)	0.00
	Oxybutynin vaginal ring 4mg q.d	[16]	1	143	2.38 (1.44,3.35)	19 (7,48)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	2.36 (1.45,3.3)	20 (7,47)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	25	2.33 (0.88,6.25)	21 (2,71)	0.01
	Darifenacin ER 15mg q.d.	[40]	2	321	2.33 (1.54,3.54)	21 (5,45)	0.00
	Propiverine IR 30mg b.i.d	[117]	1	47	2.25 (1.41,5.5)	23 (3,49)	0.00
	Solifenacin ER 5 - 15mg q.d	[34]	1	385	2.19 (1.32,3.71)	24 (5,52)	0.00
	Propiverine IR 15mg t.i.d	[116]	1	149	2.18 (1.46,3.56)	24 (7,47)	0.00
	Solifenacin ER 5 - 10mg q.d	[31]	3	1160	2.18 (1.4,3.43)	24 (6,49)	0.00
	Propiverine ER 20mg q.d	[41]	6	1458	2.15 (1.7,2.8)	25 (12,40)	0.00
	Solifenacin ER 5 - 10mg q.d + BT	[77]	1	304	2.07 (1.04,4.13)	27 (4,65)	0.00
	Darifenacin ER 7.5mg q.d	[39]	2	377	2.09 (1.35,3.23)	27 (7,51)	0.00
	Fesoterodine IR 12mg b.i.d	[122]	1	38	1.98 (1.04,3.94)	30 (5,64)	0.00
	Propiverine ER 30mg q.d	[42]	1	391	1.96 (1.22,2.84)	31 (12,55)	0.00
	Propiverine IR 15mg b.i.d	[43]	3	541	1.91 (1.26,2.65)	32 (15,54)	0.00
	Fesoterodine IR 4mg b.i.d	[120]	1	43	1.88 (0.93,3.38)	33 (7,69)	0.00
	Fesoterodine IR 8mg b.id	[121]	1	47	1.85 (0.9,3.28)	34 (7,70)	0.00
	Trospium chloride IR 45mg t.i.d	[47]	1	828	1.83 (1.22,3.1)	35 (9,56)	0.00
	Fesoterodine ER 4mg q.d	[25]	5	1368	1.82 (1.45,2.26)	35 (21,48)	0.00
	OnaBoNTA 300u trigone sparing	[76]	1	55	1.75 (0.94,3.61)	37 (6,68)	0.00
	OnaBoNTA 150u trigone sparing	[75]	1	50	1.75 (0.93,3.63)	37 (6,68)	0.00
	OnaBoNT-A 100u trigone sparing	[72]	3	186	1.73 (0.94,3.45)	38 (7,68)	0.00
	Trospium chloride IR 20mg b.i.d	[45]	2	596	1.74 (1.24,2.52)	38 (17,55)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1047	1.7 (1.28,2.21)	39 (23,54)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	4	602	1.63 (1.16,2.3)	41 (21,59)	0.00
	Trospium chloride ER 60mg q.d	[44]	2	578	1.64 (1.16,2.28)	41 (22,59)	0.00
	Trospium chloride IR 15mg t.i.d	[46]	1	22	1.61 (0.6,2.88)	42 (11,76)	0.00
	ZD0947IL 25mg/day	[58]	1	90	1.61 (0.78,3.39)	42 (6,73)	0.00
	Solifenacin ER 10mg q.d	[30]	2	375	1.61 (1.06,2.46)	42 (17,64)	0.00
	OnaBoNTA 50u trigone sparing	[74]	1	56	1.56 (0.78,2.99)	44 (10,73)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	3	186	1.56 (0.89,2.73)	44 (14,70)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	1.48 (0.77,2.83)	46 (11,74)	0.00
	Tolterodine IR 4mg b.i.d	[140]	1	16	1.45 (0.97,5.75)	47 (2,68)	0.01
	Tarafenacin 0.2mg q.d	[90]	1	79	1.47 (0.68,3.13)	47 (8,75)	0.00
	Solifenacin/trospium + placebo injection	[100]	1	127	1.42 (0.56,3.75)	48 (5,76)	0.00
	Pregabalin 150mg b.i.d	[62]	1	105	1.38 (0.74,2.56)	50 (15,74)	0.00
	Tolterodine IR 2mg b.i.d	[5]	13	1293	1.32 (1.05,1.71)	52 (38,64)	0.00
	Pregabalin 150mg b.i.d + tolterodine ER 4mg q.d	[102]	1	102	1.3 (0.72,2.38)	52 (19,74)	0.00
	Solifenacin ER 5mg q.d	[29]	6	1133	1.23 (0.87,1.75)	55 (37,72)	0.00
	Tolterodine ER 4mg q.d	[4]	19	4729	1.24 (1.08,1.42)	55 (46,63)	0.00

TABLE 7.6: Estimated posterior odds ratios (95% credible intervals) for the proportion of patients experiencing adverse events obtained from hierarchical network meta-analysis (cont.)

	Tolterodine IR 0.5mg b.i.d	[141]	1	21	1.19 (0.49,2.01)	57 (30,77)	0.00
	Tolterodine IR 1mg b.i.d	[6]	4	297	1.17 (0.79,1.57)	58 (43,73)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	2	275	1.13 (0.65,1.97)	60 (31,76)	0.00
	Lipo-BoNTA 200U	[138]	1	29	1.11 (0,245.3)	61 (1,81)	0.22
	Netupitant 200mg q.d	[112]	1	61	1.1 (0.45,2.89)	61 (10,77)	0.00
	ONO-8539 300mg b.i.d	[61]	1	88	1.11 (0.63,2.05)	61 (28,75)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	65	1.08 (0.8,1.5)	62 (45,73)	0.00
	Mirabegron ER 100mg q.d	[52]	4	1692	1.09 (0.9,1.35)	62 (50,71)	0.00
	Mirabegron ER 200mg q.d	[53]	1	167	1.1 (0.86,1.53)	62 (44,72)	0.00
	Mirabegron ER 25mg q.d	[50]	3	811	1.07 (0.87,1.34)	63 (50,72)	0.00
	Mirabegron ER 50mg q.d	[51]	9	3889	1.06 (0.9,1.26)	63 (54,71)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	1.06 (9.71,1.38)	64 (49,75)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	105	1.05 (0.56,1.91)	64 (33,77)	0.00
	Placebo	[1]	67	12078	NA	67 (59,73)	0.00
	ONO-8539 30mg b.i.d	[59]	1	88	0.96 (0.54,1.7)	68 (39,77)	0.00
	Netupitant 100mg q.d	[111]	1	61	0.94 (0.37,2.44)	69 (18,78)	0.00
	Solabegron IR 50mg b.i.d	[54]	1	88	0.89 (0.44,1.8)	70 (35,78)	0.00
	Netupitant 50mg q.d	[110]	1	62	0.89 (0.34,2.28)	71 (22,78)	0.00
	Solabegron IR 125mg b.i.d	[55]	1	85	0.87 (0.43,1.77)	71 (36,78)	0.00
	ONO-8539 100mg b.i.d	[60]	1	87	0.88 (0.48,1.55)	71 (44,77)	0.00
	Solifenacin ER 2.5mg q.d	[32]	1	41	0.82 (0.3,1.82)	72 (35,78)	0.00
	Percutaneous tibial nerve stimulation	[83]	4	188	0.72 (0.28,1.84)	74 (34,78)	0.00
	Electrostimulation	[80]	2	48	0.34 (0.02,2.01)	78 (29,81)	0.00
	Sham therapy	[3]	2	141	0.13 (0.01,0.83)	79 (73,81)	0.00
	Control	[2]	1	75	0.09 (0,1.72)	80 (39,81)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	2	103	0.04 (0,0.28)	81 (78,81)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment has the highest incidence of discontinuations due to a lack of efficacy overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Worst) denotes the probability that the intervention in question has the highest incidence of discontinuations. The probability worst is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

TABLE 7.7: Estimated posterior odds ratios (95% credible intervals) for the proportion of patients experiencing adverse events obtained from hierarchical network meta-analysis incorporating dose constraints

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Odds Ratio† (95%CrI)	Rank worst (95%CrI)	p(Worst)
	Darifenacin ER 30mg q.d	[38]	2	151	6.24 (3.51,11.64)	2 (1,7)	0.40
	Imidafenacin 0.25mg b.i.d	[37]	1	76	4.37 (2.16,8.77)	4 (1,25)	0.10
	Solifenacin ER 20mg q.d	[33]	1	37	3.57 (1.97,7.99)	6 (1,30)	0.05
	Tarafenacin 0.4mg q.d	[82]	1	81	3.52 (1.59,7.92)	6 (1,42)	0.05
	Terodiline 25mg b.i.d	[28]	1	46	3.11 (1.15,8.78)	9 (1,59)	0.05
	Propiverine IR 45mg t.i.d	[118]	1	49	3.02 (1.87,8.1)	9 (1,32)	0.03
	Propiverine ER 60mg q.d	[119]	1	43	3.03 (1.99,8.03)	9 (1,28)	0.03
	Duloxetine 60mg b.i.d	[66]	1	153	3.05 (1.59,5.94)	9 (2,43)	0.01
	Oxybutynin IR 5mg t.i.d	[7]	3	257	2.9 (2.24,4.25)	10 (4,21)	0.00
	Cizolirtine citrate 400mg b.i.d	[57]	2	81	2.72 (1.35,5.55)	13 (2,50)	0.01
	Oxbutynin patch 73.5mg	[15]	1	572	2.65 (2,3.68)	14 (5,29)	0.00
	Oxybutynin vaginal ring 6mg q.d	[17]	1	147	2.64 (1.86,3.66)	14 (5,33)	0.00
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	2.6 (1.59,4.3)	15 (4,43)	0.00
	Oxybutynin IR 5mg b.i.d	[18]	4	415	2.68 (2.1,3.64)	15 (6,27)	0.00
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	2	1018	2.55 (1.79,3.51)	16 (6,35)	0.00
	Darifenacin ER 15mg q.d.	[40]	2	321	2.5 (1.73,3.7)	17 (5,39)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	2.46 (1.54,3.31)	18 (7,45)	0.00
	Fesoterodine ER 8mg q.d	[26]	4	1187	2.4 (1.87,3.08)	20 (7,34)	0.00
	Oxybutynin vaginal ring 4mg q.d	[16]	1	143	2.37 (1.47,3.13)	21 (10,47)	0.00
	Fesoterodine IR 12mg b.i.d	[122]	1	38	2.27 (1.46,4.44)	22 (4,46)	0.00
	Propiverine IR 30mg b.i.d	[117]	1	47	2.29 (1.43,4.98)	22 (4,48)	0.00
	Propiverine ER 30mg q.d	[42]	1	391	2.3 (1.8,3.15)	22 (9,35)	0.00
	Solifenacin ER 5 - 15mg q.d	[34]	1	385	2.2 (1.31,3.78)	24 (5,52)	0.00
	Propiverine IR 15mg t.i.d	[116]	1	149	2.2 (1.42,3.5)	24 (7,49)	0.00
	Propiverine ER 20mg q.d	[41]	6	1458	2.11 (1.68,2.66)	27 (16,40)	0.00
	Solifenacin ER 10mg q.d	[30]	2	375	2.02 (1.47,2.84)	29 (11,46)	0.00
	Tolterodine IR 4mg b.i.d	[140]	1	16	1.99 (1.21,6.84)	30 (2,55)	0.02
	OnaBoNTA 300u trigone sparing	[76]	1	55	1.92 (1.1,3.59)	31 (6,57)	0.00
	Darifenacin ER 7.5mg q.d	[39]	2	377	1.94 (1.29,2.88)	31 (11,53)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	25	1.94 (0.8,4.28)	32 (5,72)	0.00
	Trospium chloride IR 15mg t.i.d	[46]	1	22	1.88 (1.32,3.53)	33 (6,51)	0.00
	Fesoterodine IR 8mg b.id	[121]	1	47	1.91 (1.13,3.04)	33 (10,59)	0.00
	Trospium chloride IR 45mg t.i.d	[47]	1	828	1.85 (1.26,3.09)	34 (9,54)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1047	1.87 (1.52,2.29)	34 (21,46)	0.00
	Propiverine IR 15mg b.i.d	[43]	3	541	1.85 (1.35,2.41)	34 (21,51)	0.00
	Solifenacin ER 5 - 10mg q.d	[31]	3	1160	1.74 (1.27,2.41)	38 (20,53)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	3	186	1.71 (1,2.86)	38 (11,63)	0.00
	Fesoterodine ER 4mg q.d	[25]	5	1368	1.69 (1.38,2.08)	40 (28,51)	0.00
	Trospium chloride ER 60mg q.d	[44]	2	578	1.67 (1.19,2.28)	40 (22,57)	0.00
	Solifenacin ER 5 - 10mg q.d + BT	[77]	1	304	1.66 (0.88,3.08)	41 (9,70)	0.00
	ZD0947IL 25mg/day	[58]	1	90	1.61 (0.77,3.42)	42 (7,73)	0.00
	Fesoterodine IR 4mg b.i.d	[120]	1	43	1.62 (0.83,2.43)	42 (20,71)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	4	602	1.62 (1.16,2.28)	42 (22,58)	0.00
	Trospium chloride IR 20mg b.i.d	[45]	2	596	1.64 (1.14,2.26)	42 (24,59)	0.00
	OnaBoNTA 150u trigone sparing	[75]	1	50	1.6 (0.92,2.66)	43 (15,67)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	79	1.48 (0.68,3.17)	46 (8,74)	0.00
	OnaBoNT-A 100u trigone sparing	[72]	3	186	1.48 (0.82,2.51)	47 (20,71)	0.00
	Pregabalin 150mg b.i.d + tolterodine ER 4mg q.d	[102]	1	102	1.39 (0.79,2.48)	49 (18,72)	0.00
	Pregabalin 150mg b.i.d	[62]	1	105	1.37 (0.73,2.59)	50 (16,74)	0.00
	Tolterodine IR 2mg b.i.d	[5]	13	1293	1.36 (1.07,1.77)	50 (37,63)	0.00
	Netupitant 200mg q.d	[112]	1	61	1.29 (0.54,3.29)	53 (7,75)	0.00
	OnaBoNTA 50u trigone sparing	[74]	1	56	1.3 (0.65,2.3)	53 (25,75)	0.00
	Solifenacin ER 5mg q.d	[29]	6	1133	1.23 (0.88,1.67)	55 (40,70)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	1.23 (0.69,1.9)	55 (33,75)	0.00
	Tolterodine ER 4mg q.d	[4]	19	4729	1.23 (1.06,1.42)	55 (46,63)	0.00

TABLE 7.7: Estimated posterior odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from hierarchical network meta-analysis incorporating dose constraints (cont.)

	Mirabegron ER 200mg q.d	[53]	1	167	1.23 (0.98,1.76)	55 (37,64)	0.00
	Solifenacin/trospium + placebo injection	[100]	1	127	1.2 (0.5,2.9)	56 (11,77)	0.00
	ONO-8539 300mg b.i.d	[61]	1	88	1.17 (0.64,2.15)	57 (25,74)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	65	1.14 (0.84,1.72)	58 (38,71)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	2	275	1.16 (0.66,2.04)	58 (29,75)	0.00
	Mirabegron ER 100mg q.d	[52]	4	1692	1.13 (0.94,1.38)	59 (49,67)	0.00
	Tolterodine IR 1mg b.i.d	[6]	4	297	1.14 (0.79,1.51)	59 (46,72)	0.00
	Lipo-BoNTA 200U	[138]	1	29	1.07 (0.01,551.1)	62 (1,81)	0.23
	Mirabegron ER 50mg q.d	[51]	9	3889	1.06 (0.89,1.26)	63 (54,70)	0.00
	Placebo	[1]	67	12078	NA	66 (58,72)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	1 (0.6,1.31)	66 (52,76)	0.00
	Mirabegron ER 25mg q.d	[50]	3	811	1 (0.8,1.2)	66 (57,73)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	105	0.99 (0.54,1.74)	66 (39,77)	0.00
	Solabegron IR 125mg b.i.d	[55]	1	85	0.97 (0.51,1.89)	67 (33,76)	0.00
	Netupitant 100mg q.d	[111]	1	61	0.96 (0.41,2.33)	67 (22,77)	0.00
	ONO-8539 100mg b.i.d	[60]	1	87	0.94 (0.58,1.58)	68 (43,76)	0.00
	Tolterodine IR 0.5mg b.i.d	[141]	1	21	0.88 (0.38,1.34)	70 (52,78)	0.00
	ONO-8539 30mg b.i.d	[59]	1	88	0.8 (0.48,1.4)	72 (50,78)	0.00
	Percutaneous tibial nerve stimulation	[83]	4	188	0.74 (0.28,1.86)	73 (34,78)	0.00
	Solifenacin ER 2.5mg q.d	[32]	1	41	0.76 (0.29,1.35)	73 (51,79)	0.00
	Solabegron IR 50mg b.i.d	[54]	1	88	0.78 (0.4,1.5)	73 (46,78)	0.00
	Netupitant 50mg q.d	[110]	1	62	0.73 (0.28,1.87)	74 (35,79)	0.00
	Electrostimulation	[80]	2	48	0.35 (0.02,2.16)	78 (25,81)	0.00
	Sham therapy	[3]	2	141	0.13 (0.01,0.85)	79 (72,81)	0.00
	Control	[2]	1	75	0.09 (0,1.56)	80 (44,81)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	2	103	0.04 (0,0.28)	81 (78,81)	0.00

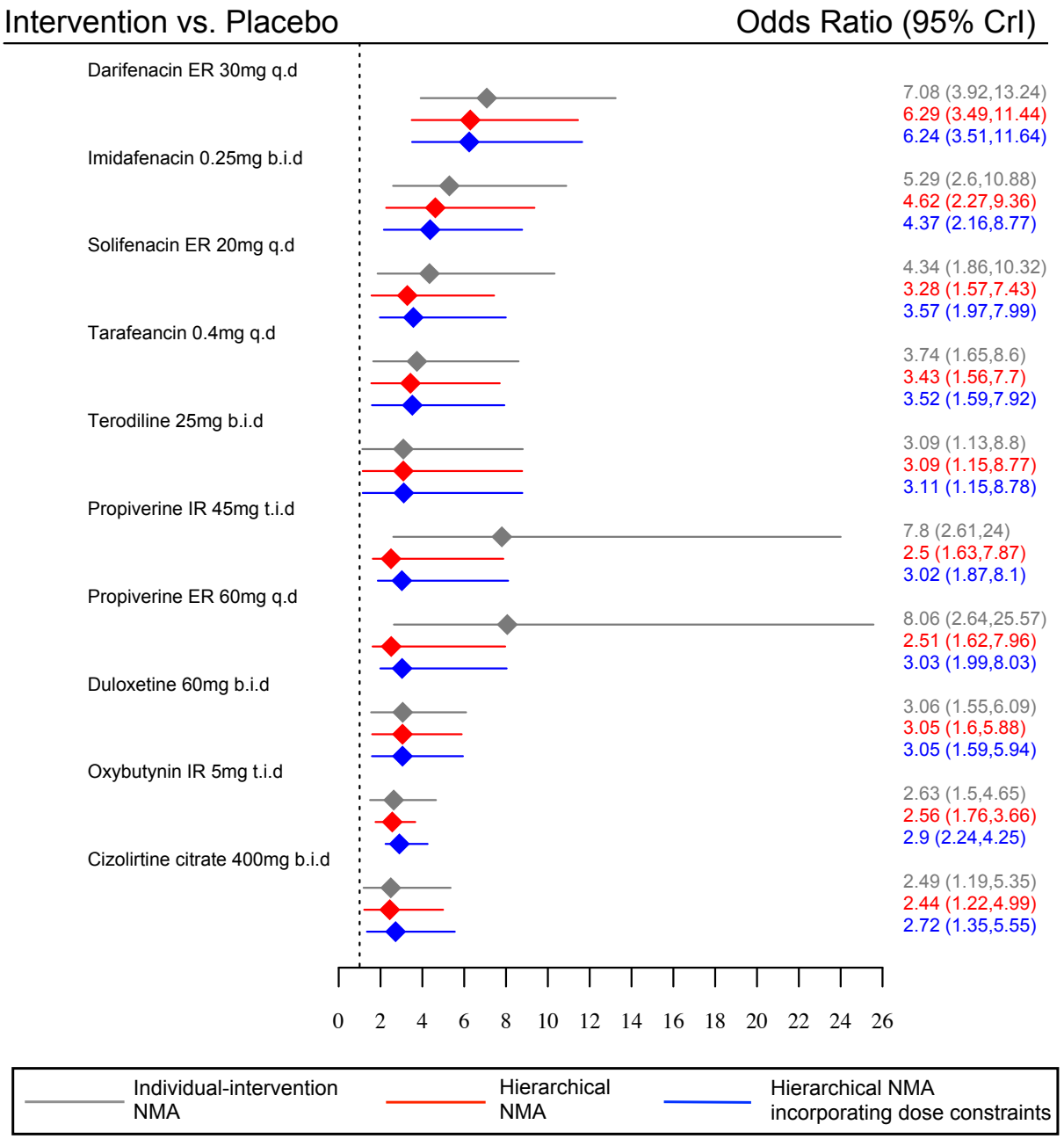
† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment has the highest incidence of discontinuations due to a lack of efficacy overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Worst) denotes the probability that the intervention in question has the highest incidence of discontinuations. The probability worst is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

Figure 7.6 illustrates the point estimates and 95% credible intervals for the estimated odds ratios obtained from individual-intervention, and hierarchical NMAs. For illustration purposes, results for the top 10 interventions are presented. It was apparent that there was a substantial benefit in adopting a hierarchical approach with regard to the reduction in posterior uncertainty, as shown through the narrower credible intervals. This was particularly evident for propiverine IR 45mg t.i.d, and propiverine ER 60mg q.d, for which there was only 1 study evaluating each of these specific formulations. Consequently, results obtained from individual-intervention NMA had a considerable amount of uncertainty in the estimated effect. However, there were many studies evaluating different formulations of propiverine, and borrowing strength between alternative formulations of propiverine interventions considerably increased the precision in the effect estimates. Furthermore, incorporating dose-response constraints further increased the precision in the treatment effect estimates.

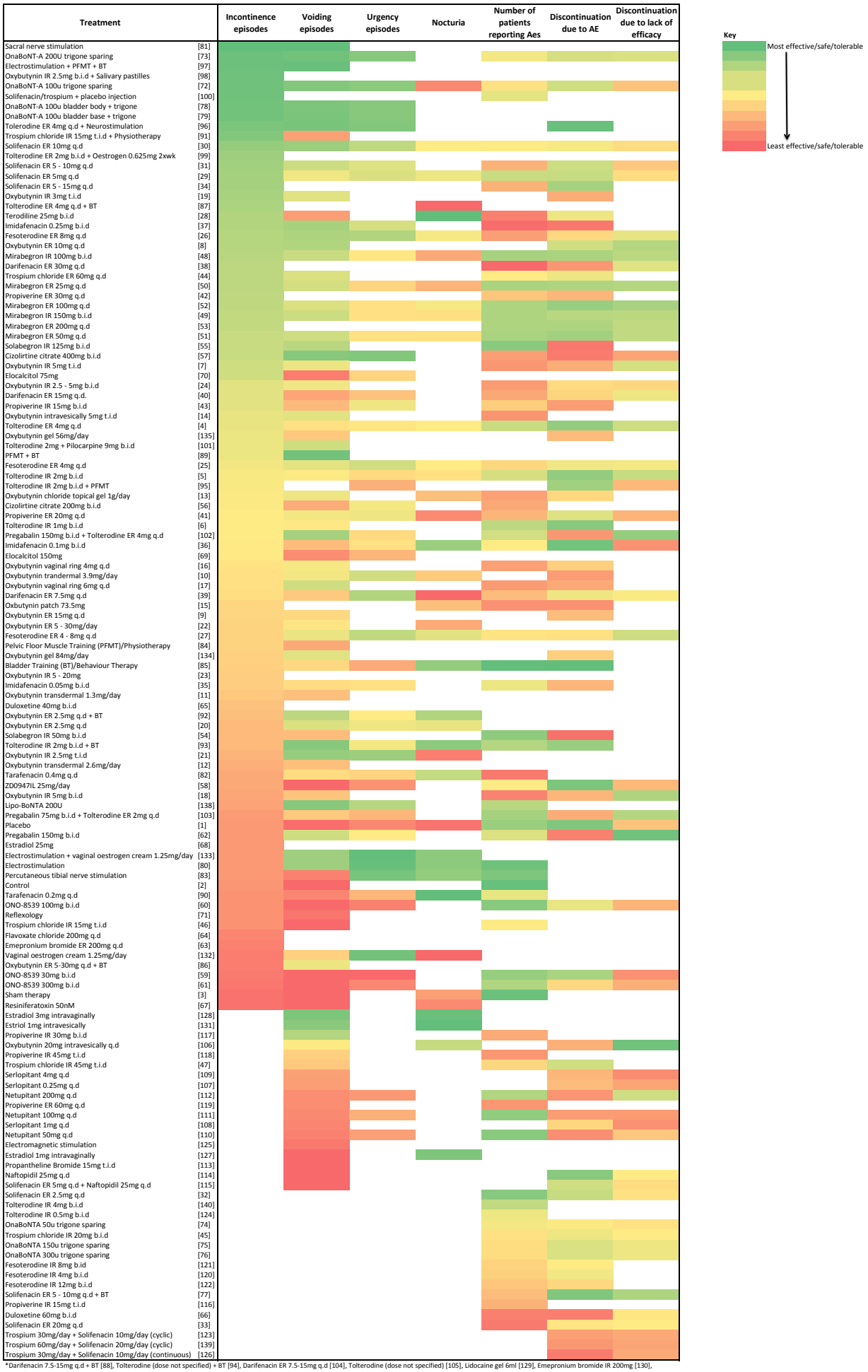
FIGURE 7.6: Comparison of the estimated odds ratios (and 95% credible intervals) in change from baseline in incontinence episodes relative to placebo between individual-intervention, hierarchical, and dose constraint models



7.4.3 Efficacy, safety and tolerability treatment profiles

Figure 7.7 illustrates the efficacy, safety and tolerability profiles of each treatment in ranked order across each of the outcome measures. For each outcome, the best fitting models according to DIC statistics were selected. For incontinence, voiding, and urgency episodes, together with, number of patients experiencing adverse events and discontinuations due to adverse events, hierarchical models were of a substantially better fit to the data than that of individual-intervention NMA. For nocturia episodes and discontinuations due to a lack of efficacy, individual-intervention NMA were more appropriate model choices. Treatments were ranked in order of efficacy for each outcome from left to right. Dark green indicates better performing interventions (i.e. the most effective, tolerable or safest intervention) and red indicates the least effective, safe, or tolerable interventions. Where blank, data were not available i.e. the interventions were not analysed or disconnected from the networks of evidence. It is clear that from a decision makers perspective, there is very little difference in the overall treatment recommendation for intervention effects obtained from hierarchical models or individual-intervention NMA. In agreement with the efficacy, safety and tolerability profiles obtained from individual-intervention NMA described in Section 6.4.3 in Chapter 6, SNS appeared to be the most effective intervention for reducing both urinary incontinence and voiding episodes, but data were not available for all other outcomes. OnaBoNT-A 200U trigone sparing had the best efficacy, safety and tolerability treatment profile across all evaluable outcomes, and mirabegron appeared to have equal efficacy to all antimuscarinic drug therapies but with improved safety and tolerability profiles. Though hierarchical NMAs broadly agree with the treatment rankings from individual-intervention NMA, the increase in precision in estimated treatment effects and relative rankings could have an important impact on future cost-effectiveness analyses and decision making.

FIGURE 7.7: Heatmap of treatment profiles across all outcomes



*Darifenacin 7.5-15mg q.d + BT [88], Tolterodine (dose not specified) + BT [94], Darifenacin ER 7.5-15mg q.d [104], Tolterodine (dose not specified) [105], Lidocaine gel 6ml [129], Emepronium bromide IR 200mg [130], Sacral nerve stimulation + Tolterodine ER 2mg q.d [136], Tolterodine ER 2mg q.d [137] were disconnected from the network

7.4.4 Model assessment

7.4.4.1 Assessing inconsistency between direct and indirect information

Inconsistencies between direct and indirect information are given in Tables D.4 and D.5 of Appendix D. The results presented here were based on the best fitting models for each outcome - for urinary incontinence episodes that is results obtained from a hierarchical NMA, and for adverse events that is results obtained from hierarchical NMA incorporating dose constraints. There appeared to be little evidence of inconsistencies between direct and indirect estimates obtained from hierarchical NMAs as assessed by methods of node-splitting. Node-splitting analyses were further described in Section 3.4.5.3. For urinary incontinence episodes, fesoterodine ER 4mg versus fesoterodine ER 8mg q.d appeared to have conflicting direct and indirect information. For number of patients experiencing adverse events, treatment comparisons between tolterodine IR 2mg b.i.d versus percutaneous tibial nerve stimulation had conflicting direct and indirect information. In terms of the number of inconsistent vertices in the networks of evidence, generally, results obtained from hierarchical NMAs had fewer inconsistencies between direct and indirect information compared to that of individual-intervention NMA presented in Chapter 6. This was particularly apparent for urinary incontinence episodes. For individual-intervention NMA there appeared to be 2 inconsistent vertices - tolterodine ER 4mg q.d versus solifenacin ER 5-10mg q.d, and fesoterodine ER 4mg q.d versus fesoterodine ER 8mg q.d. Use of hierarchical NMAs found only 1 inconsistent vertex between fesoterodine ER 4mg q.d versus fesoterodine ER 8mg q.d. However, removing potentially inconsistent studies from the NMAs had no impact on the overall result.

7.4.4.2 Convergence diagnostics

Convergence diagnostic plots for basic parameters, d_{1j}, the pooled effect estimate of intervention j relative to placebo, and the between-study standard deviations, τ,

are given in Section D.7 of Appendix D. For illustrative purposes, a small sample of diagnostic plots were presented for each outcome. Interpretation of diagnostic plots, and their use in detecting non-convergence, are described in Section 3.2.2. Overall, hierarchical NMAs did not appear to have difficulties with non-convergence. Figures D.1 and D.2 illustrates the Brooks- Gelman-Rubin plots for a random selection of parameters for efficacy and safety outcomes, respectively. The ratio of between and within chain variability, R, appeared to converge to 1, and both the between-chain, B, and within chain, G, variability appeared to reach stability. Autocorrelation plots given in Figures D.3 and D.4 suggested that there were adequate mixing of the chains with all plots tending to zero with increased lag. Figures D.5 and D.6 presents the history and trace plots which did not appear to have any systematic trends or vast differences between multiple MCMC chains with very different starting values. Density plots are given in Figures D.7 and D.8, and all parameters of interest appeared to reflect the shape of a normal distribution. Overall, the diagnostic plots appeared to suggest that there was no evidence of non-convergence or inadequate mixing of the MCMC chains with very different starting values.

7.4.4.3 Sensitivity analysis

Sensitivity analyses assessing the choice of prior distributions on τ and σ for both efficacy and safety outcomes are given in Tables D.6 and D.7, and D.8 and D.9 of Appendix D, respectively. The broad clinical decision did not change with different choices of prior distributions, and the estimated treatment effects were comparable between models. This finding suggests that the results described in this chapter are sufficiently robust to the choice of prior distributions, and decision makers may be confident in the rigour of the relative effect estimates obtained from hierarchical NMAs.

7.5 Discussion

In this chapter, hierarchical NMAs have proven to be a useful methodology that can be applied to clinical areas with which there are many interventions of interest and the evidence base is somewhat limited, both in terms of the number of trials and the number of direct treatment comparisons (Warren et al., 2014). With the development of MCMC simulation techniques within the freely available WinBUGS software, hierarchical NMAs are not only computationally feasible but also widely applicable to other clinical settings.

Characteristically, NMAs performed on large networks of treatment comparisons with relatively few trials frequently evaluate interventions using an individual-intervention NMA, thereby presenting extremely uncertain treatment effect estimates. Alternatively, and in the case of the OAB literature described in Section 2.6, NMAs will focus on analysing a specific set, or class of interventions. Both of these approaches can make it difficult to infer the most efficacious intervention overall, which in turn has implications for health policy decision making. As shown in Chapter 6, undertaking an individual-intervention NMA with a limited evidence base can produce considerable uncertainty in the treatment effect estimates, and thus any inferences regarding interventions effectiveness will remain conservative. Reducing the network by collapsing interventions in to their respective treatment classes will severely hinder the ability to specifically identify the most efficacious individual treatment overall. For example, the class-based NMA for adverse events identified terodiline as the most hazardous class of interventions with a probability of being the worst of 26%, though it was unclear which specific formulation and dose was the most harmful overall.

Use of the term "hierarchical NMA" is intermittently used to describe what is commonly known as a "random effects NMA" with variance components at two levels in the model - one at the within-study level and one at the between-study (within

intervention) level. In this chapter, a third level in the model was added, accounting for additional variance components between interventions within classes of interventions. Adding an additional level to the model changes the assumption of exchangeability and, consequently, the degree of shrinkage (Spiegelhalter et al., 2004). For this reason, there was a notable change in the estimated median treatment effects and their associated precision in a hierarchical NMA compared with that of the individual-intervention NMA described in Chapter 6. In comparison to the individual-intervention NMA, use of hierarchical models as described in this chapter have several advantages. Principally, there was a substantial increase in the precision surrounding the effect estimates, and this was particularly apparent for the interventions for which there were few trials and a limited number of direct treatment comparisons between other active interventions (Warren et al., 2014). In addition, hierarchical models maintain the interpretability of the effect estimates at an individual intervention level which is beneficial for decision making.

Nevertheless, the hierarchical models made a fundamental assumption that the intervention effects, within classes of interventions, were exchangeable, and a *judgement* of appropriateness of such an assumption has to be made (Spiegelhalter et al., 2004). If this assumption does not hold for every class of interventions, use of a hierarchical model could lead to inappropriate results; thus, it is important for researchers to classify treatments into clinically plausible classes. Of course, interventions do not have to be grouped in to classes of interventions if there is no reason to do so. A further limitation of the hierarchical model is the subjective classification of interventions when there is potential treatment overlap. For example, in the case of OAB, trospium IR 15mg t.i.d and physiotherapy as a combination therapy will overlap with both trospium and physiotherapy classes. The combination of interventions individually estimated in the NMA could be modelled as the sum of individual components, with the potential to incorporate a synergistic or subadditive interaction between the interventions (Welton et al., 2009). Furthermore, the number of interventions and trials within each class can vary substantially. For classes with which there were few interventions and a small

evidence base, the estimates will remain fairly uncertain. In situations such as these, the impact of the choice of prior distributions especially on the variance parameters could be important, and use of extensive sensitivity analyses would be crucial (Spiegelhalter et al., 2004; Lambert et al., 2005).

A further consideration is the ordering of the hierarchical structure. In the OAB example, the interventions naturally formed a biological, and clinically plausible hierarchy, whereby dose was nested within formulation, which was further nested within intervention (see Figure 7.1). If additional information was available with regard to exchangeable treatment effects across formulations of interventions, for example, if oxybutynin ER 15mg q.d was considered to have a similar efficacy to oxybutynin IR 5mg t.i.d, then formulation could be considered to be nested within dose. In this particular example, a clinically meaningful similarity between different formulations was not clear and therefore an alternative ordering of the hierarchy was not appropriate. Walsh and Mengersen (2012) explored the impact of the ordering of hierarchies in Bayesian hierarchical models, and found that different hierarchical structures can yield different clinical conclusions, especially in terms of the uncertainty in treatment effect estimates. Therefore, care must be taken in defining appropriate, and clinically meaningful hierarchies. In the absence of a clear reason to choose one hierarchical ordering over another, the user may wish to consider Bayesian model averaging (Hoeting et al., 1999).

As previously highlighted in Chapter 6, a limitation of the analyses specifically related to the OAB example was that many of the original trials did not evaluate all of the cardinal symptoms (urinary incontinence, voiding frequency, and urgency). Consequently, different interventions were evaluated for different outcomes, which makes decision making for syndromic conditions difficult. This was particularly apparent across the treatment profiles, in which very few interventions had a complete efficacy, safety and tolerability profile. To help ameliorate the potential effect of outcome reporting bias in trials that fail to report all of the outcomes of interest, hierarchical models can be further extended to incorporate

multiple outcomes (Kirkham et al., 2012; Hong et al., 2013). This approach estimates a correlation between outcomes in order to estimate a value for the missing data points, conditional on the reported outcome measures and the model (Nam et al., 2003), which is described in more detail in Chapter 8.

Chapter 6 explored the use of network meta-regression models to adjust for potential treatment effect modifiers in individual-intervention NMA. In these analyses, both exchangeable, and common, regression coefficients were explored. The former assumes that the treatment by covariate interactions were different for each active treatment but similar across all interventions, and the latter assumes that all treatment by covariate interactions were identical (see Section 6.3.3). By fitting these models, age was found to be an important covariate for voiding, and a common treatment by age interaction appeared to be the most appropriate regression coefficient. In this chapter, a further hierarchical approach was explored for voiding with which regression coefficients were assumed to be different but exchangeable between treatments within classes. In this instance, an exchangeable within-class regression coefficient did not improve model fit, and thus the results presented in Table D.1 of Appendix D were based on a hierarchical NMA model incorporating a common regression coefficient. However, in situations with which the class by covariate interactions differ in an important way e.g. if a class of interventions perform better in older populations, an exchangeable within-class regression coefficient may be of particular use.

In this example, use of hierarchical NMAs did not change the overall clinical decision in terms of the most effective intervention for reducing urinary incontinence episodes. For example, SNS appeared to be the most effective intervention for both individual-intervention and hierarchical NMAs with comparable treatment effect estimates. However, for interventions that belong to a broader class of interventions such as onaBoNT-A 200U trigone sparing, the treatment effect estimates obtained from hierarchical NMAs (-2.08; 95%CrI: -2.86,-1.45) were similar to that

of the individual-intervention NMA (-2.3; 95%CrI:-3.16,-1.42) but with a 23% increase in precision. Generally, extending the hierarchical model to incorporate ordering constraints further increased precision. For example, for onaBoNT-A 200U trigone sparing, there was a 15% increase in precision from hierarchical NMAs incorporating dose constraints (-2.19; 95%CrI: -2.91,-1.68) compared to hierarchical NMAs without dose constraints (-2.08; 95%CrI: -2.86,-1.45). For safety outcomes, there was a substantial gain in precision from using hierarchical NMAs, and consequently there was greater confidence in the relative treatment rankings. Similarly to efficacy outcomes, assuming that larger doses of an intervention had a greater or equal effect to that of lower doses did not alter the effect estimates to any noticeable extent. This approach did, however, reduce the uncertainty in the effect estimates, estimated treatment rankings, and the probability that each of the interventions were the most hazardous overall.

7.6 Chapter summary

In this chapter, use of hierarchical NMAs incorporating dose response constraints have been described and demonstrated in the context of OAB. It has been shown that application of hierarchical models can lead to increased precision in the estimated treatment effects, without hindering the interpretability of individual interventions. In this example, borrowing strength across treatments within the classes of OAB treatments led to reduced uncertainty in the individual intervention effects. However, the point estimates were broadly comparable with results obtained from individual-intervention NMAs. Therefore, the overall clinical decision did not change with the use of hierarchical NMAs but the increase in precision can have a substantial impact in the confidence of future cost-effectiveness analyses and decision making.

A recurring limitation of the analyses presented here, and in the previous chapter, was that many of the original trials did not evaluate all of the cardinal symptoms

of OAB. As a result, decision makers will be required to make inferences regarding the best allocation of interventions based on incomplete data. Chapter 8 aims to ameliorate the impact of potential outcome reporting bias by extending the methodology described in this chapter to incorporate a multivariate approach.

Chapter 8

Multivariate Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder

8.1 Chapter overview

Network meta-analysis synthesises data from a number of clinical trials of several treatments in order to assess the comparative efficacy of healthcare interventions in similar patient populations. In situations with which clinical trial data is selectively reported i.e. data are missing for one or more outcomes, synthesising such data can lead to potentially biased estimates (Kirkham et al., 2012) and disconnected networks of evidence, which can have severe implications for decision-making. In Chapters 6 and 7, treatment recommendations varied for each of the OAB symptoms, which may have been a result of the available data for each of the interventions, and outcomes of interest. In order to manage the complete symptom syndrome, patients may be exposed to unnecessary poly-pharmacy which would increase prescribing, healthcare costs, adverse reactions, and potential interactions

between interventions (Haider et al., 2009). To maximise healthcare resources and to minimise poly-pharmacy, it is imperative that all available evidence relevant to the decision problem is taken into account when summarising efficacy data. This chapter extends the methodology from the multiple outcome hierarchical network meta-analyses presented in Chapter 7 to incorporate a simultaneous multiple outcome setting.

Multivariate NMA (MVNMA) models utilise the correlation between outcomes in order to predict a distribution for missing data points (Nam et al., 2003). In this chapter, MVNMAs are developed and applied to the three cardinal symptoms of OAB: urinary incontinence, voiding frequency, and urgency (see Chapter 2 for more details). This chapter begins by briefly introducing existing MVNMA models that have recently been proposed by Efthimiou et al. (2014) and Achana et al. (2014). In the following section, the general MVNMA framework described by Achana et al. (2014) is extended to incorporate the correlation between change from baseline, baseline, and follow-up values, as described in Chapter 6, and further developed to incorporate a hierarchical approach as described in Chapter 7. The advantages of applying MVNMA in the context of HTA are described and demonstrated. Results obtained from the application of these methods to the motivating example in OAB syndrome are presented, and the chapter concludes with a discussion and final conclusions.

8.2 Introduction

In a HTA setting, one important step in the evaluation of medical interventions is the assessment of cost-effectiveness. In economic evaluations, all outcomes associated with the decision question are included, and thus, the dependence between correlated outcomes, as well as a lack of reporting of outcomes, may have a substantial impact on the ability to appropriately estimate cost-effectiveness (Sterne et al., 2009), especially in terms of uncertainty. This is particularly true

for the evaluation of OAB syndrome, with which many of the key symptoms are often under-reported. For example, in the current literature urgency is defined as "the cardinal symptom" of OAB (Cardozo et al., 2009), however, as previously highlighted in Chapter 5 there are far fewer studies evaluating interventions for urgency. This poses several limitations for decision makers; most notably, it is difficult to estimate both clinical and cost-effectiveness of interventions across the entire symptom syndrome.

Multivariate NMA has the ability to simultaneously model all outcomes of interest, and estimate a correlation between outcomes. Consequently, for the studies that fail to report certain outcomes, it is possible to obtain a predictive value for missing data using the estimated correlation. This methodology will not only increase the evidence base for decision making, but it will also limit the potential for outcome reporting bias in the original trials (Kirkham et al., 2012), and consequently increase the precision in the treatment effect estimates (Achana et al., 2014).

There are two types of correlations that need to be incorporated in the estimation of treatment effect estimates for multiple outcomes - one at the between-study level, and another at the within-study level. Between-study correlations occur due to differences in the sources of variability between-studies that have previously been encountered and discussed in Chapters 3, 5 and 6 for univariate random effects pairwise and network meta-analyses. It is these correlations that provide an indication of how the true study-specific treatment effects vary across trials, and are quantified using the between-study standard deviation. Between-study correlations occur in situations with which studies have a different distribution of potential treatment effect modifiers, such as differences in patient characteristics, trial designs, and baseline patient severity. Within-study correlations occur between outcomes within a trial, and are a consequence of differences in patient-level characteristics. These correlations are indicative of the association between multiple outcomes within a study. Within-study correlations are often difficult to

estimate and they are seldom reported in clinical trials. Thus, estimation is often required using individual participant data (IPD) (Riley, 2009) or elicited from expert opinion (Efthimiou et al., 2014).

The ability to simultaneously model healthcare outcomes in a multivariate analysis is an appealing feature in many HTA settings, as interest frequently lies in multiple and correlated outcome measures (Berkey et al., 1998). In the last 20 years, evidence synthesis methods have witnessed a rapid increase in methodological developments and applications of multivariate analyses to assess interventions with two or more outcomes of interest (Daniels and Hughes, 1997; Berkey et al., 1998; Van Houwelingen et al., 2002; Nam et al., 2003; Arends et al., 2003; Riley et al., 2007; Jackson et al., 2011; Wei and Higgins, 2013; Bujkiewicz et al., 2013; Schwarzer et al., 2015). It is often desirable to account for the correlation between outcomes in a meta-analysis framework as this has the ability to borrow strength between all reported outcomes and studies in order to inform treatment effect estimates (Bujkiewicz et al., 2013). Such an approach is commonly referred to as multivariate meta-analysis.

In more recent developments, multivariate meta-analyses have been extended to incorporate multiple treatment comparisons (Ades et al., 2010; Efthimiou et al., 2014; Achana et al., 2014). Ades et al. (2010) simultaneously modelled mutually exclusive, competing risk outcomes, using a multinomial likelihood whereby the within-study correlations were accounted for but the between-study correlations were assumed to be zero. Efthimiou et al. (2014) proposed a MVNMA model that accounts for both the between- and within-study correlations of binary outcomes. Specifically, Efthimiou et al. (2014) incorporated the within-study correlations at the study-specific treatment difference level, which can be problematic when incorporating multi-arm trials. Achana et al. (2014) extended this methodology using a more natural modelling approach whereby the within-study correlations were incorporated at the treatment-arm level. Using this approach, Achana et al.

(2014) considered the treatment arms to be independent as a consequence of randomisation, which greatly eases computation of the likelihood for multi-arm trials. Achana et al. (2014) developed this methodology to borrow information across outcomes in order to predict an estimate for missing data. This methodology allows disconnected interventions to be incorporated in to the analyses if they belong to a connected network for one or more additional outcomes; thereby, allowing all interventions to be evaluated across all outcome measures. For ease of including multi-arm trials, and the desire to borrow strength between outcome measures for missing data, the methods described in this chapter are developed from the general framework described by Achana et al. (2014). This chapter extends the current methodology to account for correlations between baseline and follow-up times, and further incorporates a hierarchical structure (see Chapter 7 for further details). This approach makes use of all available data, and has the potential to increase the precision in treatment effect estimates whilst allowing for the comparison of all interventions across all outcome measures, in order to aid decision making. Motivated by the example in OAB, a trivariate hierarchical NMA is described in this chapter, but the methods can be extended to the multivariate case for any number of outcomes.

8.3 Methods

8.3.1 Multivariate network meta-analysis

Following the random effects multivariate NMA described by Achana et al. (2014), let $\boldsymbol{Y_{ij}} = (y_{ij1}, y_{ij2}, y_{ij3})$, be the observed vector of effects for intervention j of the i^{th} study ($i = 1, 2, ..., n_s$) for each of the outcomes of interest (1 = urinary incontinence, 2 = voiding frequency, 3 = urgency episodes). Let $\boldsymbol{Y_{ij}}$ follow a

multivariate normal (MVN) distribution such that:

$$\begin{aligned} \boldsymbol{Y_{ij}} &\sim \text{MVN}(\boldsymbol{\theta_{ij}}, \boldsymbol{S_{ij}}) \\ \boldsymbol{\theta_{ij}} &= \boldsymbol{\mu_i} + \boldsymbol{\delta_{i,bk}} I_{\{j=k\}} \\ \text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{8.1}$$

where $\boldsymbol{\theta_{ij}}$ is a vector of true treatment effects, $\boldsymbol{S_{ij}}$ is the treatment-specific within-study covariance matrix assumed known, $\boldsymbol{\mu_i}$ is a vector of baseline effects in study i with baseline treatment b, and $\boldsymbol{\delta_{i,bk}}$ is a vector of treatment-specific effects in the k^{th} arm relative to the baseline treatment in arm 1 of study i. The elements of $\boldsymbol{S_{ij}}$ are expressed as:

$$\boldsymbol{S_{ij}} = \begin{pmatrix} se^2_{ij(1)} & \rho w_{(12)} se_{ij(1)} se_{ij(2)} & \rho w_{(13)} se_{ij(1)} se_{ij(3)} \\ \rho w_{(12)} se_{ij(1)} se_{ij(2)} & se^2_{ij(2)} & \rho w_{(23)} se_{ij(2)} se_{ij(3)} \\ \rho w_{(13)} se_{ij(1)} se_{ij(3)} & \rho w_{(23)} se_{ij(2)} se_{ij(3)} & se^2_{ij(3)} \end{pmatrix} \tag{8.2}$$

where $se_{ij(p)}$ denotes the observed standard errors of intervention j for outcome, $p = 1, 2, 3$, and ρw_{qr} denotes the within-study correlations, for $q = 1, 2$ and $r = 2, 3$, described in more detail in Section 8.3.3. For 2-arm trials, the elements of $\boldsymbol{\delta_{i,bk}}$ are assumed to be drawn from:

$$\begin{pmatrix} \delta_{i,bk(1)} \\ \delta_{i,bk(2)} \\ \delta_{i,bk(3)} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} d_{t_{ib}t_{ik}(1)} = d_{AB(1)} = d_{1B(1)} - d_{1A(1)} \\ d_{t_{ib}t_{ik}(2)} = d_{AB(2)} = d_{1B(2)} - d_{1A(2)} \\ d_{t_{ib}t_{ik}(3)} = d_{AB(3)} = d_{1B(3)} - d_{1A(3)} \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} \right) \tag{8.3}$$

where $d_{AB(p)}$ represents the effect of treatment A relative to B for each outcome, p. $\boldsymbol{\Sigma}$ is the between-study covariance matrix under a homogeneous between-study variance (Lumley, 2002) and correlations assumption (Achana et al., 2014) which is discussed in more detail in Section 8.3.3.

Building on the NMA framework for multi-arm trials, i.e. $k > 2$, described in Section 3.4.5, the elements of $\boldsymbol{\delta_{i,bk}}$ are expressed in terms of the following marginal and conditional distributions:

$$\begin{pmatrix}\delta_{i,b2(1)}\\ \delta_{i,b2(2)}\\ \delta_{i,b2(3)}\end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix}d_{t_{ib}t_{i2}(1)} = d_{AB(1)} = d_{1B(1)} - d_{1A(1)}\\ d_{t_{ib}t_{i2}(2)} = d_{AB(2)} = d_{1B(2)} - d_{1A(2)}\\ d_{t_{ib}t_{i2}(3)} = d_{AB(3)} = d_{1B(3)} - d_{1A(3)}\end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix}\right) \tag{8.4}$$

and for $k = 3 \ldots na_i$, the k^{th} conditional distribution is defined by:

$$\begin{pmatrix}\delta_{i,bk(1)}\\ \delta_{i,bk(2)}\\ \delta_{i,bk(3)}\end{pmatrix} \mid \begin{pmatrix}\begin{pmatrix}\delta_{i,b2(1)}\\ \delta_{i,b2(2)}\\ \delta_{i,b2(3)}\end{pmatrix}\\ \vdots\\ \begin{pmatrix}\delta_{i,b(k-1)(1)}\\ \delta_{i,b(k-1)(2)}\\ \delta_{i,b(k-1)(3)}\end{pmatrix}\end{pmatrix} \sim \text{MVN}$$

$$\left(\begin{pmatrix}\left(d_{t_{ib}t_{ik}(1)} = d_{1,t_{ik}(1)} - d_{1,t_{ib}(1)}\right) + \frac{1}{k-1}\sum_{q=1}^{k-1}\left[\delta_{i,1q(1)} - \left(d_{1,t_{iq}(1)} - d_{1,t_{ib}(1)}\right)\right]\\ \left(d_{t_{ib}t_{ik}(2)} = d_{1,t_{ik}(2)} - d_{1,t_{ib}(2)}\right) + \frac{1}{k-1}\sum_{q=1}^{k-1}\left[\delta_{i,1q(2)} - \left(d_{1,t_{iq}(2)} - d_{1,t_{ib}(2)}\right)\right]\\ \left(d_{t_{ib}t_{ik}(3)} = d_{1,t_{ik}(3)} - d_{1,t_{ib}(3)}\right) + \frac{1}{k-1}\sum_{q=1}^{k-1}\left[\delta_{i,1q(3)} - \left(d_{1,t_{iq}(3)} - d_{1,t_{ib}(3)}\right)\right]\end{pmatrix}, \frac{k}{2(k-1)}\boldsymbol{\Sigma}_{3x3}\right)$$

where

$$\boldsymbol{\Sigma}_{3x3} = \begin{pmatrix}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix} & \frac{1}{2}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix} & \frac{1}{2}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix}\\ \frac{1}{2}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix} & \begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix} & \frac{1}{2}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix}\\ \frac{1}{2}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix} & \frac{1}{2}\begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix} & \begin{pmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ . & . & \sigma_3^2\end{pmatrix}\end{pmatrix}$$

The relative effect of the study-specific reference treatment in arm 1 (the control

arm) relative to itself for outcome p, $\delta_{i,b1(p)}$, is set to 0 and as such the set of conditional univariate distributions begin with the relative effect of the intervention in arm 2 relative to the control arm, $\delta_{i,b2(p)}$. The study-specific treatment comparisons, $\delta_{i,b(k-1)(p)}$, are expressed in terms of the basic parameters of the pooled treatment effects for the intervention in arm k, $d_{1,t_{ik}(p)}$ and the basic parameters of the pooled treatment effects for the intervention in the reference treatment arm, $d_{1,t_{ib}(p)}$, as described in Equation (3.21).

Following Achana et al. (2014), in order to predict treatment effect estimates for missing data for trials that failed to report all outcomes of interest, it was assumed that the pooled effects of intervention j relative to a reference treatment, $d_{1j(p)}$, for outcome, p, can be expressed as a sum of treatment-specific effects, α_j, and outcome-specific effect, γ_p, such that:

$$d_{1j(p)} \sim \text{Normal}(\alpha_j + \gamma_p, \zeta^2) \tag{8.5}$$

The parameter ζ indicates the deviation of treatment effect profiles across outcomes. If ζ was close to zero, this would indicate a high degree of similarity between outcomes. In situations where ζ was particularly large, this would indicate a substantial deviation between treatment effect profiles across outcomes (Achana et al., 2014).

Non-informative prior distributions were specified for μ_{ip}, $\alpha_j \sim \text{Normal}(0, 10^3)$, $\sigma_p \sim \text{Uniform}(0,2)$, and $\zeta \sim \text{Uniform}(0, 2)$, for $p = 1, 2, 3$. The spherical parameterization technique was used to express ρ_{xy}, for $x = 1, 2$ and $y = 2, 3$, to ensure that values lie in the interval (-1,1) which is discussed in more detail in Section 8.3.3.

8.3.2 Multivariate hierarchical network meta-analysis

The MVNMA framework outlined above was extended to incorporate a hierarchical approach as described in the context of univariate NMAs in Chapter 7. This approach utilises the additional similarity between the same interventions with different treatment regimes. For study i evaluating intervention j belonging to class m, $\boldsymbol{Y_{ij_m}} = (y_{ij_m1}\ y_{ij_m2}\ y_{ij_m3})$, denotes the observed vector of effects for each of the outcomes of interest, and expressed in the following way:

$$\begin{aligned} \boldsymbol{Y_{ij_m}} &\sim \text{MVN}(\boldsymbol{\theta_{ij_m}}, \boldsymbol{S_{ij_m}}) \\ \boldsymbol{\theta_{ij_m}} &= \boldsymbol{\mu_i} + \boldsymbol{\delta_{i,bk}} I_{\{j_m=k\}} \\ \text{where } I_{\{u\}} &= \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{8.6}$$

where $\boldsymbol{\mu_i}$ and $\boldsymbol{\delta_{i,bk}}$ have the same interpretation as the model described in Section 8.3.1 above. The parameter $\boldsymbol{\theta_{ij_m}}$, represents a vector of true treatment effects, and $\boldsymbol{S_{ij_m}}$ represents the within-study covariance matrix, for intervention j belonging to a broader class of interventions m, such that:

$$\boldsymbol{S_{ij_m}} = \begin{pmatrix} se^2_{ij_m(1)} & \rho w_{(12)} se_{ij_m(1)} se_{ij_m(2)} & \rho w_{(13)} se_{ij_m(1)} se_{ij_m(3)} \\ \rho w_{(12)} se_{ij_m(1)} se_{ij_m(2)} & se^2_{ij_m(2)} & \rho w_{(23)} se_{ij_m(2)} se_{ij_m(3)} \\ \rho w_{(13)} se_{ij_m(1)} se_{ij_m(3)} & \rho w_{(23)} se_{ij_m(2)} se_{ij_m(3)} & se^2_{ij_m(3)} \end{pmatrix} \tag{8.7}$$

For specific interventions A and B belonging to class c, let the pooled treatment difference, $d_{t_{ib}t_{ik}}$, be denoted as $d_{A_cB_c}$. Then extending the general NMA framework outlined in Section 3.4.5, to incorporate multiple outcomes, the elements of $\boldsymbol{\delta_{i,bk}}$ for 2-arm trials are assumed to be drawn from a multivariate normal distribution:

$$\begin{pmatrix} \delta_{i,bk(1)} \\ \delta_{i,bk(2)} \\ \delta_{i,bk(3)} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} d_{t_{ib}t_{ik}(1)} = d_{A_cB_c(1)} = d_{1_1B_c(1)} - d_{1_1A_c(1)} \\ d_{t_{ib}t_{ik}(2)} = d_{A_cB_c(2)} = d_{1_1B_c(2)} - d_{1_1A_c(2)} \\ d_{t_{ib}t_{ik}(3)} = d_{A_cB_c(3)} = d_{1_1B_c(3)} - d_{1_1A_c(3)} \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} \right) \tag{8.8}$$

For multi-arm trials, $k > 2$, the elements of $\boldsymbol{\delta_{i,bk}}$ are expressed in terms of the following marginal and conditional distributions:

$$\begin{pmatrix} \delta_{i,b2(1)} \\ \delta_{i,b2(2)} \\ \delta_{i,b2(3)} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} d_{t_{ib}t_{i2}(1)} = d_{A_cB_c(1)} = d_{1_1B_c(1)} - d_{1_1A_c(1)} \\ d_{t_{ib}t_{i2}(2)} = d_{A_cB_c(2)} = d_{1_1B_c(2)} - d_{1_1A_c(2)} \\ d_{t_{ib}t_{i2}(3)} = d_{A_cB_c(3)} = d_{1_1B_c(3)} - d_{1_1A_c(3)} \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} \right) \tag{8.9}$$

and for $k = 3 \ldots na_i$, the k^{th} conditional distribution is defined by:

$$\begin{pmatrix} \delta_{i,bk(1)} \\ \delta_{i,bk(2)} \\ \delta_{i,bk(3)} \end{pmatrix} | \begin{pmatrix} \begin{pmatrix} \delta_{i,b2(1)} \\ \delta_{i,b2(2)} \\ \delta_{i,b2(3)} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \delta_{i,b(k-1)(1)} \\ \delta_{i,b(k-1)(2)} \\ \delta_{i,b(k-1)(3)} \end{pmatrix} \end{pmatrix} \sim \text{MVN}$$

$$\left(\begin{pmatrix} \left(d_{t_{ib}t_{ik}(1)} = d_{1,t_{ik}(1)} - d_{1,t_{ib}(1)}\right) + \frac{1}{k-1}\sum_{q=1}^{k-1}\left[\delta_{i,1q(1)} - \left(d_{1,t_{iq}(1)} - d_{1,t_{ib}(1)}\right)\right] \\ \left(d_{t_{ib}t_{ik}(2)} = d_{1,t_{ik}(2)} - d_{1,t_{ib}(2)}\right) + \frac{1}{k-1}\sum_{q=1}^{k-1}\left[\delta_{i,1q(2)} - \left(d_{1,t_{iq}(2)} - d_{1,t_{ib}(2)}\right)\right] \\ \left(d_{t_{ib}t_{ik}(3)} = d_{1,t_{ik}(3)} - d_{1,t_{ib}(3)}\right) + \frac{1}{k-1}\sum_{q=1}^{k-1}\left[\delta_{i,1q(3)} - \left(d_{1,t_{iq}(3)} - d_{1,t_{ib}(3)}\right)\right] \end{pmatrix}, \frac{k}{2(k-1)}\boldsymbol{\Sigma}_{3x3} \right)$$

where

$$\boldsymbol{\Sigma}_{3x3} = \begin{pmatrix} \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} & \frac{1}{2}\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} & \frac{1}{2}\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} \\ \frac{1}{2}\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} & \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} & \frac{1}{2}\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} \\ \frac{1}{2}\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} & \frac{1}{2}\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} & \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ . & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ . & . & \sigma_3^2 \end{pmatrix} \end{pmatrix}$$

The pooled treatment effect of treatment A_c relative to the pooled treatment effect B_c for outcome p, $d_{A_cB_c(p)}$, is given by:

$$d_{t_{ib}t_{ik}(p)} = d_{A_cB_c(p)} = d_{1_1B_c(p)} - d_{1_1A_c(p)} \tag{8.10}$$

As with all other NMA models described throughout this thesis, the intervention effect of the reference treatment for the entire treatment network, $j = 1_1(p)$, usually a placebo or control intervention, is set to 0 for every outcome p, such that $d_{1_11_1(p)} = 0$ (for further details see Section 3.4.5 of Chapter 3). The basic parameters for relative treatment effects, $d_{1_1j_m(p)}$, of intervention j within class m, relative to the reference treatment, were assumed to follow a normal distribution with mean equal to the treatment-specific effect, α_{j_m}, plus the outcome-specific effect, γ_p, and variance, ζ^2 (Achana et al., 2014):

$$d_{1_1j_m(p)} \sim \text{Normal}(\alpha_{j_m} + \gamma_p, \zeta^2) \tag{8.11}$$

where γ_p and ζ^2 have the same interpretation as the MVNMA model described above.

In order to incorporate the exchangeability between treatment-specific effects, α_{j_m}, within class m, α_{j_m} was assumed to follow a normal distribution with mean equal to the pooled effect estimate for the m^{th} class of interventions, β_m, with class specific between-intervention variance, ν_m^2, (Owen et al., 2015) such that:

$$\alpha_{j_m} \sim \text{Normal}(\beta_m, {\nu_m}^2) \tag{8.12}$$

Non-informative prior distributions were specified for $\beta_m \sim \text{Normal}(0, 10^3)$ and $\nu_m \sim \text{Uniform}(0, 2)$, in addition to the non-informative prior distributions selected for the MVNMA model described in Section 8.3.1. However, a prior distribution for α_j is no longer required.

In all multivariate analyses, prior distributions for the variance parameters (i.e. σ_p, ζ, and ν_m) were restricted to a Uniform(0,2) distribution. For example, a value of 2 for the class-specific between-intervention variance, ν_m, suggests that for a random pair of interventions, the difference in the mean change from baseline could be as large as 2.2 events on average. Uniform(0,2) prior distributions were considered for variance components in multivariate analyses in order to aid computation of the variance-covariance matrix. However, there is an argument to suggest that the variance parameters in Bayesian predictive distributions can be decomposed into the sum of several other variance components (Geweke and Amisano, 2014). Thus, in hierarchical NMAs with additional variance components at the class-level, narrower prior distributions for variance parameters may be reasonable.

8.3.3 Estimating between-study and within-study correlations

The between-study variance-covariance matrix, $\boldsymbol{\Sigma}$ requires specification of a prior distribution which ensures that it is non-negative definite (Mavridis and Salanti, 2012). In a multivariate meta-analysis setting, the spherical parameterization technique based on Cholesky decomposition (Watkins, 1991) has previously been adopted (Lu and Ades, 2009; Mavridis and Salanti, 2012; Wei and Higgins, 2013; Achana et al., 2014) using the following decomposition described by Barnard et al. (2000):

$$\boldsymbol{\Sigma} = \boldsymbol{V^{1/2} R V^{1/2}} \tag{8.13}$$

This approach is used to express the between-study variance-covariance matrix, $\boldsymbol{\Sigma}$ in terms of a diagonal matrix of standard deviations, $\boldsymbol{V^{1/2}}$, and positive-definite matrix of correlations, $\boldsymbol{R}$. Here, the elements of $\boldsymbol{V^{1/2}}$ represent the between-study standard deviations of $\boldsymbol{\Sigma}$ and were assigned Uniform(0,2) prior distributions which are further discussed in the context of univariate NMAs for this example in Chapter 3. $\boldsymbol{R}$ represents the correlation matrix where the diagonal elements are set to 1, and the off-diagonal elements contain the set of correlation coefficients. Estimating all between-study correlation parameters can contribute a large number of

parameters to the covariance matrix which can result in computational difficulties (Wei and Higgins, 2013). Thus, reducing the number of correlation parameters is often desirable (Wei and Higgins, 2013). In addition to assuming homogeneous between-study standard deviations across treatment comparisons, which is a common assumption of NMAs (Lumley, 2002), it is also possible to assume that the between-study correlations are equal (Wei and Higgins, 2013; Achana et al., 2014). This approach assumes that if several, independent, multivariate meta-analyses were conducted on the same outcomes, each with a different set of k versus b treatment comparisons, then the between-study correlations, ρ_{xy}, would be the same across the different sets of treatment comparisons:

$$\boldsymbol{R} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ . & 1 & \rho_{23} \\ . & . & 1 \end{bmatrix} \tag{8.14}$$

Using Cholesky decomposition, it has been shown that $\boldsymbol{R}$ can be written in terms of an upper triangular matrix, $\boldsymbol{L}$, (Lu and Ades, 2009; Wei and Higgins, 2013) such that:

$$\boldsymbol{R} = \boldsymbol{L}^{\boldsymbol{T}}\boldsymbol{L} \tag{8.15}$$

With this notation, the spherical parameterization technique can be used to express $\boldsymbol{R}$ in terms of the elements of $\boldsymbol{L}$ using sine and cosine functions (Lu and Ades, 2009), such that:

$$\boldsymbol{L} = \begin{pmatrix} 1 & \cos(\varphi_{12}) & \cos(\varphi_{13}) \\ 0 & \sin(\varphi_{12}) & \sin(\varphi_{13})\cos(\varphi_{23}) \\ 0 & 0 & \sin(\varphi_{13})\sin(\varphi_{23}) \end{pmatrix}$$

To ensure that the elements of the correlation matrix $\boldsymbol{R}$ are constrained between (-1,1), and that positive semi-definiteness of the between-study variance-covariance matrix is satisfied, Uniform(0,π) prior distributions were specified for the spherical parameters, φ_{xy}, where $\pi = 3.142$ (Lu and Ades, 2009; Achana et al., 2014).

As previously mentioned in Section 8.2, the within-study correlations are very rarely reported in clinical trials of multiple outcomes, and estimating these correlations can be difficult. In the MVNMA models described above, the treatment-specific within-study covariance matrices $\boldsymbol{S_{ij}}$ (for MVNMA) and $\boldsymbol{S_{ij_m}}$ (for multivariate hierarchical NMA) are assumed known, and thus the within-study correlations are also assumed known (Jackson et al., 2011). For the purpose of these analyses, within-study correlations were calculated using IPD obtained from the RELAX trial (further details of the RELAX trial are given in Section 4.2.1). To obtain within-study correlations, Pearson correlation coefficients were calculated between the outcomes of interest (Achana et al., 2014), using patient reported bladder diaries (see Section 2.4.1 for further details). To incorporate the uncertainty in estimating the within-study correlations, prior distributions could be specified on the parameters ρw_{xy} using a bootstrapping method (Bujkiewicz et al., 2013). Assuming that all outcomes follow a common multivariate normal distribution, it would be possible to directly obtain within-study correlations from the covariance matrix (Bujkiewicz et al., 2013). However, to aid model computation in this example, within study correlations were assumed known.

8.3.4 Model computation and convergence diagnostics

Multivariate NMA models were estimated using WinBUGS 1.4.3 (Spiegelhalter et al., 2003). Example WinBUGS code for the multivariate hierarchical NMA is given in Section E.1 of Appendix E. Samples were collected for 150,000 MCMC iterations with the first 10,000 iterations discarded in the form of a 'burn-in.' Convergence plots were assessed for a random sample of parameters of interest including treatment effect estimates and between-study standard deviations (Lunn et al., 2012). Brooks-Gelman-Rubin statistics, autocorrelation, history, trace and density plots were used to detect non-convergence for three individual MCMC chains with disparate starting values. A further explanation of these plots, together with their interpretation are given in Section 3.2.2.

8.3.5 Assessing inconsistencies between direct and indirect information

There are no known methods that are currently available to assess inconsistency between direct and indirect information obtained from MVNMA models. However, for individual-intervention NMAs and hierarchical NMAs, Chapters 6 and 7 used the method of node-splitting to assess inconsistency between direct and indirect information for each of the outcomes separately (Dias et al., 2010). Tables C.4 - C.6 of Appendix C displays the results for individual-intervention NMAs. As previously described in Section 6.4.4.1, generally, there were very few inconsistencies between direct and indirect information. For incontinence episodes, there appeared to be inconsistencies between tolterodine ER 4mg q.d versus solifenacin ER 5-10mg q.d, and fesoterodine ER 4mg q.d versus fesoterodine ER 8mg q.d. For voiding episodes, placebo versus electrostimulation, and control versus PFMT appeared to have inconsistent direct and indirect information. Urgency did not appear to have any inconsistencies between closed loops of evidence. Notably, for hierarchical NMAs there appeared to be fewer inconsistencies between treatment comparisons. For this reason, potentially inconsistent studies identified from individual-intervention NMA, were assessed for their impact on results obtained from MVNMAs by individually removing these studies from the analysis.

8.3.6 Goodness of fit and model selection

In a Bayesian framework, model comparison and fit are assessed through the use of DIC statistics and the posterior mean residual deviance (Spiegelhalter et al., 2002) (see Section 3.2.3). The DIC relies on the approximate normality of the posterior distributions for parameter estimates, and requires a plug-in estimate of the deviance for each stochastic parent (Spiegelhalter et al., 2002), for discrete nodes it is unclear as to which estimate to use (Lunn et al., 2012). For this reason, in the WinBUGS software, deviance statistics are not currently available

for mixture likelihoods (Spiegelhalter et al., 2002; Celeux et al., 2006; Spiegelhalter, 2006; Lunn et al., 2012), as the likelihood depends on discrete parameters. Celeux et al. (2006) investigated alternative methods for calculating the deviance and DIC statistics for mixture likelihoods, but found them all problematic (Lunn et al., 2012). Whilst it is not clear how to calculate deviance statistics for MVNMAs, it is possible to calculate residual deviance statistics. Though, as a result of model parameterisation, it is not clear how this would be implemented within the WinBUGS software for MVNMAs. For this particular example, calculating the residual deviance outside of WinBUGS would be both computationally and time intensive due to the large number of parameters and large proportion of missing data. It is clear that further work is required in order to assess model fit and comparison for mixture likelihoods, however, this is beyond the scope of the research presented in this chapter, where the key motivation for fitting MVNMAs was not necessarily to improve model fit, but to utilise correlations between outcomes in order to evaluate all interventions across all outcome measures with the aim to aid decision making. Future work investigating the goodness of fit and model selection is further described in Section 8.5.

8.3.7 Sensitivity analysis

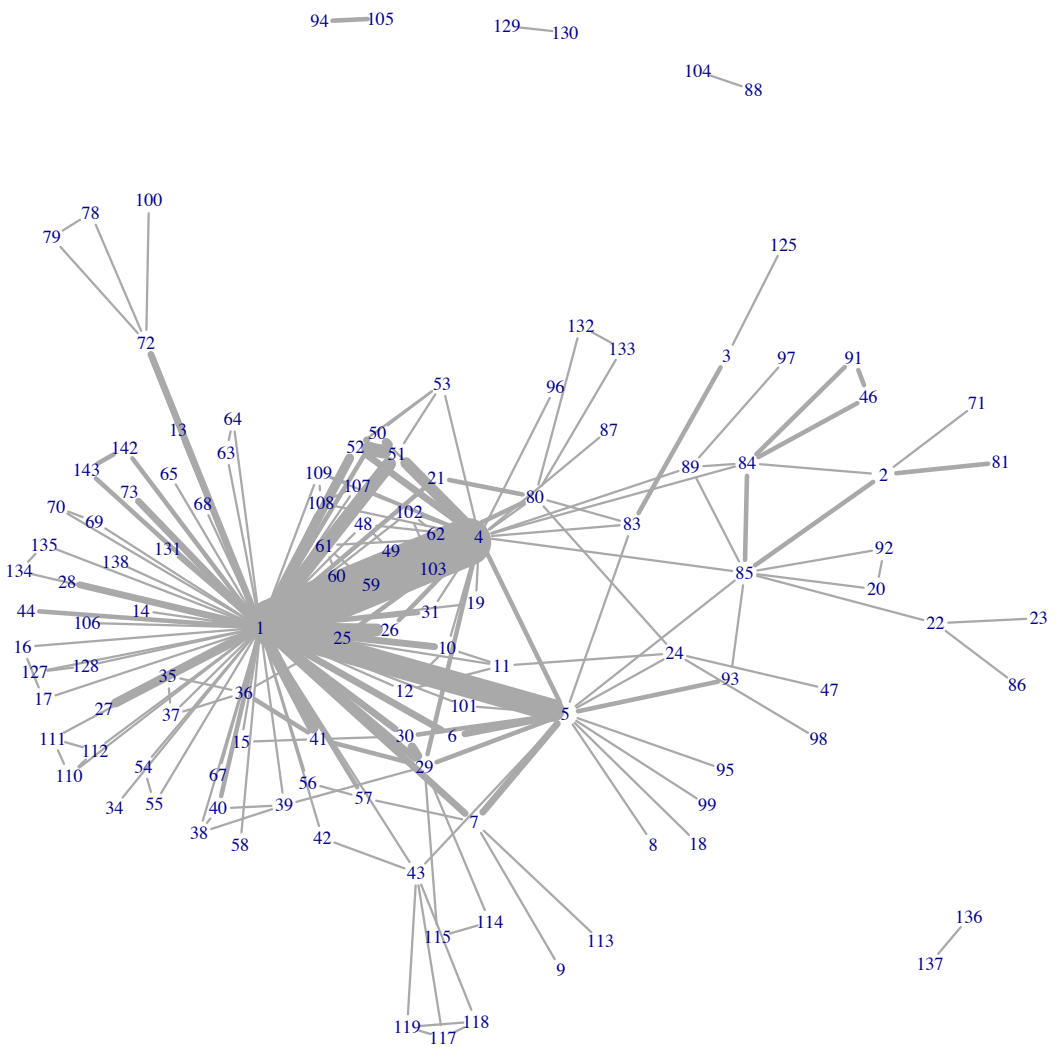
Section 3.4.2 highlighted the importance of assessing the impact of prior specifications on variance parameters. The variance parameters can have a notable impact on the overall effect estimates, and different choices of vague prior distributions should be investigated through a series of sensitivity analyses (Lambert et al., 2005). This is especially true for the elements of $\boldsymbol{V^{1/2}}$, the between-study standard deviations (Higgins et al., 2009), and the deviance of treatment effect profiles across outcomes, ζ. For both variance parameters, two alternative distributions were considered. For prior specification of $\boldsymbol{V^{1/2}}$, 1) Gamma(0.001,0.001) on the precision scale, and 2) Half-normal(0,1) on the standard deviation scale were considered. For prior specification of ζ, 1) Gamma(0.01,0.01) on the precision scale, and 2) Half-normal(0,1) on the standard deviation scale were considered.

8.4 Results

Overall, a multivariate approach allowed for the inclusion of 143 studies evaluating 115 interventions for OAB. Previously, Chapters 6 and 7 included 115, 119, and 60 studies evaluating 97, 100, and 54 interventions for univariate analyses of urinary incontinence, voiding frequency, and urgency, respectively. Figure 8.1 illustrates the network of evidence for MVNMA evaluating incontinence, voiding, and urgency episodes. Incorporating a multivariate approach sufficiently increased the inclusion of interventions for all three outcomes, compared to that of univariate analyses displayed in Figures 5.5, 5.6 and 5.7. This was particularly apparent for urgency episodes, with which all interventions evaluated for incontinence and voiding could now be evaluated for urgency. Furthermore, for interventions that were evaluated for urgency in the original trials but disconnected from the univariate network of evidence (Figure 5.7), e.g. reflexology [71], a MVNMA could borrow information between outcomes for which they were connected, in order to complete the network of evidence. However, eight treatments remained disconnected from the multivariate network of evidence: darifenacin 7.5-15mg once daily + BT [88], tolterodine + BT [94], darifenacin ER 7.5-15mg once daily [104], tolterodine [105], lidocaine gel 2x6ml [129], emepronium bromide immediate release 200mg three times a day [130], sacral nerve stimulation + tolterodine extended release 2mg once daily [136], tolterodine extended release 2mg once daily [137]. These interventions were disconnected from all networks of evidence for univariate NMAs evaluating each of the outcomes of interest; and thus, borrowing information between outcomes had little impact on the inclusion of these particular interventions in multivariate analyses.

Section 8.4.1 contains the results obtained from multivariate hierarchical NMAs which are used throughout the rest of this chapter. Results obtained from multivariate NMAs without a hierarchical approach are given in Table E.1 of Appendix E for completeness.

FIGURE 8.1: Network of evidence for multivariate network meta-analysis



8.4.1 Multivariate hierarchical network meta-analysis

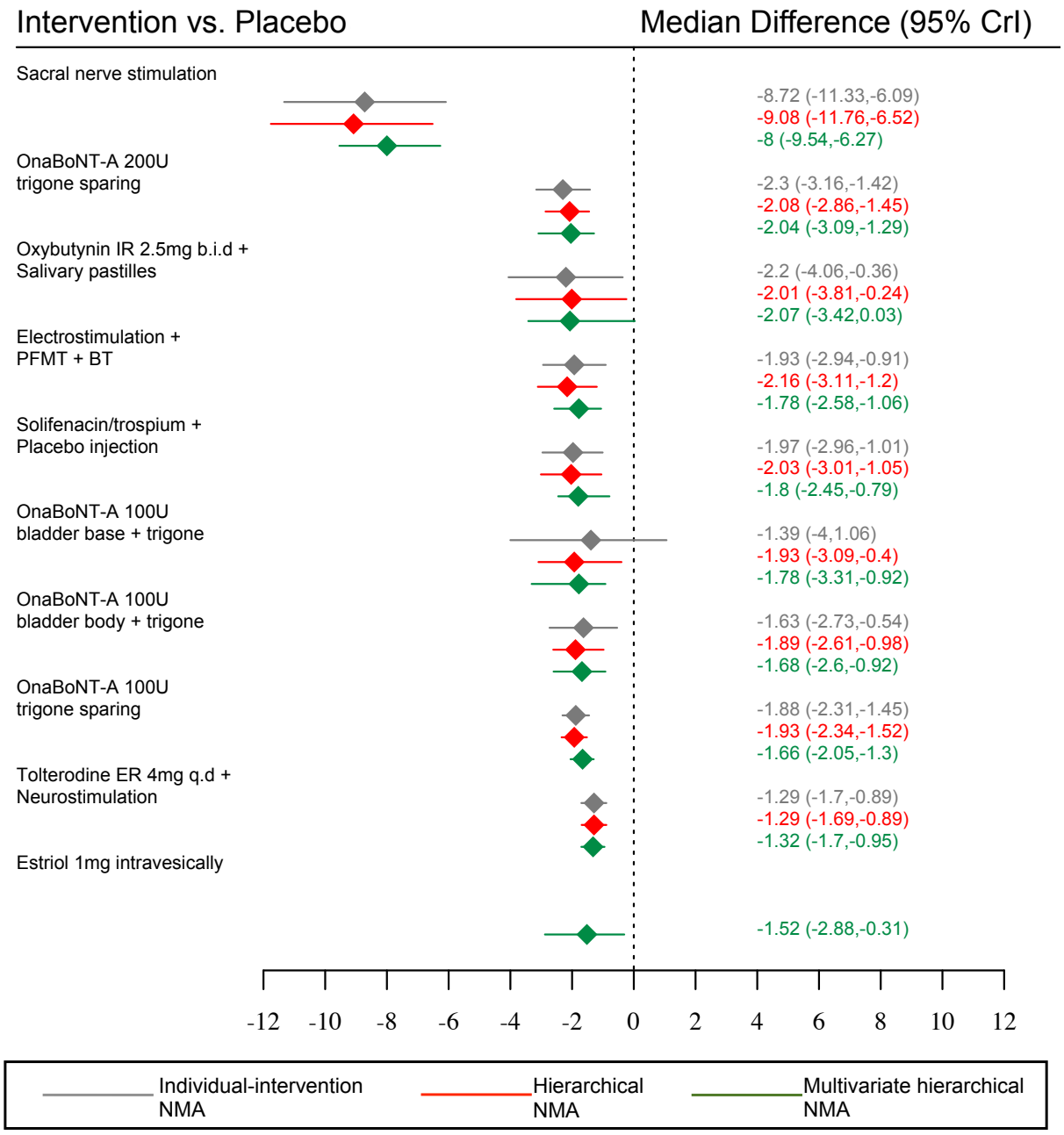
Table 8.1 displays the estimated posterior median reduction and 95% credible intervals for change from baseline in incontinence, voiding, and urgency episodes. Treatment effect estimates were first ranked according to their effectiveness in reducing incontinence episodes, then by voiding frequency, and finally by urgency episodes. Sacral nerve stimulation appeared to be the most effective intervention for the management of urinary incontinence, voiding and urgency with an estimated posterior median reduction of -8 (95%CrI: -9.54,-6.27), -8.19 (95%CrI: -9.69,-6.49) and -8.49 (95%CrI: -10.11,-6.78) episodes, relative to placebo, respectively. Sacral nerve stimulation appeared to dominate the analyses for all three

outcomes with an estimated mean rank of 1 (95%CrI: 1,1) and a probability of being the best intervention overall of 1, across all three outcomes (Table 8.2). However, there are only 2 studies informing the direct treatment comparisons of SNS and thus, these results should be interpreted with caution.

Adopting a multivariate approach changed the overall clinical decision for urgency episodes. In univariate analyses described in Chapters 6 and 7, electrostimulation with vaginal oestrogen cream 1.25mg/day appeared to be the most effective intervention for reducing urgency. Accounting for the correlation between outcomes using MVNMA, SNS appeared to be the most promising intervention. Borrowing information between outcomes generally increased the precision in treatment effect estimates compared to univariate analyses (Figures 8.2, 8.3, and 8.4). This finding was particularly apparent for SNS, with which a multivariate hierarchical approach increased precision in the estimated treatment effects by approximately 60% and 160% for urinary incontinence (Figure 8.2) and voiding frequency (Figure 8.3) respectively, compared to treatment effect estimates obtained from both univariate individual-intervention NMA (described in Chapter 6), and hierarchical NMAs (described in Chapter 7). For all interventions included in each of the univariate analyses, point estimates obtained from multivariate analyses were comparable with point estimates obtained from univariate analyses for all three outcomes (Figures 8.2, 8.3, and 8.4).

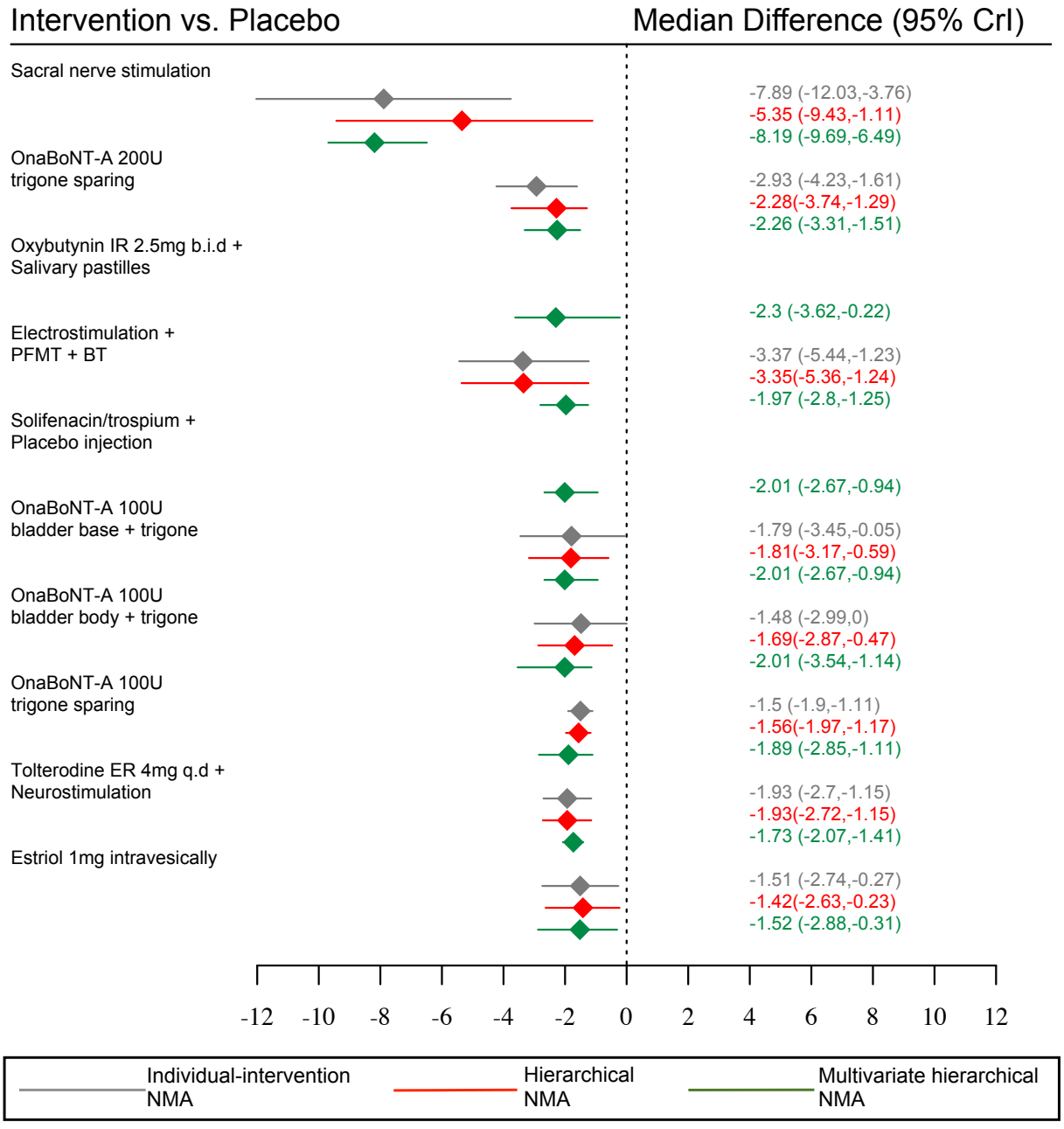
Section E.1 of Appendix E displays the results obtained from a MVNMA without incorporating a hierarchical structure. Incorporating a hierarchical approach in MVNMAs broadly increased the precision in treatment effect estimates (Table 8.1) compared to MVNMAs without a hierarchy (Table E.1 of Appendix E). For example, for SNS incorporating a hierarchical structure further increased precision in the treatment effect estimates by approximately 33%, 35%, and 28% for urinary incontinence, voiding and urgency episodes, respectively.

FIGURE 8.2: Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in incontinence episodes relative to placebo between individual-intervention, hierarchical, and multivariate hierarchical NMA models for the top 10 interventions



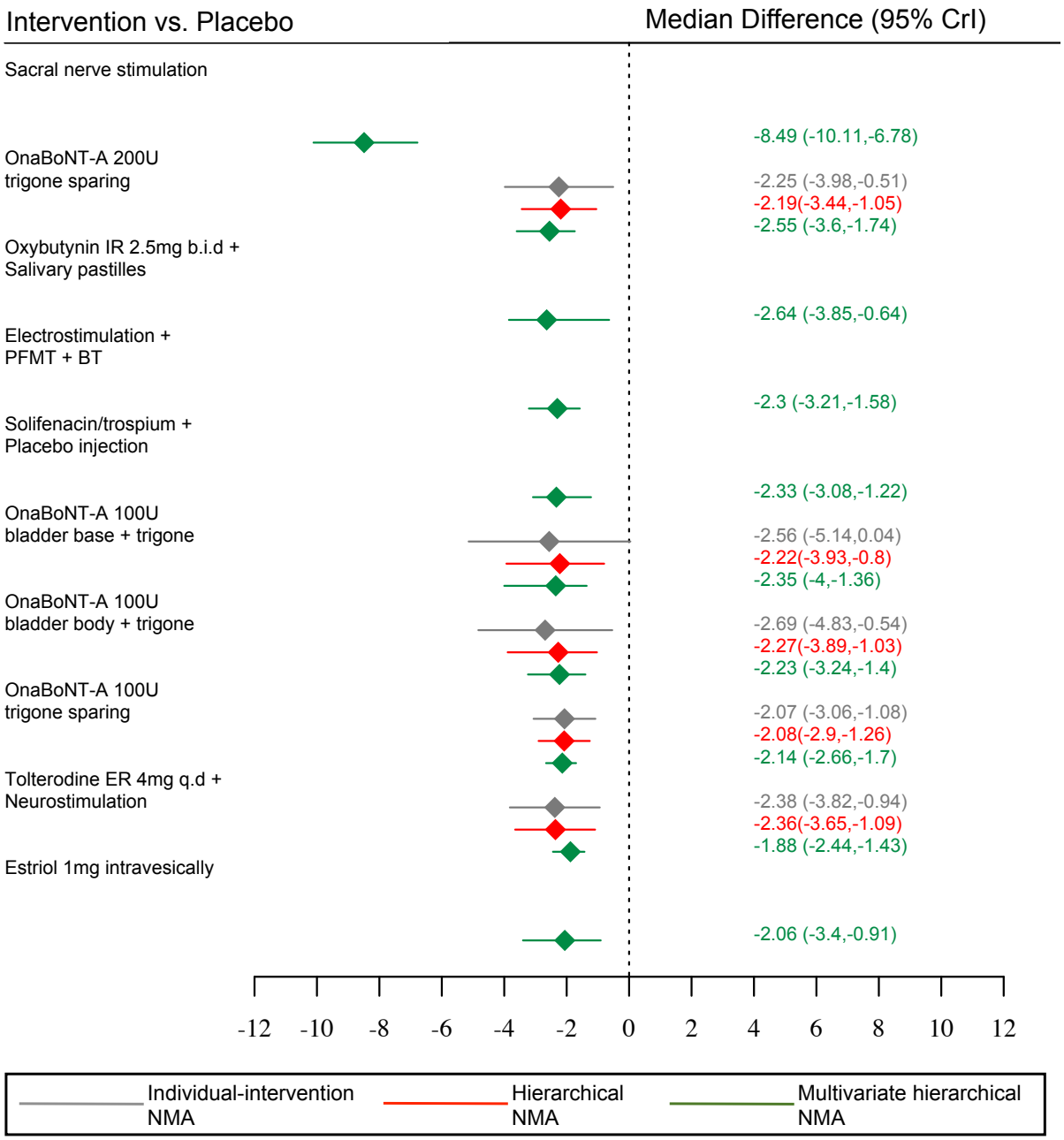
Estriol 1mg intravesically could not be evaluated for urinary incontinence in individual-intervention and hierarchical NMAs as none of the original trials evaluated this intervention for urinary incontinence episodes.

FIGURE 8.3: Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in voiding episodes relative to placebo between individual-intervention, hierarchical, and multivariate hierarchical NMA models for the top 10 interventions



Oxybutynin + salivary pastilles, and solifenacin/trospium + placebo injection could not be evaluated for voiding in individual-intervention and hierarchical NMAs as none of the original trials evaluated these intervention for voiding frequency.

FIGURE 8.4: Comparison of the estimated posterior median difference (and 95% credible intervals) in change from baseline in urgency episodes relative to placebo between individual-intervention, hierarchical, and multivariate hierarchical NMA models for the top 10 interventions



Sacral nerve stimulation, oxybutynin + salivary pastilles, electrostimulation + PFMT + BT, solifenacin/trospium + placebo injection, and estriol 1mg intravesically could not be evaluated for urgency in individual-intervention and hierarchical NMAs as none of the original trials evaluated these interventions for urgency.

TABLE 8.1: Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Incontinence episodes Median difference† (95%CrI)	Voiding episodes Median difference† (95%CrI)	Urgency episodes Median difference† (95%CrI)
	Sacral nerve stimulation	[81]	2	59	-8 (-9.54,-6.27)	-8.19 (-9.69,-6.49)	-8.49 (-10.11,-6.78)
	OnaBoNT-A 200u trigone sparing	[73]	3	114	-2.04 (-3.09,-1.29)	-2.26 (-3.31,-1.51)	-2.55 (-3.6,-1.74)
	Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	1	8	-2.07 (-3.42,0.03)	-2.3 (-3.62,-0.22)	-2.64 (-3.85,-0.64)
	Electrostimulation + PFE + Bladder training	[97]	1	25	-1.78 (-2.58,-1.06)	-1.97 (-2.8,-1.25)	-2.3 (-3.21,-1.58)
	Solifenacin/trospium + placebo injection	[100]	1	118	-1.8 (-2.45,-0.79)	-2.01 (-2.67,-0.94)	-2.33 (-3.08,-1.22)
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-1.78 (-3.31,-0.92)	-2.01 (-3.54,-1.14)	-2.35 (-4,-1.36)
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-1.68 (-2.6,-0.92)	-1.89 (-2.85,-1.11)	-2.23 (-3.24,-1.4)
	OnaBoNTA 100u trigone sparing	[72]	5	716	-1.66 (-2.05,-1.3)	-1.73 (-2.07,-1.41)	-2.14 (-2.66,-1.7)
	Tolerodine ER 4mg q.d + Neurostimulation	[96]	1	20	-1.32 (-1.7,-0.95)	-1.57 (-2.05,-1.14)	-1.88 (-2.44,-1.43)
	Estriol 1mg intravesival	[131]	1	21	-1.52 (-2.88,-0.31)	-1.76 (-3.07,-0.58)	-2.06 (-3.4,-0.91)
	Trospium IR 15mg t.i.d + Physiotherapy	[91]	2	32	-1.16 (-1.95,-0.43)	-1.4 (-2.22,-0.62)	-1.74 (-2.58,-0.93)
	Estradiol 3mg intravaginally	[128]	1	15	-0.87 (-2.13,1.08)	-1.08 (-2.31,0.82)	-1.41 (-2.64,0.53)
	Solifenacin ER 10mg q.d	[30]	5	1160	-0.82 (-1.06,-0.61)	-1.09 (-1.3,-0.9)	-1.37 (-1.74,-1.03)
	Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	1	40	-0.68 (-1.23,-0.22)	-0.89 (-1.48,-0.39)	-1.22 (-1.88,-0.66)
	Tolterodine ER 4mg q.d + BT	[87]	1	154	-0.68 (-1.31,-0.07)	-0.88 (-1.61,-0.25)	-1.23 (-1.96,-0.57)
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	37	-0.85 (-1.42,-0.2)	-1.09 (-1.65,-0.46)	-1.41 (-2.05,-0.69)
	Fesoterodine ER 8mg q.d	[26]	5	2266	-0.69 (-0.87,-0.49)	-0.92 (-1.1,-0.72)	-1.25 (-1.6,-0.9)
	Imidafenacin IR 0.25mg b.i.d	[37]	1	76	-0.72 (-1.32,-0.17)	-0.96 (-1.55,-0.34)	-1.28 (-1.98,-0.62)
	Solifenacin ER (5mg-10mg) q.d	[31]	3	1312	-0.68 (-0.9,-0.42)	-0.84 (-1.06,-0.61)	-1.2 (-1.54,-0.87)
	Solifenacin ER 5mg - 15mg q.d	[34]	1	377	-0.64 (-1,-0.27)	-0.87 (-1.29,-0.45)	-1.2 (-1.66,-0.75)
	Mirabegron 100mg b.i.d	[48]	1	37	-0.68 (-1.06,-0.4)	-0.87 (-1.27,-0.6)	-1.2 (-1.68,-0.88)
	Solabegron 125mg b.i.d	[55]	1	85	-0.62 (-0.9,-0.34)	-0.84 (-1.12,-0.59)	-1.16 (-1.58,-0.79)
	Propiverine ER 30mg b.i.d	[117]	1	45	-0.61 (-1.81,-0.1)	-0.85 (-2.06,-0.37)	-1.18 (-2.42,-0.6)
	Darifenacin ER 30mg q.d	[38]	1	115	-0.59 (-1.15,-0.05)	-0.8 (-1.41,-0.2)	-1.15 (-1.73,-0.54)
	Mirabegron ER 25mg q.d	[50]	4	763	-0.63 (-0.87,-0.42)	-0.85 (-1.05,-0.64)	-1.16 (-1.48,-0.84)
	Mirabegron IR 150mg b.i.d	[49]	1	41	-0.61 (-0.91,-0.29)	-0.87 (-1.14,-0.55)	-1.18 (-1.53,-0.81)
	Mirabegron ER 100mg q.d	[52]	6	1722	-0.62 (-0.81,-0.44)	-0.8 (-0.96,-0.61)	-1.15 (-1.47,-0.87)
	Mirabegron ER 200mg q.d	[53]	1	166	-0.62 (-0.91,-0.27)	-0.83 (-1.13,-0.47)	-1.17 (-1.53,-0.79)
	Oxybutynin ER 10mg q.d	[8]	1	185	-0.58 (-1.09,-0.25)	-0.81 (-1.27,-0.47)	-1.14 (-1.57,-0.7)
	Propiverine ER 30mg q.d	[42]	1	391	-0.56 (-1.05,-0.17)	-0.76 (-1.33,-0.37)	-1.08 (-1.72,-0.66)
	Oxybutynin IR 3mg t.i.d	[19]	1	244	-0.6 (-0.94,-0.3)	-0.75 (-1.12,-0.45)	-1.09 (-1.56,-0.7)
	Mirabegron ER 50mg q.d	[51]	9	2289	-0.59 (-0.78,-0.42)	-0.84 (-0.99,-0.67)	-1.15 (-1.45,-0.84)
	Solifenacin ER 5mg q.d	[29]	10	1152	-0.6 (-0.8,-0.38)	-0.74 (-0.92,-0.58)	-1.14 (-1.48,-0.83)
	Tolterodine IR 2mg b.i.d + BT	[93]	2	172	-0.57 (-1,0)	-0.86 (-1.28,-0.2)	-1.15 (-1.7,-0.47)
	Cizolirtine Citrate 400mg b.i.d	[57]	2	68	-0.56 (-1.1,-0.05)	-0.8 (-1.37,-0.25)	-1.16 (-1.78,-0.56)
	Trospium ER 60mg q.d	[44]	2	565	-0.55 (-0.92,-0.21)	-0.75 (-1.13,-0.41)	-1.08 (-1.59,-0.61)
	Propiverine IR 45mg t.i.d	[118]	1	48	-0.52 (-1.14,0.1)	-0.73 (-1.35,-0.15)	-1.06 (-1.73,-0.39)
	Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	1	130	-0.5 (-0.81,-0.23)	-0.74 (-1.11,-0.43)	-1.07 (-1.53,-0.68)
	Fesoterodine ER 4mg q.d	[25]	7	2971	-0.49 (-0.66,-0.32)	-0.7 (-0.86,-0.53)	-1.03 (-1.38,-0.72)
	Pregabalin 150mg b.i.d	[62]	1	41	-0.5 (-1.08,0.12)	-0.75 (-1.3,-0.18)	-1.06 (-1.74,-0.35)
	Darifenacin ER 15mg q.d	[40]	2	319	-0.46 (-0.83,0)	-0.65 (-1.05,-0.17)	-0.99 (-1.48,-0.47)
	Propiverine IR 15mg b.i.d	[43]	3	540	-0.46 (-0.88,-0.1)	-0.62 (-1.07,-0.27)	-0.98 (-1.56,-0.58)
	Tolterodine ER 4mg q.d	[4]	31	7808	-0.5 (-0.6,-0.4)	-0.62 (-0.74,-0.52)	-0.99 (-1.29,-0.75)
	Oxybutynin IR 5mg t.i.d	[7]	6	324	-0.45 (-0.78,-0.18)	-0.64 (-0.93,-0.36)	-0.98 (-1.37,-0.61)
	Propiverine ER 20mg q.d	[41]	7	1523	-0.42 (-0.6,-0.22)	-0.65 (-0.84,-0.48)	-0.98 (-1.35,-0.67)
	Tolterodine IR 2mg b.i.d	[5]	21	3469	-0.44 (-0.57,-0.3)	-0.67 (-0.82,-0.53)	-0.98 (-1.31,-0.7)
	Propiverine ER 60mg q.d	[119]	1	39	-0.41 (-0.89,0.72)	-0.6 (-1.09,0.44)	-0.92 (-1.55,0.23)
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	-0.41 (-0.97,-0.02)	-0.61 (-1.17,-0.23)	-0.96 (-1.52,-0.52)
	Oxybutynin IR 2.5-5mg b.i.d	[24]	5	804	-0.46 (-0.8,-0.16)	-0.63 (-0.95,-0.38)	-1 (-1.36,-0.6)
	Oxybutynin chloride topical gel 1g q.d	[13]	1	389	-0.43 (-0.78,-0.13)	-0.64 (-0.97,-0.35)	-0.99 (-1.35,-0.59)
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	-0.39 (-0.78,0.01)	-0.63 (-1.01,-0.23)	-0.95 (-1.4,-0.44)
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	-0.44 (-0.98,0.42)	-0.65 (-1.23,0.2)	-0.98 (-1.62,-0.2)
	PFMT + BT	[89]	2	67	-0.38 (-0.83,0.03)	-0.66 (-1.16,-0.18)	-0.95 (-1.55,-0.44)
	Tolterodine IR 1mg b.i.d	[6]	3	250	-0.39 (-0.64,-0.08)	-0.62 (-0.91,-0.31)	-0.94 (-1.35,-0.58)
	Fesoterodine ER 4mg-8mg q.d	[27]	4	1485	-0.37 (-0.59,-0.17)	-0.67 (-0.86,-0.46)	-0.96 (-1.33,-0.62)
	Oxybutynin gel 84mg/day	[134]	1	211	-0.37 (-0.76,0)	-0.63 (-0.99,-0.28)	-0.96 (-1.33,-0.51)
	Oxybutynin transdermal 3.9mg/day	[10]	3	292	-0.38 (-0.63,-0.14)	-0.61 (-0.86,-0.35)	-0.94 (-1.29,-0.58)
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	-0.37 (-0.67,-0.03)	-0.56 (-0.87,-0.22)	-0.92 (-1.31,-0.48)
	Imidafenacin 0.1mg b.i.d	[36]	4	563	-0.38 (-0.69,-0.1)	-0.55 (-0.81,-0.29)	-0.9 (-1.3,-0.5)
	Terodiline 25mg b.i.d	[28]	3	126	-0.37 (-0.78,0.14)	-0.51 (-0.9,-0.05)	-0.88 (-1.44,-0.3)
	Darifenacin ER 7.5mg q.d	[39]	2	137	-0.35 (-0.81,0.19)	-0.58 (-1.08,0.02)	-0.92 (-1.46,-0.29)
	Oxybutynin gel 56mg/day	[135]	1	198	-0.35 (-0.69,0.06)	-0.52 (-0.82,-0.1)	-0.88 (-1.3,-0.38)
	Oxbutynin patch 73.5mg	[15]	1	391	-0.34 (-0.57,-0.02)	-0.56 (-0.86,-0.15)	-0.89 (-1.27,-0.45)
	Elocalcitol 75mg	[70]	1	84	-0.37 (-0.89,0.06)	-0.5 (-1.06,-0.07)	-0.89 (-1.57,-0.38)
	Oxybutynin 20mg intravesically q.d	[106]	1	21	-0.33 (-0.75,0.42)	-0.55 (-0.95,0.27)	-0.9 (-1.34,-0.05)
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.32 (-0.76,0.23)	-0.57 (-1.03,0.03)	-0.9 (-1.42,-0.22)
	Oxybutynin ER 15mg q.d	[9]	1	53	-0.31 (-0.78,0.12)	-0.53 (-0.98,-0.07)	-0.88 (-1.32,-0.31)
	Oxybutynin IR 5-20mg	[23]	1	52	-0.29 (-0.67,0.26)	-0.52 (-0.9,0.07)	-0.85 (-1.32,-0.13)
	Oxybutynin ER 5-30mg q.d	[22]	3	109	-0.3 (-0.63,0.02)	-0.5 (-0.88,-0.12)	-0.87 (-1.27,-0.39)
	Trospium chloride IR 45mg t.i.d	[47]	1	615	-0.33 (-0.81,0.18)	-0.55 (-0.95,-0.12)	-0.89 (-1.38,-0.35)
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.22 (-1.26,0.73)	-0.46 (-1.48,0.54)	-0.77 (-1.87,0.26)
	Pelvic floor muscle training (PFMT)/Physiotherapy	[84]	4	83	-0.33 (-0.82,0.15)	-0.59 (-1.17,-0.01)	-0.89 (-1.54,-0.35)
	Oxybutynin transdermal 1.3mg/day	[11]	2	166	-0.27 (-0.56,0.14)	-0.48 (-0.78,-0.04)	-0.82 (-1.26,-0.29)
	Elocalcitol 150mg	[69]	1	87	-0.34 (-0.78,0.19)	-0.55 (-1.05,-0.04)	-0.86 (-1.48,-0.27)
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-0.27 (-0.68,0.36)	-0.54 (-0.95,0.15)	-0.86 (-1.33,-0.04)
	Duloxetine 40mg b.i.d	[65]	1	81	-0.3 (-0.86,0.26)	-0.52 (-1.16,0.09)	-0.86 (-1.56,-0.16)
	Bladder Training (BT)/ Behaviour Therapy	[85]	6	192	-0.26 (-0.56,0.05)	-0.45 (-0.8,-0.13)	-0.8 (-1.24,-0.4)
	Solabegron IR 50mg b.i.d	[54]	1	88	-0.24 (-0.52,0.05)	-0.46 (-0.74,-0.18)	-0.79 (-1.22,-0.4)
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-0.19 (-0.45,0.1)	-0.58 (-0.99,-0.17)	-0.77 (-1.17,-0.34)
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	-0.18 (-0.52,0.35)	-0.43 (-0.75,0.16)	-0.77 (-1.2,-0.11)
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	47	-0.24 (-0.67,0.24)	-0.48 (-0.9,-0.04)	-0.78 (-1.33,-0.25)
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.35 (-1.14,0.48)	-0.57 (-1.35,0.27)	-0.88 (-1.75,0.03)
	Oxybutynin IR 5mg b.i.d	[18]	1	116	-0.15 (-0.49,0.35)	-0.39 (-0.78,0.15)	-0.75 (-1.19,-0.09)
	Lipo-BoNTA 200U	[138]	1	29	-0.04 (-0.92,0.78)	-0.26 (-1.21,0.61)	-0.57 (-1.53,0.29)
	Serlopitant 0.25mg q.d	[107]	1	110	-0.08 (-0.56,0.45)	-0.31 (-0.79,0.23)	-0.64 (-1.2,-0.01)
	Serlopitant 4mg q.d	[109]	1	114	-0.07 (-0.61,0.44)	-0.3 (-0.81,0.24)	-0.63 (-1.21,0.01)
	Tarafenacin 0.4mg q.d	[82]	2	152	-0.19 (-0.82,0.53)	-0.43 (-1.04,0.29)	-0.75 (-1.4,-0.02)

TABLE 8.1: Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis (cont.)

	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-0.08 (-0.77,0.61)	-0.22 (-0.91,0.5)	-0.68 (-1.51,0.16)
	Electrostimulation	[80]	6	236	-0.05 (-0.48,0.37)	-0.35 (-0.89,0.15)	-0.76 (-1.44,-0.16)
	Serlopitant 1mg q.d	[108]	1	110	0.06 (-0.48,0.61)	-0.15 (-0.63,0.39)	-0.5 (-1.08,0.15)
	Estradiol 25mg	[68]	1	20	0.01 (-0.4,0.51)	-0.19 (-0.68,0.32)	-0.53 (-1.09,0.04)
	Placebo	[1]	89	15490	NA	NA	NA
	Netupitant 200mg q.d	[112]	1	55	-0.09 (-0.98,0.74)	-0.3 (-1.23,0.51)	-0.63 (-1.54,0.2)
	Netupitant 100mg q.d	[111]	1	59	-0.04 (-0.84,1.06)	-0.25 (-1.08,0.84)	-0.59 (-1.48,0.47)
	Tarafenacin 0.2mg q.d	[90]	2	154	0.01 (-0.66,0.7)	-0.18 (-0.89,0.51)	-0.52 (-1.25,0.14)
	Trospium IR 15mg t.i.d	[46]	2	30	-0.03 (-0.74,0.71)	-0.18 (-1.02,0.58)	-0.54 (-1.39,0.27)
	ZD0947IL 25mg/day	[58]	1	92	0.07 (-0.61,0.87)	-0.1 (-0.84,0.67)	-0.42 (-1.28,0.24)
	Netupitant 50mg q.d	[110]	1	60	0.06 (-0.61,0.81)	-0.14 (-0.83,0.51)	-0.48 (-1.33,0.22)
	Electromagnetic stimulation	[125]	1	33	-1.25 (-4.03,0.42)	-1.47 (-4.25,0.17)	-1.85 (-4.57,-0.06)
	Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	3	109	0.19 (-0.63,1.05)	-0.04 (-0.92,0.84)	-0.4 (-1.24,0.6)
	ONO-8539 100mg b.i.d	[60]	1	83	0.15 (-0.43,0.77)	-0.04 (-0.62,0.55)	-0.37 (-1,0.3)
	Percutaneous tibial nerve stimulation	[83]	5	198	-0.42 (-1.06,0.4)	-0.65 (-1.33,0.13)	-0.95 (-1.74,-0.22)
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	0.28 (-0.22,0.75)	0 (-0.57,0.52)	-0.26 (-0.91,0.34)
	Flavoxate chloride 200mg q.d	[64]	1	19	0.28 (-0.35,1.19)	0.05 (-0.59,0.97)	-0.29 (-0.93,0.64)
	Resiniferatoxin 50nM	[67]	1	34	0.08 (-1.43,1.12)	-0.15 (-1.68,0.88)	-0.47 (-2.07,0.63)
	Emepronium bromide ER 200mg q.d	[63]	1	19	0.36 (-0.65,1.17)	0.15 (-0.86,0.98)	-0.21 (-1.17,0.65)
	ONO-8539 300mg b.i.d	[61]	1	82	0.43 (-0.11,1.07)	0.2 (-0.31,0.85)	-0.17 (-0.67,0.61)
	Propantheline Bromide 15mg t.i.d	[113]	1	23	0.69 (-0.88,1.59)	0.46 (-0.99,1.35)	0.11 (-1.41,1.12)
	Estradiol 1mg intravaginally	[127]	1	15	1.02 (-0.45,2.47)	0.82 (-0.61,2.26)	0.46 (-0.95,1.94)
	ONO-8539 30mg b.i.d	[59]	1	87	0.58 (-0.1,1.2)	0.35 (-0.32,0.98)	0.02 (-0.67,0.7)
	Control	[2]	5	180	0.49 (0,1.09)	0.33 (-0.1,0.94)	-0.02 (-0.62,0.6)
	Reflexology	[71]	1	54	0.46 (-0.12,1.22)	0.22 (-0.45,0.97)	-0.06 (-0.79,0.7)
	Sham Therapy	[3]	3	157	0.2 (-0.45,0.96)	0 (-0.65,0.73)	-0.34 (-1.08,0.34)
	Naftopidil 25mg q.d	[114]	1	22	4.06 (2.41,5.02)	3.8 (2.21,4.79)	3.48 (1.81,4.49)
	Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	1	21	5.19 (3.67,7.37)	4.95 (3.47,7.15)	4.6 (3.22,6.91)

† median relative to a placebo intervention

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

TABLE 8.2: Estimated intervention rank (95% credible intervals) and probability that each intervention is the best for the management of urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Incontinence episodes		Voiding episodes		Urgency episodes	
					Rank (95%CrI)	p(Best)	Rank(95%CrI)	p(Best)	Rank(95%CrI)	p(Best)
	Sacral nerve stimulation	[81]	2	59	1 (1,1)	1	1 (1,1)	1	1 (1,1)	1
	OnaBoNT-A 200u trigone sparing	[73]	3	114	3 (2,9)	0	3 (2,9)	0	3 (2,9)	0
	Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	1	8	4 (2,90)	0	4 (2,90)	0	4 (2,90)	0
	Electrostimulation + PFE + Bladder training	[97]	1	25	4 (2,12)	0	5 (2,12)	0	5 (2,12)	0
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	6 (2,24)	0	6 (2,23)	0	6 (2,24)	0
	Solifenacin/trospium + placebo injection	[100]	1	118	6 (2,21)	0	6 (2,22)	0	6 (2,22)	0
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	7 (3,27)	0	7 (3,28)	0	7 (3,28)	0
	OnaBoNTA 100u trigone sparing	[72]	5	716	7 (4,11)	0	8 (4,11)	0	7 (4,11)	0
	Tolerodine ER 4mg q.d + Neurostimulation	[96]	1	20	9 (5,15)	0	9 (4,15)	0	9 (4,15)	0
	Estriol 1mg intravesival	[131]	1	21	10 (2,88)	0	10 (2,88)	0	10 (2,88)	0
	Trospium IR 15mg t.i.d + Physiotherapy	[91]	2	32	12 (4,69)	0	12 (4,69)	0	12 (4,70)	0
	Estradiol 3mg intravaginally	[128]	1	15	14 (2,102)	0	14 (2,101)	0	14 (2,101)	0
	Solifenacin ER 10mg q.d	[30]	5	1160	17 (11,30)	0	16 (11,26)	0	17 (11,36)	0
	Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	1	40	22 (10,74)	0	22 (10,77)	0	22 (10,77)	0
	Tolterodine ER 4mg q.d + BT	[87]	1	154	22 (9,83)	0	22 (9,85)	0	23 (9,84)	0
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	37	23 (8,88)	0	22 (8,86)	0	23 (8,88)	0
	Fesoterodine ER 8mg q.d	[26]	5	2266	23 (14,42)	0	23 (14,41)	0	23 (13,48)	0
	Imidafenacin IR 0.25mg b.i.d	[37]	1	76	24 (10,78)	0	24 (10,79)	0	25 (10,79)	0
	Solifenacin ER (5mg-10mg) q.d	[31]	3	1312	26 (14,52)	0	28 (15,56)	0	27 (14,59)	0
	Solifenacin ER 5mg - 15mg q.d	[34]	1	377	27 (12,73)	0	27 (12,76)	0	27 (12,76)	0
	Mirabegron 100mg b.i.d	[48]	1	37	27 (12,59)	0	28 (12,59)	0	29 (12,61)	0
	Solabegron 125mg b.i.d	[55]	1	85	28 (14,61)	0	28 (14,59)	0	28 (14,66)	0
	Propiverine ER 30mg b.i.d	[117]	1	45	30 (4,81)	0	28 (4,79)	0	30 (4,80)	0
	Darifenacin ER 30mg q.d	[38]	1	115	30 (10,86)	0	31 (10,87)	0	31 (10,87)	0
	Mirabegron ER 25mg q.d	[50]	4	763	30 (16,58)	0	31 (16,58)	0	32 (16,63)	0
	Mirabegron IR 150mg b.i.d	[49]	1	41	32 (14,68)	0	30 (14,65)	0	31 (14,67)	0
	Mirabegron ER 100mg q.d	[52]	6	1722	32 (18,55)	0	35 (20,59)	0	33 (18,62)	0
	Mirabegron ER 200mg q.d	[53]	1	166	33 (15,71)	0	32 (15,72)	0	32 (15,72)	0
	Oxybutynin ER 10mg q.d	[8]	1	185	33 (11,74)	0	33 (11,74)	0	35 (11,77)	0
	Propiverine ER 30mg q.d	[42]	1	391	33 (12,74)	0	35 (12,78)	0	35 (12,77)	0
	Oxybutynin IR 3mg t.i.d	[19]	1	244	33 (13,69)	0	38 (14,75)	0	38 (14,76)	0
	Mirabegron ER 50mg q.d	[51]	9	2289	34 (20,56)	0	31 (19,51)	0	33 (19,62)	0
	Solifenacin ER 5mg q.d	[29]	10	1152	34 (18,60)	0	39 (21,65)	0	35 (17,64)	0
	Tolterodine IR 2mg b.i.d + BT	[93]	2	172	36 (11,89)	0	32 (11,87)	0	35 (11,89)	0
	Cizolirtine Citrate 400mg b.i.d	[57]	2	68	36 (12,90)	0	34 (11,90)	0	34 (11,90)	0
	Trospium ER 60mg q.d	[44]	2	565	37 (14,79)	0	39 (15,80)	0	39 (14,82)	0
	Propiverine IR 45mg t.i.d	[118]	1	48	42 (10,98)	0	42 (10,97)	0	42 (10,97)	0
	Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	1	130	44 (17,80)	0	43 (17,80)	0	43 (16,82)	0
	Fesoterodine ER 4mg q.d	[25]	7	2971	44 (25,67)	0	45 (25,67)	0	44 (22,72)	0
	Pregabalin 150mg b.i.d	[62]	1	41	45 (13,95)	0	42 (13,92)	0	44 (13,94)	0
	Darifenacin ER 15mg q.d	[40]	2	319	45 (12,91)	0	47 (13,93)	0	46 (12,93)	0
	Propiverine IR 15mg b.i.d	[43]	3	540	46 (19,81)	0	49 (19,86)	0	47 (18,84)	0
	Tolterodine ER 4mg q.d	[4]	31	7808	46 (33,61)	0	55 (39,72)	0	50 (30,73)	0
	Oxybutynin IR 5mg t.i.d	[7]	6	324	48 (19,78)	0	50 (21,80)	0	50 (19,81)	0
	Propiverine ER 20mg q.d	[41]	7	1523	52 (31,76)	0	50 (28,74)	0	50 (26,79)	0
	Tolterodine IR 2mg b.i.d	[5]	21	3469	52 (36,69)	0	50 (33,69)	0	52 (30,77)	0
	Propiverine ER 60mg q.d	[119]	1	39	52 (14,104)	0	53 (15,104)	0	51 (15,103)	0
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	53 (13,89)	0	54 (14,89)	0	54 (14,89)	0
	Oxybutynin IR 2.5-5mg b.i.d	[24]	5	804	53 (19,84)	0	55 (21,85)	0	54 (20,85)	0
	Oxybutynin chloride topical gel 1g q.d	[13]	1	389	55 (22,84)	0	55 (22,84)	0	56 (22,85)	0
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	56 (19,86)	0	54 (18,85)	0	55 (18,86)	0
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	56 (14,98)	0	55 (13,99)	0	56 (13,98)	0
	PFMT + BT	[89]	2	67	57 (14,95)	0	51 (13,94)	0	54 (13,95)	0
	Tolterodine IR 1mg b.i.d	[6]	3	250	57 (29,88)	0	56 (27,88)	0	56 (26,89)	0
	Fesoterodine ER 4mg-8mg q.d	[27]	4	1485	58 (31,82)	0	51 (28,77)	0	54 (26,83)	0
	Oxybutynin gel 84mg/day	[134]	1	211	59 (21,88)	0	55 (20,85)	0	57 (20,87)	0
	Oxybutynin transdermal 3.9mg/day	[10]	3	292	59 (30,81)	0	58 (21,81)	0	59 (28,84)	0
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	61 (23,89)	0	62 (24,90)	0	61 (23,90)	0
	Imidafenacin 0.1mg b.i.d	[36]	4	563	61 (31,85)	0	65 (35,87)	0	63 (30,88)	0
	Terodiline 25mg b.i.d	[28]	3	126	61 (18,95)	0	67 (22,96)	0	64 (19,96)	0
	Darifenacin ER 7.5mg q.d	[39]	2	137	63 (14,98)	0	63 (14,97)	0	61 (14,97)	0
	Oxybutynin gel 56mg/day	[135]	1	198	63 (24,92)	0	68 (29,93)	0	65 (26,93)	0
	Oxbutynin patch 73.5mg	[15]	1	391	64 (33,88)	0	64 (29,89)	0	63 (29,89)	0
	Elocalcitol 75mg	[70]	1	84	65 (15,100)	0	71 (16,100)	0	67 (15,100)	0
	Oxybutynin 20mg intravesically q.d	[106]	1	21	66 (23,102)	0	67 (24,101)	0	66 (23,101)	0
	Imidafenacin 0.05mg b.i.d	[35]	1	91	67 (22,97)	0	65 (21,96)	0	66 (21,97)	0
	Oxybutynin ER 15mg q.d	[9]	1	53	67 (25,96)	0	68 (25,96)	0	67 (25,95)	0
	Oxybutynin IR 5-20mg	[23]	1	52	68 (24,99)	0	68 (24,98)	0	67 (24,98)	0
	Oxybutynin ER 5-30mg q.d	[22]	3	109	68 (32,93)	0	70 (32,94)	0	68 (30,93)	0
	Trospium chloride IR 45mg t.i.d	[47]	1	615	70 (20,101)	0	70 (21,98)	0	70 (20,99)	0
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	70 (9,110)	0	70 (9,110)	0	70 (9,110)	0
	Pelvic floor muscle training (PFMT)/Physiotherapy	[84]	4	83	71 (19,98)	0	69 (16,98)	0	70 (17,98)	0
	Oxybutynin transdermal 1.3mg/day	[11]	2	166	71 (33,96)	0	72 (33,96)	0	71 (32,96)	0
	Elocalcitol 150mg	[69]	1	87	71 (15,102)	0	72 (15,102)	0	72 (15,102)	0
	Oxybutynin ER 2.5mg q.d	[20]	1	16	73 (32,102)	0	69 (28,101)	0	71 (29,100)	0
	Duloxetine 40mg b.i.d	[65]	1	81	73 (20,101)	0	73 (18,101)	0	73 (18,101)	0
	Bladder Training (BT)/ Behaviour Therapy	[85]	6	192	73 (35,95)	0	74 (32,96)	0	73 (31,96)	0
	Solabegron IR 50mg b.i.d	[54]	1	88	74 (41,93)	0	73 (41,92)	0	73 (37,94)	0
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	76 (46,95)	0	66 (24,92)	0	75 (41,95)	0
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	78 (42,101)	0	77 (40,100)	0	76 (38,100)	0
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	47	78 (29,101)	0	77 (28,100)	0	78 (27,101)	0
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	79 (13,108)	0	79 (13,108)	0	79 (13,108)	0
	Oxybutynin IR 5mg b.i.d	[18]	1	116	81 (43,104)	0	79 (40,104)	0	79 (39,103)	0
	Lipo-BoNTA 200U	[138]	1	29	81 (12,109)	0	80 (12,109)	0	80 (12,109)	0
	Serlopitant 0.25mg q.d	[107]	1	110	82 (30,104)	0	82 (32,102)	0	82 (30,102)	0
	Serlopitant 4mg q.d	[109]	1	114	83 (30,103)	0	82 (33,102)	0	82 (30,102)	0
	Tarafenacin 0.4mg q.d	[82]	2	152	83 (19,104)	0	82 (18,103)	0	82 (18,103)	0

TABLE 8.2: Estimated intervention rank (95% credible intervals) and probability that each intervention is the best for the management of urinary incontinence, voiding and urgency episodes obtained from multivariate hierarchical network meta-analysis (cont.)

Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	84 (19,106)	0	87 (22,107)	0	83 (17,105)	0
Electrostimulation	[80]	6	236	86 (42,103)	0	83 (33,101)	0	80 (26,100)	0
Serlopitant 1mg q.d	[108]	1	110	91 (47,108)	0	91 (52,108)	0	90 (47,107)	0
Estradiol 25mg	[68]	1	20	91 (54,106)	0	91 (47,106)	0	91 (48,105)	0
Placebo	[1]	89	15490	91 (82,100)	0	99 (93,106)	0	106 (99,111)	0
Netupitant 200mg q.d	[112]	1	55	93 (17,111)	0	92 (18,111)	0	93 (18,111)	0
Netupitant 100mg q.d	[111]	1	59	93 (18,111)	0	93 (19,111)	0	92 (18,111)	0
Tarafenacin 0.2mg q.d	[90]	2	154	94 (38,109)	0	94 (37,109)	0	94 (36,109)	0
Trospium IR 15mg t.i.d	[46]	2	30	94 (30,111)	0	95 (33,112)	0	94 (31,111)	0
ZD0947IL 25mg/day	[58]	1	92	96 (29,111)	0	96 (31,111)	0	96 (29,111)	0
Netupitant 50mg q.d	[110]	1	60	97 (24,111)	0	96 (25,111)	0	96 (25,111)	0
Electromagnetic stimulation	[125]	1	33	97 (4,113)	0	96 (4,113)	0	96 (4,113)	0
Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	3	109	98 (26,112)	0	97 (25,112)	0	97 (25,112)	0
ONO-8539 100mg b.i.d	[60]	1	83	98 (46,110)	0	98 (49,110)	0	98 (46,110)	0
Percutaneous tibial nerve stimulation	[83]	5	198	99 (20,112)	0	97 (19,112)	0	97 (19,112)	0
Vaginal oestrogen cream 1.25mg/day	[132]	1	98	100 (73,110)	0	99 (69,110)	0	100 (72,110)	0
Flavoxate chloride 200mg q.d	[64]	1	19	102 (57,112)	0	102 (52,112)	0	101 (53,112)	0
Resiniferatoxin 50nM	[67]	1	34	102 (20,113)	0	102 (20,113)	0	101 (19,113)	0
Emepronium bromide ER 200mg q.d	[63]	1	19	103 (65,112)	0	103 (63,112)	0	103 (62,112)	0
ONO-8539 300mg b.i.d	[61]	1	82	105 (83,113)	0	105 (83,113)	0	104 (81,113)	0
Propantheline Bromide 15mg t.i.d	[113]	1	23	106 (27,113)	0	106 (28,113)	0	105 (27,113)	0
Estradiol 1mg intravaginally	[127]	1	15	107 (20,113)	0	106 (20,113)	0	106 (20,113)	0
ONO-8539 30mg b.i.d	[59]	1	87	107 (83,113)	0	107 (82,113)	0	106 (82,113)	0
Control	[2]	5	180	107 (91,112)	0	108 (92,113)	0	107 (90,113)	0
Reflexology	[71]	1	54	108 (82,113)	0	107 (78,113)	0	107 (80,113)	0
Sham Therapy	[3]	3	157	111 (84,113)	0	111 (87,113)	0	111 (85,113)	0
Naftopidil 25mg q.d	[114]	1	22	114 (113,115)	0	114 (113,115)	0	114 (113,115)	0
Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	1	21	115 (114,115)	0	115 (114,115)	0	115 (114,115)	0

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average.

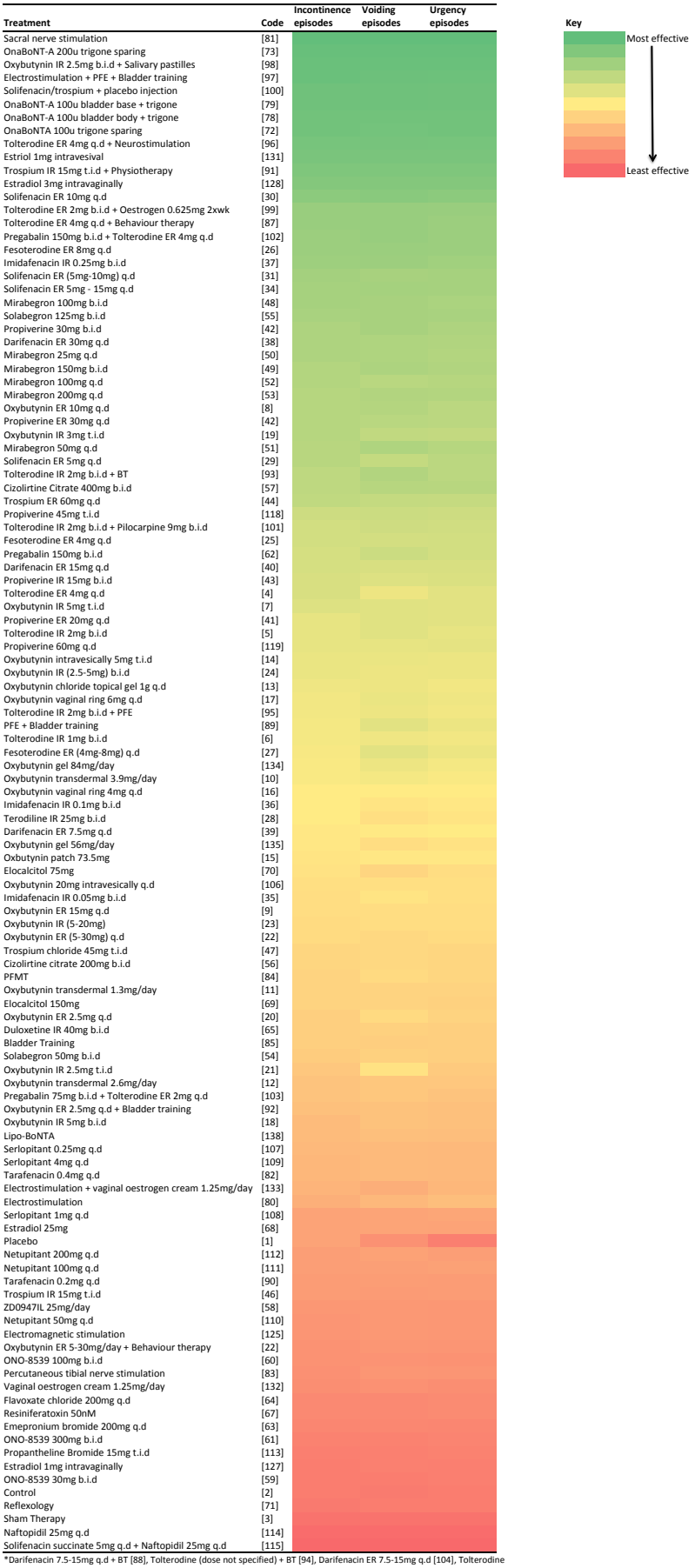
p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

8.4.2 Treatment profiles

Figure 8.5 illustrates the treatment profiles for all of the cardinal symptoms of OAB. Use of a multivariate hierarchical NMA allowed for the comparison of all interventions across all outcome measures, and thus completed the intervention profiles for efficacy outcomes. Generally, the treatment rankings were broadly similar to those obtained from univariate analyses (Figures 6.2 and 7.7). Sacral nerve stimulation appeared to be the most effective intervention across all three outcomes. Using multivariate analyses, estriol 1mg intravesically appeared to be amongst the top ten interventions (Figure 8.5). Previously, estriol 1mg intravesically was only evaluated for voiding and nocturia outcomes, and ranked amongst the top interventions for both outcomes (see Chapters 6 and 7). Borrowing information between outcomes allowed for estimation of treatment effects for both urinary incontinence and urgency episodes, and consequently estriol 1mg intravesically was ranked in tenth place across all outcome measures with a mean rank of 10 (95%CrI: 2,88) for all outcomes (Table 8.2). However, the remaining top ten interventions remained unchanged.

FIGURE 8.5: Heatmap of intervention profiles for the cardinal symptoms of OAB



*Darifenacin 7.5-15mg q.d + BT [88], Tolterodine (dose not specified) + BT [94], Darifenacin ER 7.5-15mg q.d [104], Tolterodine (dose not specified) [105], Lidocaine gel 6ml [129], Emepronium bromide IR 200mg [130], Sacral nerve stimulation+Tolterodine ER 2mg q.d [136], Tolterodine ER 2mg q.d [137] were disconnected from the network

8.4.3 Model assessment

8.4.3.1 Assessing inconsistency between direct and indirect information

Potentially inconsistent studies identified from individual-intervention NMA in Chapter 6 were assessed for their impact on the results obtained from multivariate hierarchical NMAs. Overall, removing potentially inconsistent studies had very little impact on the treatment rankings and clinical conclusions obtained from the multivariate hierarchical NMA. Results obtained from removing studies A157 (Wang et al., 2006) and A158 (Wang et al., 2009), respectively, are given in Table E.2 of Appendix E. In univariate analyses, these studies were identified as potentially contributing to inconsistent estimates of placebo versus electrostimulation for voiding frequency (see Section 6.4.4.1), and previously highlighted as potentially biased studies for urgency outcomes (see Section 5.7 for further details). However, in the univariate analyses of urgency, it was not possible to infer whether these studies contributed to inconsistencies in the network of evidence, as the interventions evaluated by studies A157 and A158 did not belong to a closed loop in the network, and thus it was not possible to obtain indirect estimates. In multivariate analyses, removing these studies had very little impact on the treatment effect estimates and clinical interpretation for all three outcomes of interest (Table E.2 of Appendix E).

8.4.3.2 Convergence diagnostics

Convergence diagnostic plots for a small selection of basic parameters, d_{1j}, the pooled effect estimate of intervention j relative to placebo, are given in Section E.3 of Chapter 8. Figure E.1 illustrates the Brooks-Gelman-Rubin plots. Both the between, G, and within chain variability, B, appeared to reach stability, and the ratio, R, appeared to converge to 1. Autocorrelation plots are given in Figure E.2. All plots appeared to show reducing autocorrelation with increased lag. History and trace plots are presented in Figure E.3, and took the appearance of random

noise, with no obvious difference between multiple MCMC chains with very different starting values. Density plots are given in Figure E.4. All parameters appeared to have a characteristically bell-shaped appearance resembling a normal distribution. Overall all diagnostic plots appeared to suggest that there was no evidence of non-convergence and both d_{1j} and ζ were estimated from samples which appeared to have a reasonable degree of mixing of the chains.

8.4.3.3 Sensitivity analysis

Sensitivity analyses assessing different choices of prior distributions for the between-study standard deviations, $\boldsymbol{V^{1/2}}$, and the variability of treatment effect profiles across outcomes, ζ, are given in Tables E.3 to E.5 of Appendix E. Table E.3 provides the treatment effect estimates for urinary incontinence episodes, Table E.4 provides the treatment effect estimates for voiding frequency, and Table E.5 provides the treatment effect estimates for urgency episodes. Different choices of prior distributions for variance parameters had very little impact in the treatment effect estimates for all three outcomes, and therefore the overall clinical decisions regarding interventions effectiveness remained the same. Thus suggesting that treatment effect estimates were fairly robust regardless of choice of prior distribution.

8.5 Discussion

The methods described in this chapter were an extension of the hierarchical NMA models, presented in Chapter 7, with which a multiple outcome framework was incorporated. This approach makes use of the correlations between multiple outcomes in order to predict and impute treatment effect estimates for missing data, and therefore, has the potential to limit the impact of outcome reporting bias. Furthermore, a multivariate hierarchical approach borrows strength across outcomes, classes of treatments, and studies, which has the potential to increase precision in the treatment effect estimates.

In the OAB example, the datasets used in univariate analyses for each outcome independently, included 115, 119 and 60 studies evaluating 97, 100, and 54 interventions for incontinence, voiding, and urgency episodes, respectively (see Chapter 5 for further information). Despite urgency being documented as "the cardinal symptom" of OAB (Cardozo et al., 2009), it was sufficiently under-reported in the original trials, and consequently fewer interventions were able to be evaluated in univariate analyses. Adopting a multivariate hierarchical NMA borrowed information across outcomes and consequently included 143 studies evaluating all 115 interventions for the management of incontinence, voiding and urgency. Using this methodology completed the treatment profiles for all prominent symptoms of OAB, which in turn allows decision makers to make inferences regarding the potential treatment benefit for all interventions, across all salient outcomes. In doing so, SNS appeared to be the most effective intervention for reducing incontinence, voiding and urgency episodes with an estimated posterior median reduction of -8 (95%CrI: -9.54, -6.27), -8.19 (95%CrI: -9.69,-6.49) and -8.49(95%CrI: -10.11,-6.78) episodes, relative to placebo, respectively. However, due to the limited number of studies and participants in which SNS was evaluated, these results should be interpreted with caution.

Sacral nerve stimulation was not evaluated for urgency in the original trials and thus could not be assessed in univariate analyses. Using a multivariate approach therefore changed the overall clinical decision for the management of urgency, where electrostimulation in combination with vaginal oestrogen cream 1.25mg/day was found to be the most effective intervention from univariate analyses (Chapters 6 and 7). Similarly, from univariate analyses, estriol 1mg intravesically appeared to be a promising intervention for voiding and nocturia, but there was no data for all other outcomes. Using a multivariate approach, estriol 1mg intravesically ranked in the top 10 interventions for all three cardinal symptoms of OAB.

A key assumption of MVNMA is that all data are assumed to be missing at random (Jackson et al., 2011), which in the case of the OAB example may not be

plausible (Globerman and Robert, 2015). It is likely that there is an element of selective reporting in the original trials, where outcomes with which interventions perform particularly well are more likely to be reported (Chan et al., 2004). In this situation, treatment effects may be exaggerated (Jackson et al., 2011), though there is an argument to suggest that in certain circumstances, a multivariate meta-analysis can lead to a more appropriate estimate of treatment effect in the presence of outcome reporting bias (Kirkham et al., 2012). In order to obtain more accurate estimates of treatment effects for decision making, data are needed for all interventions, across all outcome measures. Following the Core Outcome Measures in Effectiveness Trials (COMET) initiative (Williamson et al., 2011), there is a clear need to define a core outcome set (COS) for the future reporting of OAB trials (Globerman and Robert, 2015).

To ameliorate the impact of outcome reporting bias, the correlation between outcomes were used to obtain a predictive value for missing data. In order to achieve this, an assumption of constant relative effectiveness across outcomes was assumed for the basic parameters of the pooled treatment effect estimates, $d_{(1j)p}$ and $d_{(1_1 j_m)p}$, as described in Equations (8.5) and (8.11) for MVNMA and multivariate hierarchical NMA, respectively. If interest lies in the difference between active interventions, this may be a strong assumption as the outcome-specific effect, γ_p, will cancel. For example, the relative treatment effects of intervention A relative to intervention B obtained from MVNMA are expressed in terms of the basic parameters such that $d_{AB} = (d_{(1B)p} - d_{(1A)p}) \sim \text{Normal}(\alpha_B - \alpha_A, 2\zeta^2)$. Thus, in these situations, alternative methods should be explored. An assumption of constant relative effectiveness across outcomes is of less importance if interest lies in the relative rankings of the interventions, as these are calculated based on the basic parameters, $d_{(1j)p}$ (see Section 3.4.5.2 for further details).

In this example, the outcome-specific effect, γ_p, was assumed to be constant across all interventions. Further work could extend this model to incorporate the exchangeability of outcome-specific effects within classes of interventions, such that

$\gamma_{p_m} \sim$Normal(κ_m, Ψ^2_m), where κ_m denotes the pooled outcome effect for the m^{th} class of interventions, and Ψ^2_m denotes the class-specific between intervention variance. However this approach is likely to substantially increase the number of parameters to be estimated in the model and could lead to computational difficulties. Furthermore, if there is evidence to suggest that there is a disparity in outcome-specific effects between classes, the assumption of homogeneous between-study correlations is also unlikely to be satisfied and alternative parameterisations of the between-study covariance matrix would need to be considered. This could also be a potential area of further work.

One limitation of implementing MVNMA in WinBUGS, is the difficultly in calculating deviance statistics for the assessment of model fit and comparison. Residual deviance could be calculated by monitoring the estimated true treatment effects, $\boldsymbol{\theta_{ij}}$, for each study $i = 1, ..., n_s$ and intervention $j = 1, ..., n_t$, and calculating the difference of these true treatment effects relative to the observed treatment effects, $\boldsymbol{Y_{ij}}$, using the following equation: total residual deviance $= \sum_{i=1,j=1}^{i=n_s,j=n_t} (\boldsymbol{Y_{ij}} - \boldsymbol{\theta_{ij}})^2 \boldsymbol{S_{ij}}$, where $\boldsymbol{S_{ij}}$ denotes the treatment-specific within-study covariance matrix. As mentioned in Section 8.2, the within-study correlations were incorporated at the treatment-arm level in order to appropriately account for multi-arm trials (Achana et al., 2014), thus the within-study model was parametrised at the arm-level whereas the between-study model, with which $\boldsymbol{\theta_{ij}}$ is estimated, was parametrised at the study-level. This makes calculation of residual deviances more difficult, especially in the presence of large amounts of missing data. In the OAB example, there were a large number of observed data-points with intermittently missing data, therefore, such an approach would be both computational-, and time-intensive. Further work is needed to explore re-parameterisation of MVNMA models, and alternative methods in order to adequately assess model fit and comparison.

Chapters 6 and 7 assessed inconsistency between direct and indirect information for each of the outcomes individually, using a method of node-splitting (Dias et al.,

2010) . In this chapter, potentially inconsistent studies identified from univariate analyses were removed from the multivariate hierarchical NMA in order to assess their impact on the overall results. Removing potentially inconsistent studies had very little impact on the treatment effect estimates and clinical conclusions obtained from multivariate analyses. However, consistency of direct and indirect information obtained from MVNMAs were partially assessed due to the inability to generalise current node-splitting methods to the multivariate case. Further work is needed to extend this framework to incorporate a multivariate approach, and assess its use in detecting inconsistencies between direct and indirect evidence when information is borrowed across multiple outcomes. To do this, a simulation study could be used, implementing the methodology described in Section 4.5.2 in order to adequately capture the complexity of both the model and the data.

In this example, use of a multivariate hierarchical NMA was illustrated using the three cardinal symptoms of OAB (incontinence, voiding, and urgency). However, interest lies in both efficacy and safety outcomes. Incorporating additional outcomes results in an exponential increase in the number of parameters to be estimated in the model. This is particularly true for estimation of the parameters involved in both the within-study, and between-study covariance matrix. This substantial increase in the number of parameters can often result in computational difficulties for complex multivariate models such MVNMAs. Furthermore, estimating the within-study correlation structures can be particularly difficult for mixed (Bujkiewicz et al., 2013) and binary outcomes (Wei and Higgins, 2013). This is because an analytic solution is not possible (Wei and Higgins, 2013). A further limitation of using MVNMA for imputing missing data with mixed outcomes, is the assumption that intervention effects are exchangeable across outcomes. This assumption may not be reasonable if the outcomes differ in an important way, e.g, if the outcomes were measured on different scales. In the OAB example, binary outcomes were measured on a log-odds scale and continuous outcomes were measured on a median difference scale, therefore, intervention effect estimates will

differ in terms of the uncertainty with which they were estimated.

In this analysis, a homogeneity of correlations assumption was used to simplify the number of parameters in the model and to aid computation. In situations with which there are fewer interventions, and more information for each treatment comparison, it may be desirable to incorporate treatment-specific between-study correlations. It may also be desirable to incorporate treatment-specific within-study correlations, however, data for every pair of treatment comparisons may be difficult to obtain, and thus within-study correlations may be difficult to estimate. In this example, the within-study correlations were estimated from IPD obtained from the RELAX trial. Uncertainty in estimating the within-study correlations needs to be further accounted for. For continuous outcomes that follow a multivariate normal distribution, it would be possible to obtain estimates of the within-study correlations directly from the covariance matrix (Bujkiewicz et al., 2013). Estimates of the within-study correlations, together with their uncertainty, could be incorporated in to the MVNMA model by applying prior distributions to the within-study correlation parameters, pw_{xy}, using a bootstrapping method (Bujkiewicz et al., 2013).

8.6 Chapter summary

This chapter extends the hierarchical NMA framework presented in Chapter 7 to incorporate a multivariate approach. Accounting for the correlation between outcomes, multivariate hierarchical NMAs allowed for the evaluation of all interventions across all outcome measures. Including this additional information changed the overall clinical conclusions. Estriol 1mg intravesically was ranked in the top 10 interventions for the management of all cardinal symptoms of OAB, and SNS was found to be the most effective intervention for reducing urgency. Borrowing information across outcomes generally increased the precision in the treatment effect estimates, and relative treatment rankings. This precision was further increased

by incorporating a hierarchical structure where similarities between interventions that belong to the same class of interventions were accounted for. Multivariate hierarchical NMAs have proven to be a useful methodology, which have the potential to aid decision making in situations with which interest lies in the complete evaluation of multi-morbid, or syndromic conditions.

Chapter 9

Discussion and Conclusions

9.1 Summary

In the last 25 years, evidence-based medicine has become a paradigm of healthcare decision making and reimbursement in the UK. At the outset, an evidence-based approach set out to ensure that clinical practice was established from empirical evidence in order to achieve safer, more effective, and consistent medical care (Guyatt et al., 1992). At the heart of all evidence-based medicine, however, are the patients and the public, whose informed choices play an important role in healthcare delivery and service. Inevitably, an evidence-based approach mirrors the process of evidence creation, and usually begins with the analysis of well-conducted clinical trials. These trials are then systematically collated, and the results are summarised using evidence synthesis methods. In more recent years, evidence-based medicine has seen a rapid development in statistical methodologies, and evolved to embrace a wider range of disciplines. This thesis adds to this established body of literature and has developed and applied novel methodologies starting with the analysis of original evidence, such as RCTs, through to the comprehensive synthesis of clinical trial data. The methodological developments presented in this thesis were primarily motivated by the challenges of assessing healthcare interventions for OAB; however, these approaches can be applied to

different clinical settings facing the same methodological difficulties. Therefore, the findings of this thesis are of both clinical and methodological importance.

9.1.1 Methodological summary

This thesis presented a novel application of Bayesian methodology to evaluate real, and complex clinical scenarios. Motivated by the RELAX trial (Tincello et al., 2012), interest was in the evaluation of repeat treatment of onaBoNT-A in patients with interval censored data. Using a Bayesian flexible parametric frailty model allowed for computation of predictive posterior distributions in which to sample unreported event times, whilst fully capturing the complex underlying hazard rate. Simulation studies found that Bayesian prediction models generally performed well with up to 50% of interval censored data when an appropriate distributional form was selected. To my knowledge, this is the first study to adopt a fully Bayesian framework to model interval censored data using poly-Weibull models.

Synthesising trial data in individual-intervention NMA resulted in considerable uncertainty in the treatment effect estimates and relative treatment rankings. Collapsing interventions in to endonodal treatment classes increased the precision in the treatment effect estimates but restricted the interpretability of the individual interventions. Both of these approaches can have important implications for healthcare decision making, as it is difficult to infer the most efficacious intervention for the management of OAB. The development, and use of, hierarchical NMA models incorporating dose-response constraints sufficiently increased the precision in treatment effect estimates, without hindering the interpretability of individual interventions (Owen et al., 2015). In the presence of treatment effect modifiers, such as age, exchangeable treatment-by-covariate interactions within classes may be of particular use in situations where families of interventions differ in terms of the class-by-covariate interaction. This approach may be of particular use to get an insight in to class-by-covariate interactions when generalising healthcare

interventions to broader patient populations.

Extending the hierarchical model to incorporate a multivariate approach allowed for the comparison of all interventions across the complete symptom syndrome. This approach helped to ameliorate the impact of outcome reporting bias in the original trials, and borrowing information between outcomes increased the precision in the treatment effect estimates by up to 160% compared to that of univariate hierarchical NMAs.

9.1.2 Clinical summary

This thesis commenced by exploring whether there was a potential waning effect of repeated injections of onaBoNT-A in a large RCT in OAB - the RELAX trial (Tincello et al., 2012) (Chapter 4). The data represented the largest cohort of patients receiving up to three active injections of onaBoNT-A, and were analysed using novel and robust statistical methods to appropriately account for interval censored data. To my knowledge, this is the first analysis in the OAB literature to use time-to-symptom-recurrence as a measure of treatment effect for repeatedly administered onaBoNT-A injections (Owen et al., 2016a). There appeared to be a slow cumulative effect after second and third active injections of onaBoNT-A, but this difference was not of clinical or patient importance. This result adds to findings in the current OAB literature, which suggests that there is no difference in clinical efficacy of repeated injections of onaBoNT-A based on symptom profiles obtained from smaller samples of OAB patients (Sahai et al., 2010; Gousse et al., 2011; Dowson et al., 2012; Granese et al., 2012). However, there did appear to be a potential selection effect and extended placebo-effect from the open-label extension study, which to my knowledge, has not been reported in studies of OAB patients to date. This finding may have severe implications on future decision making for OAB as many drug studies have a pooled open-label extension with which to generate additional support for product licensing and reimbursement.

In the current literature, there is no evidence of a coherent comparison between interventions of different treatment modalities for the management of OAB. Consequently, there is little evidence of a superior, preventative intervention. This thesis reports the largest and most comprehensive comparison of interventions for the management of OAB to date. To my knowledge, this is the first analysis to compare all conservative and minimally invasive interventions for the management of OAB in a single coherent analysis of the most salient outcomes - clinical effectiveness, safety, and tolerability. In these analyses, SNS appeared to be the most effective intervention for reducing urinary incontinence and voiding frequency, with effect sizes much larger than the next ranked interventions. Electrostimulation and vaginal oestrogen cream, as a combination therapy, appeared to be the most effective intervention for minimising urgency episodes, with each of the individual interventions ranking in second and third place (Owen et al., 2017). Across the entire symptom profile, the top 5 most effective treatments included different doses and injection regimes of SNS, electrostimulation, and onaBoNT-A. In terms of oral medication, antimuscarinics barely featured in the most effective interventions, however, newly emerging β_3-agonists appeared to have equivalent efficacy but with an improved safety and tolerability profile.

One of the data limitations identified in this thesis was that of selective reporting in the original trials, which led to different interventions being evaluated for different outcomes. To ameliorate the impact of outcome reporting bias on healthcare decision making, a multivariate approach was incorporated (Kirkham et al., 2012). In doing so, the overall clinical conclusion as to which intervention was the most effective, changed for the management of urgency episodes. In multivariate analyses, SNS appeared to the be the most effective intervention for all three cardinal symptoms of OAB (urinary incontinence, voiding, and urgency) and estriol 1mg intravesically ranked amongst the top 10 interventions for the entire symptom profile.

9.2 Discussion and limitations

This thesis presents the largest, most comprehensive, cross-modality treatment comparison for the management of OAB to date. With plans to update the NICE clinical guidance on urinary incontinence in the near future (National Institute for Health and Care Excellence, 2013b), it is anticipated that this work could be used to further evidence-based practice in the context of OAB. The methodologies developed as part of this thesis contribute to a growing body of methodological literature in the evolving area of HTA. Whilst this thesis focuses on the implementation of these methods in OAB research, the models can be easily generalisable to alternative clinical areas posing the same methodological concerns. Each of the individual chapters of this thesis concluded with a detailed discussion, highlighting, where appropriate, the specific methodological and clinical limitations of each chapter. The following section gives an overview of the broader discussion points encompassing all analyses described in this thesis.

9.2.1 Methodological discussion

One of the overarching limitations of the analyses presented in this thesis is that of data availability and quality. Starting with initial analyses of RCT data described in Chapter 4 through to evidence synthesis methods presented in Chapter 8, availability and quality of data appeared to be a challenge throughout the HTA process. This thesis aimed to address many of these challenges by developing novel methodologies to minimise the impact of sparse, and selectively reported data. However, data quality still posed several limitations. Chapter 4 developed and applied Bayesian poly-Weibull models to analyse a large RCT of patients with interval censored data. In this chapter, one of the limitations in terms of model computation, was that of sparsity of known events. As a general rule of thumb, at least 10 events are required for every parameter to be estimated in the model (Concato et al., 1995; Peduzzi et al., 1995, 1996). Consequently, in the example presented in this thesis, the limited number of known events hindered the

sophistication of the Bayesian prediction models. The availability of data was a further limitation throughout the evidence synthesis analyses described in Chapters 6 and 7, for which, selective reporting in the original trials appeared to be of most concern. In this example, outcome reporting bias in the original trials had a substantial impact on the ability of healthcare decision making, as many of the intervention profiles were incomplete across the entire symptom syndrome. This finding motivated the development of multivariate hierarchical NMAs described in Chapter 8. Whilst these models help to ameliorate the impact of outcome reporting bias, they do not eliminate the possibility of exaggerated treatment effects (Jackson et al., 2011). For this reason, further data is required for intervention effects across all outcome measures.

Use of data synthesis methods raise a number of concerns, one of which is the quality of included data (Jansen et al., 2011). Sources of bias, both in terms of publication bias, and study quality were further explored in Chapter 5. Sensitivity analyses assessed the impact of potentially biased studies on the pooled treatment effect estimates obtained from NMAs presented in Chapters 6, 7, and 8. Removing potentially biased studies, A157 (Wang et al., 2006) and A158 (Wang et al., 2009), changed the overall clinical conclusion for effect estimates obtained from univariate analyses of urgency; and thus, these effect estimates should be interpreted with caution. An alternative, potentially more useful approach than individually removing studies in the form of a sensitivity analysis, is to incorporate the various sources of bias by down-weighting potentially biased studies (Turner et al., 2009).

The analyses presented in this thesis evaluated a broad range of interventions for the management of OAB; some of which included interventions and doses that are not currently licensed (e.g. ZD0947IL 25mg/day). Arguably these interventions will not be considered by regulatory bodies such as NICE, and thus this raises the question of whether these particular interventions should be included in network meta-analyses. There is much controversy in the current network meta-analysis

literature over the design of the most efficient evidence space in considering all relevant indirect information (Hawkins et al., 2009a,b; Hoaglin et al., 2011; Jansen et al., 2011; Sturtz and Bender, 2012; Dequen et al., 2014; Caldwell, 2014). In a paper by Hawkins et al. (2009a), the authors proposed an iterative search strategy to more efficiently identify indirect information. In this paper, the authors make use of a staged approach, whereby the interventions of key interest are referred to as primary comparators. Interventions that have directly been compared with primary comparators are referred to as secondary comparators, and interventions that have been compared to secondary comparators are referred to as tertiary comparators. Extending the network in this way may connect interventions that may otherwise be disconnected from the network, and thus may provide valuable indirect information. However, including additional interventions may introduce inconsistency between direct and indirect information, as well as, increase between-study heterogeneity and uncertainty in the treatment effect estimates (Hawkins et al., 2009a,b; Jansen et al., 2011; Hoaglin et al., 2011). It is clear that there remains a trade-off between defining the evidence- and decision-space (Sturtz and Bender, 2012). In the OAB example, all tertiary interventions were included in the analyses and thus, interventions and doses that are not currently licensed were included. In this particular example, including doses that are not currently licensed added valuable information to the networks of evidence. This is particularly true for onaBoNT-A, as initially, onaBoNT-A was administered at 200 units for the management of OAB, and thus there is a vast amount of trial information solely evaluating onaBoNT-A 200U. In more recent guidance (National Institute for Health and Care Excellence, 2013b), onaBoNT-A is recommended at 100 units. In Chapters 7 and 8, including all interventions allowed information to be borrowed across alternative doses of onaBoNT-A using partial exchangeability and dose-response constraints. Using this methodology increased the precision in the treatment effect estimates for all formulations and doses of onaBoNT-A.

From a HTA perspective, interventions must form discrete classifications in order to assign costs. Therefore, interpretation of individual intervention effects are

of utmost importance. Chapter 7 developed a three-level hierarchical NMA incorporating dose-response constraints, with the aim to increase precision in the treatment effect estimates whilst maintaining the interpretability of individual intervention effects. Hierarchical NMA models make a fundamental assumption that the intervention effects, within families of interventions, are exchangeable, and a judgement of the suitability of this assumption is required. If this assumption is not satisfied, use of hierarchical models may introduce inaccurate results. A further limitation of hierarchical NMA models is the subjective classification of interventions in to broader families of interventions. This is especially true for interventions that span more than one family of interventions, such as combination therapies. In this situation, a potential synergistic or sub-additive relationship between the individual interventions could be further explored. Use of hierarchical models in such settings also raises the issue of the ordering of the hierarchy (Walsh and Mengersen, 2012). For example, in the OAB example presented in this thesis, interventions could be grouped first by formulation, such as immediate release (IR) and extended release (ER), and then by dose, or vice versa.

Throughout the analyses presented in this thesis, a number of strong assumptions were made, namely, the assumptions of additivity, consistency, and exchangeability. As with all other analyses, these assumptions should be validated. Throughout this thesis, the assumptions of additivity and consistency were assessed, where possible, using node-splitting methods. However, additivity and consistency could not be assessed for a number of treatment comparisons as they did not form closed loops in the networks of evidence (see Section 3.4.5.3), and should therefore be interpreted with caution. Assumptions of additivity and consistency are of most concern if there appears to be outlying studies with large treatment effects (Mills et al., 2013b). The impact of outlying studies should be examined in a series of sensitivity analyses. In the example presented in this thesis, removing outlying studies with large treatment effects had no impact on the remaining treatment effect estimates. Furthermore, the analyses described in Chapter 6 make a fundamental assumption that treatment effects are exchangeable. Assuming exchangeability

of treatment effects across treatments with different doses and different levels of invasiveness may not be reasonable. Chapter 7 relaxes this assumption and assumes partial exchangeability of treatment effects, and incorporates dose-response constraints on increasing doses of an intervention. In this chapter, variance components are compartmentalised to families of treatments, and constraints are placed on increasing doses. However, for SNS there is only one study evaluating each of the urinary incontinence and voiding outcomes, and thus estimates of the variance parameter for this class of intervention will largely be driven by the prior distribution. It is therefore crucial to undertake sensitivity analyses assessing the impact of the choice of prior distribution for variance parameters.

9.2.2 Clinical discussion

The results of this thesis show that SNS, electrostimulation, and onaBoNT-A are promising interventions for the management of OAB, though these interventions vary in terms of clinical effectiveness, cost, and level of invasiveness. The analyses and results presented in this thesis take no account of cost or the natural treatment pathway outlined by the National Institute for Health and Care Excellence (2013b). This thesis aimed to aid decision making in OAB from a patient, physician's, and health policy decision maker's perspective. Particularly from a patient and physician's perspective, it is useful to understand the expected patient benefit of an intervention at the detriment of an increased burden elicited from a more invasive treatment. Therefore, all interventions were included in this thesis regardless of where they fall in the natural treatment pathway (National Institute for Health and Care Excellence, 2013b). Treatment pathways could be seen as a classification of interventions, with which interventions belonging to each stage in the treatment pathway may be considered exchangeable; and therefore, treatment pathways could be incorporated in to the hierarchical model as an additional level in the hierarchy.

Whilst this thesis endeavoured to consider, and adjust for, potential treatment effect modifiers such as age, and baseline severity, through the exploration of prediction models and network meta-regression analyses, an element of between study-heterogeneity remained unexplained. Alternative study-level characteristics such as proportion of female participants, and publication date, may further be assessed as potential reasons for between-study heterogeneity that is more than that attributable to chance alone. If IPD were available, to increase the power in detecting, and adjusting for potential treatment effect modifiers, individual patient characteristics could be incorporated in to the random effects network meta-analysis framework (Riley et al., 2010). In the presence of IPD, potential subgroup effects could further be investigated using subgroup analyses to identify patient characteristics for which interventions work particularly well. With an emphasis on personalised medicine, use of subgroup analyses could provide information on treatment efficacy for targeted interventions (Hamburg and Collins, 2010).

One of the clinical limitations of the work presented in this thesis was the generalisability of the findings to the male population. Many of the published trials in idiopathic DO and OAB evaluate interventions in a mixed population of both males and females, though almost all of the trials were predominantly female. In the systematic literature review of RCTs in OAB described in Chapter 5, the average proportion of females in each of the trial populations was 84.9% (SD: 16.8). The evaluation of interventions predominantly in female populations is a consequence of the prevalence of OAB in this subgroup. As a result, many of the interventions were designed for use in the female population only. This can make generalisability of treatment recommendations particularly difficult for the male population. However, there is not enough available evidence to evaluate the effect of interventions in men with idiopathic DO and OAB alone. There is, however, a large body of literature documenting intervention effects for men with OAB due to benign prostatic hypertrophy and concomitant bladder outlet obstruction (Dmochowski and Gomelsky, 2009). In the presence of trial data in a male only population with idiopathic DO and OAB, it may be useful to use meta-regression

techniques to establish whether there is a difference in the treatment-by-covariate interaction for males versus females. This was not possible in the example presented in this thesis, as there were only two studies (Burgio et al., 2011; Orešković et al., 2012) evaluating interventions in males only. As explained above, proportion of females in the patient population may be incorporated in to the network meta-regression models to help explain heterogeneity. However, this would only give an indication of the impact of intervention effects on different proportions of male and female populations, and would not necessarily give a direct interpretation of treatment effects in males alone. As there is little variation in the proportion of females across studies, this approach may also lack power to detect an effect (Debray et al., 2015).

9.3 Further work

In this thesis, synthesising data in secondary analyses such as NMAs proved to be a difficult task in the presence of selective reporting in the original trials; and thus, one of the main findings of this thesis was the necessity to define core outcome sets for the future reporting of OAB trials. With the recent momentum of the COMET initiative (Williamson et al., 2011), core outcome sets for OAB have been developed during the writing of this thesis (International Consortium for Health Outcome Measurements, 2016). Further work, therefore, relies on the publication of future OAB trials reporting all of the core outcomes outlined by the International Consortium for Health Outcome Measurements (2016). These trials can then be incorporated in to multivariate analyses in order to strengthen the evidence base and consequently, increase the precision in the treatment effect estimates.

The scope of this thesis was restricted to RCTs in order to maintain randomisation in the decision-making process. Further information can be gained from alternative sources of evidence, such as single-armed trials and observational studies. In

this thesis, one study (Burgio et al., 2002) was excluded from the analyses as it did not present data on a comparator arm, and thus was considered a single-armed study (see Section 5.4.1 for further details). To strengthen the evidence base, single-armed studies could be incorporated in to the NMA models described in Chapters 6, 7 and 8. The NICE Decision Support Unit currently recommend the use of match-adjusted indirect comparisons for incorporating single-armed studies in evidence synthesis models (Phillippo et al., 2016). Matching aims to replicate randomisation by identifying matched controls who are similar, in terms of one or more individual characteristics, to the participants treated in the single-armed, or observational study. However, this approach breaks randomisation (Jansen et al., 2011), and as with all observational relationships, single-armed trials and observational studies may introduce bias which needs to be considered and accounted for in the NMA models (Schmitz et al., 2013).

The performance of univariate and multivariate hierarchical NMA models, described in Chapters 7 and 8, could be further investigated through a series of simulation studies (McCarron et al., 2011; Takwoingi et al., 2015) under different levels of individual-intervention (Pibouleau and Chevret, 2013) and class variability. In doing so, recommendations of scenarios for which hierarchical NMA models may outperform standard NMA models, in the context of healthcare decision making, may be quantified. It is anticipated that in situations where interventions do not greatly differ in effectiveness, a random effects NMA model may be an adequate model choice. A hierarchical NMA model may be of more use in situations where the difference between interventions within classes are modest, but there is a substantial difference in between class variability. Pibouleau and Chevret (2013) investigated the performance of pairwise hierarchical meta-analyses under different levels of variability between individual-interventions and found that hierarchical models were selected over standard meta-analyses for more than 95% of scenarios. Further work could extend these simulation studies to evaluate hierarchical models in a NMA setting, and further investigate the impact of differing levels of within class heterogeneity and between class variability. However, even in the presence of

simulation studies - which will quantify scenarios for which hierarchical NMAs will be of most use in terms of the optimal between and within class heterogeneity/ variability - in practice, hierarchical NMAs would need to be estimated in order to obtain such information; and thus, there is an argument to suggest that it may be more informative to use residual deviance and DIC statistics to assess model fit and choice of model (Spiegelhalter et al., 2002) for specific clinical examples if indeed hierarchical models can be estimated at all due to the sparsity of treatments and/or data.

Further methodological developments are required in order to assess inconsistencies between direct and indirect information, model fit and comparison, for MVN-MAs described in Chapter 8. This work would require consideration of model re-parametrisation and generalisability of node-splitting analyses outlined by Dias et al. (2010) to the multivariate setting. Certainly, for the example presented in this thesis, further work remains in the calculation of residual deviance statistics in order to assess model fit and suitably.

Building on the work outlined in Chapter 8, with which all cardinal symptoms of OAB were simultaneously analysed, the multivariate hierarchical NMA model could be extended to include safety and tolerability measures. However, this approach could be both computationally intensive and time-consuming as the number of parameters to be estimated in the model would exponentially increase with every additional outcome. In the example presented in this thesis, a further concern with incorporating binary outcomes is that the assumption underpinning the imputation technique for missing data - that is the assumption of exchangeable intervention effects across outcomes (Achana et al., 2014) - may not be satisfied due to differences in the measurement and scale of continuous and binary outcomes. In this context, there may be a benefit of borrowing information between safety and tolerability outcomes alone, using a trivariate hierarchical NMA. However, careful consideration of the estimation of within-study correlation matrices would

be necessary (Wei and Higgins, 2013).

Chapter 8 of this thesis endeavoured to obtain treatment effect estimates across the entire symptom syndrome of OAB using MVNMA, which allowed for comparison of all interventions across all outcome measures. With a large number of potential treatments, and a multitude of outcome measures, identifying the 'best' intervention overall can still be problematic for multi-morbid conditions. Further work could develop and compare methods to identify the most important clinical outcomes from the perspective of society and the patients. There are two potential methods that could be used for this purpose; the first method uses multi-criteria decision analysis (MCDA) to combine the utility of each individual outcome measure by inter-outcome comparison, and thus allows for the identification of the preferred overall measure (Thokala and Duenas, 2012). The second approach could evaluate the net-benefit, which aims to identify the overall health gain - which would result from the implementation of the assigned utility for each outcome - that would be required to justify the cost incurred (Sutton et al., 2005; Glasziou and Irwig, 1995; Laska et al., 1999). For both of these methods, elicitation of utilities would be required. This could be obtained from two possible approaches; the first, from a societal perspective on each of the outcome measures using EQ-5D data collected from the RELAX study (Tincello et al., 2012); and the second, could involve elicitation of the patient's perspective using time trade off analyses. Here, the patients would be presented with two scenarios: 1) a chronic symptom of OAB for an arbitrary time, t, or 2) a healthy condition for a shorter lifetime, $x < t$. Time x will vary until the patients are indifferent between the two scenarios, at which point a corresponding utility can be given (Drummond et al., 2015).

Throughout this thesis, a wider HTA perspective has been considered with the anticipation that the findings could be used to inform future cost-effectiveness analyses. There are a number of published economic decision models in the current literature which evaluated the cost-effectiveness of interventions for OAB from

a UK National Health Service perspective (Hughes and Dubois, 2004; Guest et al., 2004; Aballéa et al., 2015; Nazir et al., 2015b,a; Freemantle et al., 2015). All of these studies compared the cost-effectiveness of interventions of the same class and thus, to my knowledge, there is no economic decision model comparing interventions across multiple treatment modalities. Future work is needed to assess the cost-effectiveness of interventions for the management of OAB of varying levels of invasiveness and cost. To achieve this, a comprehensive economic decision model would need to be developed and implemented. Effect estimates from the multivariate hierarchical NMA described in Chapter 8, together with the elicitation of patient preferences and utilities, described above, could be used to populate a Markov model over a lifetime horizon. It is here that the effects of long term and repeat treatment, described in Chapter 4, would be of particular use.

Furthermore, value of information techniques could be used to estimate the expected increase in utility that added information would give to the cost-effectiveness decision (Eckermann and Willan, 2007). These methods would give an indication of the supporting evidence surrounding the overall cost-effective decision. Therefore, where evidence is particularly absent, the design and application of a future adaptive trial with (Chen and Willan, 2013) or without (Kairalla et al., 2012) the application of value of information techniques, could be directly implemented to identify the most clinically effective intervention overall.

9.4 Conclusions

In conclusion, this thesis has demonstrated the use of novel Bayesian methods to answer real clinical questions in OAB. The models developed as part of this thesis are widely generalisable to other research areas facing the same methodological challenges with missing, and sparse data. In the context of OAB, this thesis has enabled the evaluation of all conservative and minimally invasive interventions for

the management of OAB, and found that SNS appeared to be a promising intervention. Though, building on this body of work, a comprehensive cost-effectiveness analysis is still required. The work presented in this thesis has the ability to inform public healthcare policy, and aid decision making from a patient, physician's, and decision maker's perspective.

Bibliography

Aballéa, S., Maman, K., Thokagevistk, K., Nazir, J., Odeyemi, I. A., Hakimi, Z., Garnham, A., and Toumi, M. Cost effectiveness of mirabegron compared with tolterodine extended release for the treatment of adults with overactive bladder in the united kingdom. *Clinical drug investigation*, 35(2):83–93, 2015.

Abrams, K. R., Gillies, C. L., and Lambert, P. C. Meta-analysis of heterogeneously reported trials assessing change from baseline. *Statistics in medicine*, 24(24):3823–3844, 2005.

Abrams, P., Cardozo, L., Fall, M., Griffiths, D., Rosier, P., Ulmsten, U., Van Kerrebroeck, P., Victor, A., and Wein, A. Standardisation subcommittee of the international continence society. the standardisation of terminology of lower urinary tract function: report from the standardisation sub-committee of the international continence society. *Neurourol Urodyn*, 21(2):167–178, 2002.

Abrams, P. Describing bladder storage function: overactive bladder syndrome and detrusor overactivity. *Urology*, 62(5):28–37, 2003a.

Abrams, P. Describing bladder storage function: overactive bladder syndrome and detrusor overactivity. *Urology*, 62(5):28–37, 2003b.

Abrams, P. Urgency: the key to defining the overactive bladder. *BJU international*, 96(s1):1–3, 2005.

Abrams, P., Kelleher, C., Kerr, L. A., and Rogers, R. G. Overactive bladder significantly affects quality of life. *Am J Manag Care*, 6(11 Suppl):S580–S590, 2000.

Abrams, P., Avery, K., Gardener, N., Donovan, J., and Board, I. A. The international consultation on incontinence modular questionnaire: www. iciq. net. *The Journal of urology*, 175(3):1063–1066, 2006.

Abrams, P., Kelleher, C., Staskin, D., Rechberger, T., Kay, R., Martina, R., Newgreen, D., Paireddy, A., van Maanen, R., and Ridder, A. Combination treatment with mirabegron and solifenacin in patients with overactive bladder: efficacy and safety results from a randomised, double-blind, dose-ranging, phase 2 study (symphony). *Eur Urol*, 67(3):577–588, Mar 2015.

Achana, F. A., Cooper, N. J., Dias, S., Lu, G., Rice, S. J., Kendrick, D., and Sutton, A. J. Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. *Statistics in medicine*, 32(5):752–771, 2013.

Achana, F. A., Cooper, N. J., Bujkiewicz, S., Hubbard, S. J., Kendrick, D., Jones, D. R., and Sutton, A. J. Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. *BMC medical research methodology*, 14(1):1, 2014.

Ades, A., Mavranezouli, I., Dias, S., Welton, N. J., Whittington, C., and Kendall, T. Network meta-analysis with competing risk outcomes. *Value in Health*, 13(8):976–983, 2010.

Amaro, J. L., Gameiro, M. O., and Padovani, C. R. Effect of intravaginal electrical stimulation on pelvic floor muscle strength. *Int Urogynecol J Pelvic Floor Dysfunct*, 16(5):355–358, 2005.

Apostolidis, A., Dasgupta, P., and Fowler, C. J. Proposed mechanism for the efficacy of injected botulinum toxin in the treatment of human detrusor overactivity. *European urology*, 49(4):644–650, 2006.

Arends, L. R., Vokó, Z., and Stijnen, T. Combining multiple outcome measures in a meta-analysis: an application. *Statistics in medicine*, 22(8):1335–1353, 2003.

Avery, K., Donovan, J., Peters, T. J., Shaw, C., Gotoh, M., and Abrams, P. Iciq: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourology and urodynamics*, 23(4):322–330, 2004.

Awad, S., Al-Zahrani, H., Gajewski, J., and Bourque-Kehoe, A. Long-term results and complications of augmentation ileocystoplasty for idiopathic urge incontinence in women. *British journal of urology*, 81:569–573, 1998.

Banerjee, S. and Carlin, B. P. Parametric spatial cure rate models for interval-censored time-to-relapse data. *Biometrics*, 60(1):268–275, 2004.

Barnard, J., McCulloch, R., and Meng, X.-L. Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica*, pages 1281–1311, 2000.

Basu, M. and Duckett, J. R. Update on duloxetine for the management of stress urinary incontinence. *Clinical interventions in aging*, 4:25, 2009.

Battista, R. N. and Hodge, M. J. The evolving paradigm of health technology assessment: reflections for the millennium. *CMAJ: Canadian Medical Association Journal*, 160(10):1464, 1999.

Bayes, M. and Price, M. An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfrs. *Philosophical Transactions (1683-1775)*, pages 370–418, 1763.

Bellette, P. O., Rodrigues-Palma, P. C., Hermann, V., Riccetto, C., Bigozzi, M., and Olivares, J. M. Posterior tibial nerve stimulation in the management of overactive bladder: a prospective and controlled study. *Actas Urol Esp*, 33(1): 58–63, Jan 2009.

Berger, J. O. and Sun, D. Bayesian analysis for the poly-weibull distribution. *Journal of the American Statistical Association*, 88(424):1412–1418, 1993.

Berkey, C., Hoaglin, D., Antczak-Bouckoms, A., Mosteller, F., and Colditz, G. Meta-analysis of multiple outcomes by regression with random effects. *Statistics in medicine*, 17(22):2537–2550, 1998.

Berlin, J. A., Santanna, J., Schmid, C. H., Szczech, L. A., and Feldman, H. I. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statistics in Medicine*, 21(3):371–387, 2002. ISSN 1097-0258. doi: 10.1002/sim.1023. URL `http://dx.doi.org/10.1002/sim.1023`.

Bø, K. Pelvic floor muscle training is effective in treatment of female stress urinary incontinence, but how does it work? *International Urogynecology Journal*, 15(2):76–84, 2004.

Booth, A., Clarke, M., Ghersi, D., Moher, D., Petticrew, M., and Stewart, L. An international registry of systematic-review protocols. *The Lancet*, 377(9760): 108–109, 2011.

Bower, P., Macdonald, W., Harkness, E., Gask, L., Kendrick, T., Valderas, J. M., Dickens, C., Blakeman, T., and Sibbald, B. Multimorbidity, service organization and clinical decision making in primary care: a qualitative study. *Family practice*, 28(5):579–587, 2011.

Brazzelli, M., Murray, A., and Fraser, C. Efficacy and safety of sacral nerve stimulation for urinary urge incontinence: a systematic review. *The Journal of urology*, 175(3):835–841, 2006.

Briggs, A. H. et al. A bayesian approach to stochastic cost-effectiveness analysis. *Health economics*, 8(3):257–261, 1999.

Brooks, S. P. and Gelman, A. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*, 7(4): 434–455, 1998.

Buckley, B. S., Grant, A. M., Tincello, D. G., Wagg, A. S., and Firkins, L. Prioritizing research: patients, carers, and clinicians working together to

identify and prioritize important clinical uncertainties in urinary incontinence. *Neurourology and urodynamics*, 29(5):708–714, 2010.

Bujkiewicz, S., Thompson, J. R., Sutton, A. J., Cooper, N. J., Harrison, M. J., Symmons, D. P., and Abrams, K. R. Multivariate meta-analysis of mixed outcomes: a bayesian approach. *Statistics in Medicine*, 32(22):3926–3943, 2013.

Burgio, K. L., Locher, J. L., and Goode, P. S. Combined behavioral and drug therapy for urge incontinence in older women. *Journal of the American Geriatrics Society*, 48(4):370–374, 2000.

Burgio, K. L., Goode, P. S., Locher, J. L., Umlauf, M. G., Roth, D. L., Richter, H. E., Varner, R. E., and Lloyd, L. K. Behavioral training with and without biofeedback in the treatment of urge incontinence in older women: a randomized controlled trial. *JAMA*, 288(18):2293–2299, Nov 2002.

Burgio, K. L., Kraus, S. R., Menefee, S., Borello-France, D., Corton, M., Johnson, H. W., Mallett, V., Norton, P., FitzGerald, M. P., Dandreo, K. J., Richter, H. E., Rozanski, T., Albo, M., Zyczynski, H. M., Lemack, G. E., Chai, T. C., Khandwala, S., Baker, J., Brubaker, L., Stoddard, A. M., Goode, P. S., Nielsen-Omeis, B., Nager, C. W., Kenton, K., Tennstedt, S. L., Kusek, J. W., Chang, T. D., Nyberg, L. M., Steers, W., and , U. I. T. N. Behavioral therapy to enable women with urge incontinence to discontinue drug treatment: a randomized trial. *Ann Intern Med*, 149(3):161–169, Aug 2008.

Burgio, K. L., Goode, P. S., Richter, H. E., Markland, A. D., Johnson, T. M., and Redden, D. T. Combined behavioral and individualized drug therapy versus individualized drug therapy alone for urge urinary incontinence in women. *The Journal of urology*, 184(2):598–603, 2010.

Burgio, K. L., Goode, P. S., Johnson, T. M., Hammontree, L., Ouslander, J. G., Markland, A. D., Colli, J., Vaughan, C. P., and Redden, D. T. Behavioral versus drug treatment for overactive bladder in men: the male overactive

bladder treatment in veterans (motive) trial. *Journal of the American Geriatrics Society*, 59(12):2209–2216, 2011.

Burton, A., Altman, D. G., Royston, P., and Holder, R. L. The design of simulation studies in medical statistics. *Stat Med*, 25(24):4279–4292, Dec 2006.

Buser, N., Ivic, S., Kessler, T. M., Kessels, A. G., and Bachmann, L. M. Efficacy and adverse events of antimuscarinics for treating overactive bladder: network meta-analyses. *European urology*, 62(6):1040–1060, 2012.

But, I. Conservative treatment of female urinary incontinence with functional magnetic stimulation. *Urology*, 61(3):558–561, 2003.

Caldwell, D. M. An overview of conducting systematic reviews with network meta-analysis. *Systematic reviews*, 3(1):109, 2014.

Caldwell, D. M., Ades, A., and Higgins, J. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *British Medical Journal*, 7521:897, 2005.

Calle, M. L. and Gómez, G. Nonparametric bayesian estimation from interval-censored data using monte carlo methods. *Journal of Statistical Planning and Inference*, 98(1):73–87, 2001.

Cao, Y., Lv, J., Zhao, C., Li, J., and Leng, J. Cholinergic antagonists combined with electrical stimulation or bladder training treatments for overactive bladder in female adults: A meta-analysis of randomized controlled trials. *Clinical drug investigation*, 36(10):801–808, 2016.

Cardozo, L., Lisec, M., Millard, R., KUZMIN, I., Drogendijk, T., HUANG, M., Ridder, A., et al. Randomized, double-blind placebo controlled trial of the once daily antimuscarinic agent solifenacin succinate in patients with overactive bladder. *The Journal of urology*, 172(5):1919–1924, 2004.

Cardozo, L., Chapple, C., and Wein, A. Urgency as the cardinal symptom of overactive bladder: a critical analysis. *World journal of urology*, 27(6): 701–703, 2009.

Caremel, R., Loutochin, O., and Corcos, J. What do we know and not know about mirabegron, a novel $\beta 3$ agonist, in the treatment of overactive bladder? *International urogynecology journal*, 25(2):165–170, 2014.

Carlin, B. P. and Louis, T. A. Bayes and empirical bayes methods for data analysis. *Statistics and Computing*, 7(2):153–154, 1997.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. Stan: A probabilistic programming language. *J Stat Softw*, 2016.

Celeux, G., Forbes, F., Robert, C. P., Titterington, D. M., et al. Deviance information criteria for missing data models. *Bayesian analysis*, 1(4):651–673, 2006.

Chaimani, A., Higgins, J. P., Mavridis, D., Spyridonos, P., and Salanti, G. Graphical tools for network meta-analysis in stata. *PloS one*, 8(10):e76654, 2013.

Chan, A.-W., Hróbjartsson, A., Haahr, M. T., Gøtzsche, P. C., and Altman, D. G. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *Jama*, 291(20):2457–2465, 2004.

Chancellor, M., Kianifard, F., Beamer, E., Mongay, L., Ebinger, U., Hicks, G., and DelConte, A. A comparison of the efficacy of darifenacin alone vs. darifenacin plus a behavioural modification programme upon the symptoms of overactive bladder. *International journal of clinical practice*, 62(4):606–613, 2008.

Chancellor, M. B. and Chartier-Kastler, E. J. Principles of sacral nerve stimulation (sns) for the treatment of bladder and urethral sphincter dysfunctions. *Neuromodulation: Technology at the Neural Interface*, 3(1): 16–26, 2000.

Chapple, C. R., Khullar, V., Gabriel, Z., Muston, D., Bitoun, C. E., and Weinstein, D. The effects of antimuscarinic treatments in overactive bladder: an update of a systematic review and meta-analysis. *European urology*, 54(3): 543–562, 2008a.

Chapple, C. R., Khullar, V., Gabriel, Z., Muston, D., Bitoun, C. E., and Weinstein, D. The effects of antimuscarinic treatments in overactive bladder: an update of a systematic review and meta-analysis. *European urology*, 54(3): 543–562, 2008b.

Chen, M. H. and Willan, A. R. Determining optimal sample sizes for multistage adaptive randomized clinical trials from an industry perspective using value of information methods. *Clinical Trials*, 10(1):54–62, 2013.

Chen, M.-H. and Shao, Q.-M. Monte carlo estimation of bayesian credible and hpd intervals. *Journal of Computational and Graphical Statistics*, 8(1):69–92, 1999.

Choo, M.-S., Ku, J. H., Lee, J. B., Lee, D. H., Kim, J. C., Kim, H.-J., Lee, J. J., and Park, W.-H. Cross-cultural differences for adapting overactive bladder symptoms: results of an epidemiologic survey in korea. *World journal of urology*, 25(5):505–511, 2007.

Chowdhry, A. K., Dworkin, R. H., and McDermott, M. P. Meta-analysis with missing study-level sample variance data. *Statistics in medicine*, 2016.

Clayton, D. Some approaches to the analysis of recurrent event data. *Statistical Methods in Medical Research*, 3(3):244–262, 1994.

Collins, L., Schafer, J., and Kam, C. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol Methods*, 6(4):330–351, Dec 2001.

Concato, J., Peduzzi, P., Holford, T. R., and Feinstein, A. R. Importance of events per independent variable in proportional hazards analysis i.

background, goals, and general strategy. *Journal of clinical epidemiology*, 48 (12):1495–1501, 1995.

Cooper, N. J., Sutton, A. J., Morris, D., Ades, A., and Welton, N. J. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Statistics in medicine*, 28(14):1861–1881, 2009.

Corcos, J. and Schick, E. Prevalence of overactive bladder and incontinence in canada. *The Canadian journal of urology*, 11(3):2278–2284, 2004.

Coyne, K., Revicki, D., Hunt, T., Corey, R., Stewart, W., Bentkover, J., Kurth, H., and Abrams, P. Psychometric validation of an overactive bladder symptom and health-related quality of life questionnaire: the oab-q. *Quality of life research*, 11(6):563–574, 2002.

Coyne, K. S., Matza, L. S., Kopp, Z., and Abrams, P. The validation of the patient perception of bladder condition (ppbc): a single-item global measure for patients with overactive bladder. *European urology*, 49(6):1079–1086, 2006.

Coyne, K. S., Margolis, M. K., Jumadilova, Z., Bavendam, T., Mueller, E., and Rogers, R. Overactive bladder and women's sexual health: What is the impact? *The journal of sexual medicine*, 4(3):656–666, 2007.

Coyne, K. S., Sexton, C. C., Vats, V., Thompson, C., Kopp, Z. S., and Milsom, I. National community prevalence of overactive bladder in the united states stratified by sex and age. *Urology*, 77(5):1081–1087, 2011.

Coyne, K., Zhou, Z., Thompson, C., and Versi, E. The impact on health-related quality of life of stress, urge and mixed urinary incontinence. *BJU international*, 92(7):731–735, 2003.

Cui, Y., Wang, L., Liu, L., Zeng, F., Niu, J., Qi, L., and Chen, H. Botulinum toxin-a injections for idiopathic overactive bladder: a systematic review and meta-analysis. *Urologia internationalis*, 91(4):429–438, 2013.

Cui, Y., Zong, H., Yang, C., Yan, H., and Zhang, Y. The efficacy and safety of mirabegron in treating oab: a systematic review and meta-analysis of phase iii trials. *International urology and nephrology*, 46(1):275–284, 2014.

Cui, Y., Zhou, X., Zong, H., Yan, H., and Zhang, Y. The efficacy and safety of onabotulinumtoxina in treating idiopathic oab: A systematic review and meta-analysis. *Neurourology and urodynamics*, 34(5):413–419, 2015.

Dakin, H. A., Welton, N. J., Ades, A., Collins, S., Orme, M., and Kelly, S. Mixed treatment comparison of repeated measurements of a continuous endpoint: an example using topical treatments for primary open-angle glaucoma and ocular hypertension. *Statistics in medicine*, 30(20):2511–2535, 2011.

Daniels, M. J. and Hughes, M. D. Meta-analysis for the evaluation of potential surrogate markers. *Statistics in medicine*, 16(17):1965–1982, 1997.

Davies, S. Annual report of the chief medical officer 2014: women's health. *London: Department of Health*, 2015.

Davis, N., Burke, J., Redmond, E., Elamin, S., Brady, C., and Flood, H. Trigonal versus extratrigonal botulinum toxin-a: a systematic review and meta-analysis of efficacy and adverse events. *International urogynecology journal*, 26(3):313–319, 2015.

Debray, T., Moons, K. G., Valkenhoef, G., Efthimiou, O., Hummel, N., Groenwold, R. H., and Reitsma, J. B. Get real in individual participant data (ipd) meta-analysis: a review of the methodology. *Research synthesis methods*, 6(4):293–309, 2015.

Demiris, N., Lunn, D., and Sharples, L. D. Survival extrapolation using the poly-weibull model. *Stat Methods Med Res*, 24(2):287–301, Apr 2015.

Dequen, P., Sutton, A. J., Scott, D. A., and Abrams, K. R. Searching for indirect evidence and extending the network of studies for network meta-analysis: case study in venous thromboembolic events prevention following elective total knee replacement surgery. *Value in Health*, 17(4):416–423, 2014.

Diamond, P., Hassonah, S., Alarab, M., Lovatsis, D., and Drutz, H. The prevalence of detrusor overactivity amongst patients with symptoms of overactive bladder: a retrospective cohort study. *International urogynecology journal*, 23(11):1577–1580, 2012.

Dias, S., Welton, N., Caldwell, D., and Ades, A. Checking consistency in mixed treatment comparison meta-analysis. *Statistics in medicine*, 29(7-8):932–944, 2010.

Dias, S., Welton, N. J., Sutton, A. J., and Ades, A. Nice dsu technical support document 1: Introduction to evidence synthesis for decision making. *University of Sheffield, Decision Support Unit*, pages 1–24, 2011a.

Dias, S., Welton, N. J., Sutton, A. J., and Ades, A. Nice dsu technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. *National Institute for Health and Clinical Excellence, London, UK*, 2011b.

Dias, S., Sutton, A. J., Welton, N. J., and Ades, A. Nice dsu technical support document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. *Report by the Decision Support Unit*, 2012.

Diener, H.-C. Placebo effects in treating migraine and other headaches. *Current opinion in investigational drugs (London, England: 2000)*, 11(7):735–739, 2010.

Digesu, G. A., Khullar, V., Cardozo, L., and Salvatore, S. Overactive bladder symptoms: do we need urodynamics? *Neurourology and urodynamics*, 22(2): 105–108, 2003.

Dmochowski, R. R. and Gomelsky, A. Overactive bladder in males. *Therapeutic advances in urology*, 1(4):209, 2009.

Dmochowski, R. R., Sand, P. K., Zinner, N. R., Gittelman, M. C., Davila, G. W., Sanders, S. W., Group, T. O. S., et al. Comparative efficacy and safety of transdermal oxybutynin and oral tolterodine versus placebo in previously

treated patients with urge and mixed urinary incontinence. *Urology*, 62(2): 237–242, 2003.

Dowson, C., Sahai, A., Watkins, J., Dasgupta, P., and Khan, M. The safety and efficacy of botulinum toxin-a in the management of bladder oversensitivity: a randomised double-blind placebo-controlled trial. *Int J Clin Pract*, 65(6): 698–704, Jun 2011.

Dowson, C., Watkins, J., Khan, M. S., Dasgupta, P., and Sahai, A. Repeated botulinum toxin type a injections for refractory overactive bladder: medium-term outcomes, safety profile, and discontinuation rates. *European urology*, 61(4):834–839, 2012.

Drummond, M. and Sorenson, C. Nasty or nice? a perspective on the use of health technology assessment in the united kingdom. *Value in Health*, 12(s2): S8–S13, 2009.

Drummond, M. F., Schwartz, J. S., Jönsson, B., Luce, B. R., Neumann, P. J., Siebert, U., and Sullivan, S. D. Key principles for the improved conduct of health technology assessments for resource allocation decisions. *International journal of technology assessment in health care*, 24(03):244–258, 2008.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., and Torrance, G. W. *Methods for the economic evaluation of health care programmes*. Oxford university press, 2015.

Duckett, J. and Balachandran, A. Tolerability and persistence in a large, prospective case series of women prescribed mirabegron. *International urogynecology journal*, pages 1–5, 2016.

Dumoulin, C. and Hay-Smith, J. Pelvic floor muscle training versus no treatment for urinary incontinence in women. a cochrane systematic review. *European journal of physical and rehabilitation medicine*, 44(1):47–63, 2008.

Dumoulin, C., Hay-Smith, J., et al. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. *Cochrane Database Syst Rev*, 1(1), 2010.

Eckermann, S. and Willan, A. R. Expected value of information and decision making in hta. *Health economics*, 16(2):195–209, 2007.

Edwards, A. L. *Probability and statistics*. Holt, Rinehart and Winston, 1971.

Efthimiou, O., Mavridis, D., Cipriani, A., Leucht, S., Bagos, P., and Salanti, G. An approach for modelling multiple correlated outcomes in a network of interventions using odds ratios. *Statistics in Medicine*, 33(13):2275–2287, 2014.

Eustice, S., Roe, B., and Paterson, J. Prompted voiding for the management of urinary incontinence in adults. *The Cochrane Library*, 2000.

Fitzgerald, M., Lemack, G., Wheeler, T., Litman, H., and , U. I. T. N. Nocturia, nocturnal incontinence prevalence, and response to anticholinergic and behavioral therapy. *Int Urogynecol J Pelvic Floor Dysfunct*, 19(11): 1545–1550, Nov 2008.

Fleurence, R. L., Naci, H., and Jansen, J. P. The critical role of observational evidence in comparative effectiveness research. *Health Aff (Millwood)*, 29(10): 1826–1833, Oct 2010.

Flood, H. D., Malhotra, S. J., O'Connell, H. E., Ritchey, M. J., Bloom, D. A., and McGuire, E. J. Long-term results and complications using augmentation cystoplasty in reconstructive urology. *Neurourology and urodynamics*, 14(4): 297–309, 1995.

Freedman, L. Bayesian statistical methods. *BMJ: British Medical Journal*, 313 (7057):569, 1996.

Freemantle, N., Khalaf, K., Loveman, C., Stanisic, S., Gultyaev, D., Lister, J., and Drake, M. Onabotulinumtoxina in the treatment of overactive bladder: a cost-effectiveness analysis versus best supportive care in england and wales. *The European Journal of Health Economics*, pages 1–11, 2015.

Freemantle, N., Ginsberg, D. A., McCool, R., Fleetwood, K., Arber, M., Khalaf, K., Loveman, C., Ni, Q., and Glanville, J. Comparative assessment of onabotulinumtoxina and mirabegron for overactive bladder: an indirect treatment comparison. *BMJ open*, 6(2):e009122, 2016.

Friedrich, J. O., Adhikari, N. K., and Beyene, J. Inclusion of zero total event trials in meta-analyses maintains analytic consistency and incorporates all available data. *BMC medical research methodology*, 7(1):1, 2007.

Gargon, E., Williamson, P. R., Altman, D. G., Blazeby, J. M., and Clarke, M. The comet initiative database: progress and activities update (2014). *Trials*, 16(1):1, 2015.

Gelfand, A. E. and Smith, A. F. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410): 398–409, 1990.

Gelman, A. and Rubin, D. B. Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472, 1992.

Gelman, A., Shirley, K., et al. Inference from simulations and monitoring convergence. *Handbook of Markov chain Monte Carlo*, pages 163–174, 2011.

Geman, S. and Geman, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741, 1984.

Geoffrion, R., Lovatsis, D., Walter, J.-E., Chou, Q., Easton, W., Epp, A., Harvey, M.-A., Larochelle, A., Maslow, K., Neustaedter, G., et al. Treatments for overactive bladder: focus on pharmacotherapy. *Journal of Obstetrics and Gynaecology Canada*, 34(11):1092–1101, 2012.

Geweke, J. and Amisano, G. Analysis of variance for bayesian inference. *Econometric Reviews*, 33(1-4):270–288, 2014.

Gilks, W. R. *Markov chain monte carlo*. Wiley Online Library, 2005.

Girman, C. J., Jacobsen, S. J., Tsukamoto, T., Richard, F., Garraway, W. M., Sagnier, P.-P., Guess, H. A., Rhodes, T., Boyle, P., and Lieber, M. M. Health-related quality of life associated with lower urinary tract symptoms in four countries. *Urology*, 51(3):428–436, 1998.

Glass, G. V. Primary, secondary, and meta-analysis of research. *Educational researcher*, 5(10):3–8, 1976.

Glasziou, P. P. and Irwig, L. M. An evidence based approach to individualising treatment. *BMJ: British Medical Journal*, 311(7016):1356, 1995.

Globerman, D. and Robert, M. Heterogeneity in post-intervention prolapse and urinary outcome reporting: a one-year review of the international urogynecology journal. *International urogynecology journal*, 26(9):1373–1378, 2015.

Goggins, W. B., Finkelstein, D. M., Schoenfeld, D. A., and Zaslavsky, A. M. A markov chain monte carlo em algorithm for analyzing interval-censored data under the cox proportional hazards model. *Biometrics*, pages 1498–1507, 1998.

Goodman, C. S. Introduction to health technology assessment. *National Library of Medicine, USA*, 2004.

Gormley, E. A., Lightner, D. J., Burgio, K. L., Chai, T. C., Clemens, J. Q., Culkin, D. J., Das, A. K., Foster, H. E., Scarpero, H. M., Tessier, C. D., et al. Diagnosis and treatment of overactive bladder (non-neurogenic) in adults: Aua/sufu guideline. *The Journal of urology*, 188(6):2455–2463, 2012.

Goulart, B. H. L., Ramsey, S. D., and Parvathaneni, U. Observational study designs for comparative effectiveness research: an alternative approach to close evidence gaps in head-and-neck cancer. *Int J Radiat Oncol Biol Phys*, 88(1): 106–114, Jan 2014.

Gousse, A. E., Kanagarajah, P., Ayyathurai, R., Handa, P., Dabas, N., and Gomez, C. S. Repeat intradetrusor injections of onabotulinum toxin a for

refractory idiopathic overactive bladder patients: a single-center experience. *Female pelvic medicine & reconstructive surgery*, 17(5):253–257, 2011.

Granese, R., Adile, G., Gugliotta, G., Cucinella, G., Saitta, S., and Adile, B. Botox® for idiopathic overactive bladder: efficacy, duration and safety. effectiveness of subsequent injection. *Archives of gynecology and obstetrics*, 286 (4):923–929, 2012.

Guest, J. F., Abegunde, D., and Ruiz, F. J. Cost effectiveness of controlled-release oxybutynin compared with immediate-release oxybutynin and tolterodine in the treatment of overactive bladder in the uk, france and austria. *Clinical drug investigation*, 24(6):305–321, 2004.

Gusi, N., Olivares, P., and Rajendram, R. The eq-5d health-related quality of life questionnaire. In *Handbook of disease burdens and quality of life measures*, pages 87–99. Springer, 2010.

Guyatt, G., Cairns, J., Churchill, D., and et al. Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA*, 268(17):2420–2425, 1992.

Hagstroem, S., Mahler, B., Madsen, B., Djurhuus, J. C., and Rittig, S. Transcutaneous electrical nerve stimulation for refractory daytime urinary urge incontinence. *The Journal of urology*, 182(4):2072–2078, 2009.

Haider, S. I., Johnell, K., Weitoft, G. R., Thorslund, M., and Fastbom, J. The influence of educational level on polypharmacy and inappropriate drug use: A register-based study of more than 600,000 older people. *Journal of the American Geriatrics Society*, 57(1):62–69, 2009.

Hamburg, M. A. and Collins, F. S. The path to personalized medicine. *New England Journal of Medicine*, 363(4):301–304, 2010.

Hanson, T. and Johnson, W. O. A bayesian semiparametric aft model for interval-censored data. *Journal of Computational and Graphical Statistics*, 2012.

Hardy, R. J. and Thompson, S. G. Detecting and describing heterogeneity in meta-analysis. *Statistics in medicine*, 17(8):841–856, 1998.

Harrell, F. E., Lee, K. L., and Mark, D. B. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in medicine*, 15: 361–387, 1996.

Hastings, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

Hawkins, N., Scott, D. A., and Woods, B. How far do you go? efficient searching for indirect evidence. *Medical decision making*, 29(3):273–281, 2009a.

Hawkins, N., Scott, D. A., Woods, B. S., and Thatcher, N. No study left behind: a network meta-analysis in non–small-cell lung cancer demonstrating the importance of considering all relevant data. *Value in health*, 12(6):996–1003, 2009b.

Haylen, B. T., De Ridder, D., Freeman, R. M., Swift, S. E., Berghmans, B., Lee, J., Monga, A., Petri, E., Rizk, D. E., Sand, P. K., et al. An international urogynecological association (iuga)/international continence society (ics) joint report on the terminology for female pelvic floor dysfunction. *International urogynecology journal*, 21(1):5–26, 2010.

Hempel, S., Miles, J. N., Booth, M. J., Wang, Z., Morton, S. C., and Shekelle, P. G. Risk of bias: a simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic reviews*, 2(1):1, 2013.

Henriet, B. and Roumeguère, T. Botulinum toxin injection for refractory non-neurogenic overactive bladder. systematic review. *Revue medicale de Bruxelles*, 36(1):29–37, 2014.

Herderschee, R., Hay-Smith, E. J. C., Herbison, G. P., Roovers, J. P., and Heineman, M. J. Feedback or biofeedback to augment pelvic floor muscle training for urinary incontinence in women. *The Cochrane Library*, 2011.

Herschorn, S., Becker, D., Miller, E., Thompson, M., and Forte, L. Impact of a health education intervention in overactive bladder patients. *Can J Urol*, 11 (6):2430–2437, Dec 2004.

Herschorn, S., Gajewski, J., Schulz, J., and Corcos, J. A population-based study of urinary symptoms and incontinence: the canadian urinary bladder survey. *BJU international*, 101(1):52–58, 2008.

Hextall, A. Oestrogens and lower urinary tract function. *Maturitas*, 36(2):83–92, 2000.

Higgins, J., Thompson, S. G., and Spiegelhalter, D. J. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1):137–159, 2009.

Higgins, J. P., Green, S., et al. *Cochrane handbook for systematic reviews of interventions*, volume 5. Wiley Online Library, 2008.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Juni, P., Moher, D., Oxman, A. D., SavoviÄ, J., Schulz, K. F., Weeks, L., and Sterne, J. A. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ: British Medical Journal*, 343, 2011.

Hoaglin, D. C., Hawkins, N., Jansen, J. P., Scott, D. A., Itzler, R., Cappelleri, J. C., Boersma, C., Thompson, D., Larholt, K. M., Diaz, M., et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ispor task force on indirect treatment comparisons good research practices: part 2. *Value in health*, 14(4):429–437, 2011.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. Bayesian model averaging: a tutorial. *Statistical science*, pages 382–401, 1999.

Hoffman, M. D. and Gelman, A. The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15 (1):1593–1623, 2014.

Homma, Y. and Kawabe, K. Health-related quality of life of japanese patients with overactive bladder treated with extended-release tolterodine or immediate-release oxybutynin: a randomized, placebo-controlled trial. *World J Urol*, 22(4):251–256, Oct 2004.

Homma, Y., Kakizaki, H., Yamaguchi, O., Yamanishi, T., Nishizawa, O., Yokoyama, O., Takeda, M., Seki, N., and Yoshida, M. Assessment of overactive bladder symptoms: comparison of 3-day bladder diary and the overactive bladder symptoms score. *Urology*, 77(1):60–64, 2011.

Hong, H., Carlin, B. P., Shamliyan, T. A., Wyman, J. F., Ramakrishnan, R., Sainfort, F., and Kane, R. L. Comparing bayesian and frequentist approaches for multiple outcome mixed treatment comparisons. *Medical Decision Making*, page 0272989X13481110, 2013.

Hu, T. and Wagner, T. H. Economic considerations in overactive bladder. *Am J Manag Care*, 6(11 Suppl):S591–S598, 2000.

Huang, W., Zong, H., Zhou, X., and Zhang, Y. Efficacy and safety of imidafenacin for overactive bladder in adult: a systematic review and meta-analysis. *International urology and nephrology*, 47(3):457–464, 2015.

Hughes, D. A. and Dubois, D. Cost-effectiveness analysis of extended-release formulations of oxybutynin and tolterodine for the management of urge incontinence. *Pharmacoeconomics*, 22(16):1047–1059, 2004.

Hutcheon, J. A., Chiolero, A., and Hanley, J. A. Random measurement error and regression dilution bias. *Bmj*, 340:c2289, 2010.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. *Bayesian survival analysis*. Wiley Online Library, 2005.

International Consortium for Health Outcome Measurements. Ichom overactive bladder data collection reference guide, 2016. URL `http://www.ichom.org/medical-conditions/overactive-bladder/`.

Ioannidis, J., Haidich, A., Pappa, M., Pantazis, N., Kokori, S., Tektonidou, M., Contopoulos-Ioannidis, D., and Lau, J. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA*, 286(7):821–830, Aug 2001.

Irwin, D. E., Milsom, I., Hunskaar, S., Reilly, K., Kopp, Z., Herschorn, S., Coyne, K., Kelleher, C., Hampel, C., Artibani, W., et al. Population-based survey of urinary incontinence, overactive bladder, and other lower urinary tract symptoms in five countries: results of the epic study. *European urology*, 50(6):1306–1315, 2006.

Irwin, D. E., Kopp, Z. S., Agatep, B., Milsom, I., and Abrams, P. Worldwide prevalence estimates of lower urinary tract symptoms, overactive bladder, urinary incontinence and bladder outlet obstruction. *BJU international*, 108 (7):1132–1138, 2011.

Jackson, D., Riley, R., and White, I. R. Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, 30(20):2481–2498, 2011.

Jansen, J. P., Crawford, B., Bergman, G., and Stam, W. Bayesian meta-analysis of multiple treatment comparisons: An introduction to mixed treatment comparisons. *Value in Health*, 11(5):956–964, 2008.

Jansen, J. P., Fleurence, R., Devine, B., Itzler, R., Barrett, A., Hawkins, N., Lee, K., Boersma, C., Annemans, L., and Cappelleri, J. C. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ispor task force on indirect treatment comparisons good research practices: part 1. *Value in Health*, 14(4):417–428, 2011.

Jara, A., Lesaffre, E., De Iorio, M., Quintana, F., et al. Bayesian semiparametric inference for multivariate doubly-interval-censored data. *The Annals of Applied Statistics*, 4(4):2126–2149, 2010.

Johnson, N., Kotz, S., and Balakrishnan, N. Continuous univariate distributions. volume 1, 2nd edition, 1994.

Johnson, N. L., Kemp, A. W., and Kotz, S. *Univariate Discrete Distributions, Set*, volume 444. John Wiley & Sons, 2005.

Johnston, N., Kotz, S., and Balakrishnan, N. Continuous univariate distributions (volume 2), 1994.

Kadane, J. B. Prime time for bayes. *Controlled Clinical Trials*, 16(5):313–318, 1995.

Kafri, R., Deutscher, D., Shames, J., Golombp, J., and Melzer, I. Randomized trial of a comparison of rehabilitation or drug therapy for urgency urinary incontinence: 1-year follow-up. *Int Urogynecol J*, 24(7):1181–1189, Jul 2013.

Kairalla, J. A., Coffey, C. S., Thomann, M. A., and Muller, K. E. Adaptive trial designs: a review of barriers and opportunities. *Trials*, 13(1):1, 2012.

Kanters, S., Mills, E., Thorlund, K., Bucher, H., and Ioannidis, J. Antiretroviral therapy for initial human immunodeficiency virus/aids treatment: critical appraisal of the evidence from over 100 randomized trials and 400 systematic reviews and meta-analyses. *Clinical Microbiology and Infection*, 20(2):114–122, 2014.

Karademir, K., Baykal, K., Sen, B., Senkul, T., Iseri, C., and Erden, D. A peripheric neuromodulation technique for curing detrusor overactivity: Stoller afferent neurostimulation. *Scand J Urol Nephrol*, 39(3):230–233, 2005.

Kaya, S., Akbayrak, T., and BeksaÃ§, S. Comparison of different treatment protocols in the treatment of idiopathic detrusor overactivity: a randomized controlled trial. *Clin Rehabil*, 25(4):327–338, Apr 2011.

Kelleher, C., Cardozo, L., Khullar, V., and Salvatore, S. A medium term analysis of the subjective efficacy of treatment for women with detrusor instability and low bladder compliance. *BJOG: An International Journal of Obstetrics & Gynaecology*, 104(9):988–993, 1997a.

Kelleher, C., Cardozo, L., Khullar, V., and Salvatore, S. A medium term analysis of the subjective efficacy of treatment for women with detrusor instability and

low bladder compliance. *BJOG: An International Journal of Obstetrics & Gynaecology*, 104(9):988–993, 1997b.

Kelleher, C., Cardozo, L., Khullar, V., and Salvatore, S. A new questionnaire to assess the quality of life of urinary incontinent women. *BJOG: An International Journal of Obstetrics & Gynaecology*, 104(12):1374–1379, 1997c.

Kelleher, C. J., Reese, P. R., Pleil, A. M., and Okano, G. J. Health-related quality of life of patients receiving extended-release tolterodine for overactive bladder. *Am J Manag Care*, 8(19 Suppl):S608–S615, Dec 2002.

Kim, Y., Yoshimura, N., Masuda, H., de Miguel, F., and Chancellor, M. B. Antimuscarinic agents exhibit local inhibitory effects on muscarinic receptors in bladder-afferent pathways. *Urology*, 65(2):238–242, 2005.

Kirkham, J. J., Riley, R. D., and Williamson, P. R. A multivariate meta-analysis approach for reducing the impact of outcome reporting bias in systematic reviews. *Statistics in medicine*, 31(20):2179–2195, 2012.

Klotz, T., Brüggenjürgen, B., Burkart, M., and Resch, A. The economic costs of overactive bladder in germany. *European urology*, 51(6):1654–1663, 2007.

Komárek, A., Lesaffre, E., Härkänen, T., Declerck, D., and Virtanen, J. I. A bayesian analysis of multivariate doubly-interval-censored dental data. *Biostatistics*, 6(1):145–155, 2005.

Kosilov, K., Loparev, S., Ivanovskaya, M., and Kosilova, L. A randomized, controlled trial of effectiveness and safety of management of oab symptoms in elderly men and women with standard-dosed combination of solifenacin and mirabegron. *Arch Gerontol Geriatr*, 61(2):212–216, 2015.

Kosilov, K. V., Loparev, S. A., Ivanovskaya, M. A., and Kosilova, L. V. Comparative effectiveness of combined low- and standard-dose trospium and solifenacin for moderate overactive bladder symptoms in elderly men and women. *Urol Int*, 93(4):470–473, 2014.

Kotz, S., Balakrishnan, N., and Johnson, N. L. *Continuous multivariate distributions, models and applications*. John Wiley & Sons, 2004.

Kreder, K. and Dmochowski, R. *The Overactive Bladder: Evaluation and Management*. CRC Press, 2007. ISBN 9780203931622.

Kuo, H.-C. Bladder base/trigone injection is safe and as effective as bladder body injection of onabotulinumtoxina for idiopathic detrusor overactivity refractory to antimuscarinics. *Neurourology and urodynamics*, 30(7):1242–1248, 2011.

Lambert, P. C., Sutton, A. J., Abrams, K. R., and Jones, D. R. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of clinical epidemiology*, 55(1):86–94, 2002.

Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., and Jones, D. R. How vague is vague? a simulation study of the impact of the use of vague prior distributions in mcmc using winbugs. *Stat Med*, 24(15):2401–2428, Aug 2005.

Laska, E. M., Meisner, M., Siegel, C., and Stinnett, A. A. Ratio-based and net benefit-based approaches to health care resource allocation: proofs of optimality and equivalence. *Health Economics*, 8(2):171–174, 1999.

Leng, W. W. and Chancellor, M. B. How sacral nerve stimulation neuromodulation works. *Urologic Clinics*, 32(1):11–18, 2005.

Lie, K. Y., Wong, M. Y., and Ng, L. G. Botulinum toxin a for idiopathic detrusor overactivity. *Annals Academy of Medicine Singapore*, 39(9):714, 2010.

Lilford, R. and Braunholtz, D. The statistical basis of public policy: a paradigm shift is overdue. *BMJ: British Medical Journal*, 313(7057):603, 1996.

Lindsey, J. C. and Ryan, L. M. Methods for interval-censored data. *Statistics in medicine*, 17(2):219–238, 1998.

Liu, B., Li, P., Li, J., Wang, Y., and Wu, Y. Comparisons of therapeutic efficacy and safety of solifenacin versus tolterodine in patients with overactive bladder: a meta-analysis of outcomes. *Zhonghua yi xue za zhi*, 94(30):2350–2354, 2014.

Lose, G., Mattiasson, A., Walter, S., Lalos, O., Van Kerrebroeck, P., Abrams, P., and Freeman, R. Clinical experiences with desmopressin for long-term treatment of nocturia. *The Journal of urology*, 172(3):1021–1025, 2004.

Lu, G. and Ades, A. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in medicine*, 23(20):3105–3124, 2004.

Lu, G. and Ades, A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics*, 10(4):792–805, 2009.

Lumley, T. Network meta-analysis for indirect treatment comparisons. *Statistics in medicine*, 21(16):2313–2324, 2002.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. The bugs project: Evolution, critique and future directions. *Statistics in medicine*, 28(25): 3049–3067, 2009.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. *The BUGS book: A practical introduction to Bayesian analysis*. CRC press, 2012.

Luo, D., Liu, L., Han, P., Wei, Q., and Shen, H. Solifenacin for overactive bladder: a systematic review and meta-analysis. *International urogynecology journal*, 23(8):983–991, 2012.

MacDiarmid, S. A. Maximizing the treatment of overactive bladder in the elderly. *Reviews in urology*, 10(1):6, 2008.

Madhuvrata, P., Cody, J. D., Ellis, G., Herbison, G. P., and Hay-Smith, E. J. C. Which anticholinergic drug for overactive bladder symptoms in adults. *The Cochrane Library*, 2012.

Maman, K., Aballea, S., Nazir, J., Desroziers, K., Neine, M.-E., Siddiqui, E., Odeyemi, I., and Hakimi, Z. Comparative efficacy and safety of medical treatments for the management of overactive bladder: a systematic literature review and mixed treatment comparison. *European urology*, 65(4):755–765, 2014.

Mangera, A., Apostolidis, A., Andersson, K. E., Dasgupta, P., Giannantoni, A., Roehrborn, C., Novara, G., and Chapple, C. An updated systematic review and statistical comparison of standardised mean outcomes for the use of botulinum toxin in the management of lower urinary tract disorders. *European urology*, 65(5):981–990, 2014.

Marinkovic, S. P., Rovner, E. S., Moldwin, R. M., Stanton, S. L., Gillen, L. M., and Marinkovic, C. M. The management of overactive bladder syndrome. *BMJ-British Medical Journal*, 344(7):e2365, 2012.

Marko, N. F. and Weil, R. J. The role of observational investigations in comparative effectiveness research. *Value Health*, 13(8):989–997, Dec 2010.

Mathew, N. T., Frishberg, B. M., Gawel, M., Dimitrova, R., Gibson, J., and Turkel, C. Botulinum toxin type a (botox®) for the prophylactic treatment of chronic daily headache: A randomized, double-blind, placebo-controlled trial. *Headache: The Journal of Head and Face Pain*, 45(4):293–307, 2005.

Mavridis, D. and Salanti, G. A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research*, page 0962280211432219, 2012.

McCarron, C. E., Pullenayegum, E. M., Thabane, L., Goeree, R., and Tarride, J.-E. Bayesian hierarchical models combining different study types and adjusting for covariate imbalances: a simulation study to assess model performance. *PloS one*, 6(10):e25635, 2011.

McGrother, C., Donaldson, M., Shaw, C., Matthews, R. J., Hayward, T., Dallosso, H., Jagger, C., Clarke, M., and Castleden, C. Storage symptoms of the bladder: prevalence, incidence and need for services in the uk. *BJU international*, 93(6):763–769, 2004.

Millard, A. P. T. S. G., RJ. Clinical efficacy of tolterodine with or without a simplified pelvic floor exercise regimen. *Neurourol Urodyn*, 23(1):48–53, 2004.

Mills, E. J., Kanters, S., Thorlund, K., Chaimani, A., Veroniki, A.-A., and Ioannidis, J. P. The effects of excluding treatments from network meta-analyses: survey. *Bmj*, 347:f5195, 2013a.

Mills, E. J., Thorlund, K., and Ioannidis, J. P. Demystifying trial networks and network meta-analysis. *Bmj*, 346:f2914, 2013b.

Milsom, I., Abrams, P., Cardozo, L., Roberts, R., Thüroff, J., and Wein, A. How widespread are the symptoms of an overactive bladder and how are they managed? a population-based prevalence study. *BJU international*, 87(9): 760–766, 2001.

Mullins, C. D. and Sanchez, R. J. It is important to note that rwd will never replace the more traditional and more robust rct data; however, the emerging trend is to incorporate data that are more generalizable. introduction. *J Manag Care Pharm*, 17(9 Suppl A):S03–S04, 2011.

Nam, I.-S., Mengersen, K., and Garthwaite, P. Multivariate meta-analysis. *Statistics in medicine*, 22(14):2309–2333, 2003.

Nascimento-Correia, G., Santos-Pereira, V., Tahara, N., and Driusso, P. Effects of pelvic floor muscle training on quality of life of a group of women with urinary incontinence: randomized controlled trial. *Actas Urol Esp*, 36(4): 216–221, Apr 2012.

National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. *London: NICE*, 2013a.

National Institute for Health and Care Excellence. Urinary incontinence in women: management (cg171), 2013b. URL `http://www.nice.org.uk/guidance/cg171/resources/urinary-incontinence-in-women-management-35109747194821`.

National Institute for Health and Care Excellence. Urinary incontinence in women (qs77), 2015. URL `https://www.nice.org.uk/guidance/qs77/resources/urinary-incontinence-in-women-2098853147077`.

Nazir, J., Maman, K., Neine, M.-E., Briquet, B., Odeyemi, I. A., Hakimi, Z., Garnham, A., and Aballéa, S. Cost-effectiveness of mirabegron compared with antimuscarinic agents for the treatment of adults with overactive bladder in the united kingdom. *Value in Health*, 18(6):783–790, 2015a.

Nazir, J., Posnett, J., Walker, A., Odeyemi, I. A., Hakimi, Z., and Garnham, A. Economic evaluation of pharmacological treatments for overactive bladder from the perspective of the uk national health service. *Journal of medical economics*, 18(5):390–397, 2015b.

Newman, D. K., Hanno, P. M., Dmochowski, R. R., Rudy, D. C., Thomas, H., Caramelli, K. E., and Hoel, G. Effects of oxybutynin chloride topical gel on health-related quality of life in adults with overactive bladder: a randomized, double-blind, placebo-controlled study. *Clinical Medicine Insights: Therapeutics*, 2:889–896, 11 2010.

Nitti, V., Khullar, V., Kerrebroeck, P., Herschorn, S., Cambronero, J., Angulo, J., Blauwet, M., Dorrepaal, C., Siddiqui, E., and Martin, N. Mirabegron for the treatment of overactive bladder: a prespecified pooled efficacy analysis and pooled safety analysis of three randomised, double-blind, placebo-controlled, phase iii studies. *International journal of clinical practice*, 67(7):619–632, 2013.

Nixon, A., Colman, S., Sabounjian, L., Sandage, B., Schwiderski, U. E., Staskin, D. R., and Zinner, N. A validated patient reported measure of urinary urgency severity in overactive bladder for use in clinical trials. *The Journal of urology*, 174(2):604–607, 2005.

Nixon, R., Bansback, N., and Brennan, A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Statistics in medicine*, 26(6): 1237–1254, 2007.

Novara, G., Galfano, A., Secco, S., DElia, C., Cavalleri, S., Ficarra, V., and Artibani, W. A systematic review and meta-analysis of randomized controlled

trials with antimuscarinic drugs for overactive bladder. *European urology*, 54 (4):740–764, 2008a.

Novara, G., Galfano, A., Secco, S., D'Elia, C., Cavalleri, S., Ficarra, V., and Artibani, W. A systematic review and meta-analysis of randomized controlled trials with antimuscarinic drugs for overactive bladder. *European urology*, 54 (4):740–764, 2008b.

Odell, P. M., Anderson, K. M., and D'Agostino, R. B. Maximum likelihood estimation for interval-censored data using a weibull-based accelerated failure time model. *Biometrics*, pages 951–959, 1992.

O'Donnell, M., Lose, G., Sykes, D., Voss, S., and Hunskaar, S. Help-seeking behaviour and associated factors among women with urinary incontinence in france, germany, spain and the united kingdom. *European urology*, 47(3): 385–392, 2005.

Olivera, C. K., Meriwether, K., El-Nashar, S., Grimes, C. L., Chen, C. C. G., Orejuela, F., Antosh, D., Gleason, J., Kim-Fine, S., Wheeler, T., et al. Nonantimuscarinic treatment for overactive bladder: a systematic review. *American journal of obstetrics and gynecology*, 2016.

Orasanu, B. and Mahajan, S. T. The use of botulinum toxin for the treatment of overactive bladder syndrome. *Indian journal of urology: IJU: journal of the Urological Society of India*, 29(1):2, 2013.

Orešković, S., But, I., Banović, M., and Šprem Goldštajn, M. The efficacy and safety of solifenacin in patients with overactive bladder syndrome. *Collegium antropologicum*, 36(1):243–248, 2012.

Owen, R. K., Tincello, D. G., and Keith, R. A. Network meta-analysis: development of a three-level hierarchical modeling approach incorporating dose-related constraints. *Value in Health*, 18(1):116–126, 2015.

Owen, R. K., Abrams, K. R., Mayne, C., Slack, M., and Tincello, D. G. Comparison of the effectiveness of repeated injections of onabotulinum toxin a

for refractory idiopathic detrusor overactivity: analysis of an open label extension of a randomized trial (the relax study). *Neurourology and Urodynamics*, 2016a.

Owen, R. K., Abrams, K. R., Mayne, C., Slack, M., and Tincello, D. G. Patient factors associated with onabotulinum toxin a treatment outcome in women with detrusor overactivity. *Neurourology and urodynamics*, 2016b.

Owen, R. K., Tincello, D. G., and Abrams, K. R. Comparative efficacy, safety and tolerability of interventions for idiopathic detrusor overactivity and overactive bladder in adults: a comprehensive systematic review and network meta-analysis. *BMJ (submitted)*, 2017.

Ozdedeli, S., Karapolat, H., and Akkoc, Y. Comparison of intravaginal electrical stimulation and trospium hydrochloride in women with overactive bladder syndrome: a randomized controlled study. *Clin Rehabil*, 24(4):342–351, Apr 2010.

Peduzzi, P., Concato, J., Feinstein, A. R., and Holford, T. R. Importance of events per independent variable in proportional hazards regression analysis ii. accuracy and precision of regression estimates. *Journal of clinical epidemiology*, 48(12):1503–1510, 1995.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., and Feinstein, A. R. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, 49(12):1373–1379, 1996.

Pennell, M. L. and Dunson, D. B. Bayesian semiparametric dynamic frailty models for multiple event time data. *Biometrics*, 62(4):1044–1052, Dec 2006.

Perry, S., Gardner, E., and Thamer, M. The status of health technology assessment worldwide: Results of an international survey. *International journal of technology assessment in health care*, 13(01):81–98, 1997.

Peters, K. M., Feber, K. M., and Bennett, R. C. Sacral versus pudendal nerve stimulation for voiding dysfunction: a prospective, single-blinded, randomized, crossover trial. *Neurourol Urodyn*, 24(7):643–647, 2005.

Peto, R. Experimental survival curves for interval-censored data. *Applied Statistics*, pages 86–91, 1973.

Phillippo, D. M., Ades, A. E., Dias, S., Palmer, S., Abrams, K. R., and Welton, N. J. Nice dsu technical support document 18: methods for population-adjusted indirect comparisons in submissions to nice. *School of Health and Related Research, University of Sheffield, Sheffield, UK*, 2016.

Pibouleau, L. and Chevret, S. Bayesian hierarchical meta-analysis model for medical device evaluation: application to intracranial stents. *International journal of technology assessment in health care*, 29(02):123–130, 2013.

Pleil, A., Reese, P., Kelleher, C., and Okano, G. Health-related quality of life of patients with overactive bladder receiving immediate-release tolterodine. *The European Journal of Health Economics (HEPAC)*, 2(2):69–75, 2001. ISSN 1618-7598.

Plummer, M. Jags: A program for analysis of bayesian graphical models using gibbs sampling, 2003.

Prevost, T., Abrams, K., and Jones, D. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. *Stat Med*, 19(24):3359–3376, Dec 2000.

Price, N., Dawood, R., and Jackson, S. R. Pelvic floor exercise for urinary incontinence: a systematic literature review. *Maturitas*, 67(4):309–315, 2010.

Radke, B. R. A demonstration of interval-censored survival analysis. *Preventive veterinary medicine*, 59(4):241–256, 2003.

Rawlins, M. D. and Culyer, A. J. National institute for clinical excellence and its value judgments. *Bmj*, pages 224–227, 2004.

Reynolds, W. S., McPheeters, M., Blume, J., Surawicz, T., Worley, K., Wang, L., and Hartmann, K. Comparative effectiveness of anticholinergic therapy for overactive bladder in women: a systematic review and meta-analysis. *Obstetrics & Gynecology*, 125(6):1423–1432, 2015.

Riley, R. D. Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(4):789–811, 2009.

Riley, R. D., Abrams, K., Lambert, P., Sutton, A., and Thompson, J. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Statistics in medicine*, 26(1):78–97, 2007.

Riley, R. D., Lambert, P. C., and Abo-Zaid, G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *Bmj*, 340:c221, 2010.

Robinson, D. and Cardozo, L. Antimuscarinic drugs to treat overactive bladder. *BMJ*, 344:e2130, 2012.

Robinson, D. and Cardozo, L. D. The role of estrogens in female lower urinary tract dysfunction. *Urology*, 62(4):45–51, 2003.

Robinson, D., Cardozo, L., Terpstra, G., Bolodeoku, J., and , T. S. G. A randomized double-blind placebo-controlled multicentre study to explore the efficacy and safety of tamsulosin and tolterodine in women with overactive bladder syndrome. *BJU Int*, 100(4):840–845, Oct 2007.

Rogers, R. G., Bachmann, G., Scarpero, H., Jumadilova, Z., Sun, F., Morrow, J. D., Guan, Z., and Bavendam, T. Effects of tolterodine er on patient-reported outcomes in sexually active women with overactive bladder and urgency urinary incontinence. *Curr Med Res Opin*, 25(9):2159–2165, Sep 2009.

Rosa, G. M., Ferrero, S., Nitti, V. W., Wagg, A., Saleem, T., and Chapple, C. R. Cardiovascular safety of β 3-adrenoceptor agonists for the treatment of

patients with overactive bladder syndrome. *European urology*, 69(2):311–323, 2016.

Rothstein, H., Sutton, A. J., and Borenstein, M. *Publication Bias in Meta-analysis : Prevention, Assessment and Adjustments.* Wiley, 2005. ISBN 9780470870143.

Royston, P. and Parmar, M. K. 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 medical research methodology*, 13(1):152, 2013.

Royston, P., Lambert, P. C., et al. Flexible parametric survival analysis using stata: beyond the cox model. 2011.

Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., and Richardson, W. S. Evidence based medicine: what it is and what it isn't. *Bmj*, 312(7023): 71–72, 1996.

Sahai, A., Dowson, C., Khan, M. S., and Dasgupta, P. Improvement in quality of life after botulinum toxin-a injections for idiopathic detrusor overactivity: results from a randomized double-blind placebo-controlled trial. *BJU Int*, 103 (11):1509–1515, Jun 2009.

Sahai, A., Dowson, C., Khan, M. S., Dasgupta, P., Group, G. B. S., et al. Repeated injections of botulinum toxin-a for idiopathic detrusor overactivity. *Urology*, 75(3):552–558, 2010.

Salanti, G., Ades, A., and Ioannidis, J. P. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of clinical epidemiology*, 64(2):163–171, 2011.

Schmitz, S., Adams, R., and Walsh, C. Incorporating data from various trial designs into a mixed treatment comparison model. *Statistics in medicine*, 32 (17):2935–2949, 2013.

Schwarzer, G., Carpenter, J. R., and Rücker, G. Multivariate meta-analysis. In *Meta-Analysis with R*, pages 165–185. Springer, 2015.

Shaban, A. M. and Drake, M. J. Botulinum toxin treatment for overactive bladder: risk of urinary retention. *Current urology reports*, 9(6):445–451, 2008.

Siegel, S. W., Richardson, D. A., Miller, K. L., Karram, M. M., Blackwood, N. B., Sand, P. K., Staskin, D. R., and Tuttle, J. P. Pelvic floor electrical stimulation for the treatment of urge and mixed urinary incontinence in women. *Urology*, 50(6):934–940, 1997.

Silberstein, S. D., Stark, S. R., Lucas, S. M., Christie, S. N., Degryse, R. E., Turkel, C. C., Group, B.-. S., et al. Botulinum toxin type a for the prophylactic treatment of chronic daily headache: a randomized, double-blind, placebo-controlled trial. In *Mayo Clinic Proceedings*, volume 80, pages 1126–1137. Elsevier, 2005.

Singh, G., Wilkinson, J., and Thomas, D. Supravesical diversion for incontinence: a long-term follow-up. *British journal of urology*, 79(3):348–353, 1997.

Sinha, D., Chen, M.-H., and Ghosh, S. K. Bayesian analysis and model selection for interval-censored survival data. *Biometrics*, 55(2):585–590, 1999.

Souto, S. C., Reis, L. O., Palma, T., Palma, P., and Denardi, F. Prospective and randomized comparison of electrical stimulation of the posterior tibial nerve versus oxybutynin versus their combination for treatment of women with overactive bladder syndrome. *World J Urol*, 32(1):179–184, Feb 2014.

Spiegelhalter, D. Some dic slides. *MRC Biostatistics Unit, Cambridge. IceBUGS: Finland*, 2006.

Spiegelhalter, D., Myles, J. P., Jones, D. R., and Abrams, K. R. An introduction to bayesian methods in health technology assessment. *British Medical Journal*, 319(7208):508, 1999a.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. Winbugs user manual, 2003.

Spiegelhalter, D. J., Myles, J. P., Jones, D. R., and Abrams, K. R. *Bayesian methods in health technology assessment: a review*. NCCHTA, 2000.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.

Spiegelhalter, D. J., Abrams, K. R., and Myles, J. P. *Bayesian approaches to clinical trials and health-care evaluation*, volume 13. John Wiley & Sons, 2004.

Spiegelhalter, D., Myles, J., Jones, D., and Abrams, K. Methods in health service research. an introduction to bayesian methods in health technology assessment. *BMJ*, 319(7208):508–512, Aug 1999b.

StataCorp. *Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.*, 2015.

Steers, W. D., Herschorn, S., Kreder, K. J., Moore, K., Strohbehn, K., Yalcin, I., and Bump, R. C. Duloxetine compared with placebo for treating women with symptoms of overactive bladder. *BJU international*, 100(2):337–345, 2007.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., and Carpenter, J. R. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj*, 338: b2393, 2009.

Stewart, W., Van Rooyen, J., Cundiff, G., Abrams, P., Herzog, A., Corey, R., Hunt, T., and Wein, A. Prevalence and burden of overactive bladder in the united states. *World journal of urology*, 20(6):327–336, 2003.

Sturtz, S. and Bender, R. Unsolved issues of mixed treatment comparison meta-analysis: network size and inconsistency. *Research Synthesis Methods*, 3 (4):300–311, 2012.

Sun, J. *The statistical analysis of interval-censored failure time data*. Springer Science & Business Media, 2007.

Sun, Y., Luo, D., Tang, C., Yang, L., and Shen, H. The safety and efficiency of onabotulinumtoxina for the treatment of overactive bladder: a systematic

review and meta-analysis. *International urology and nephrology*, 47(11): 1779–1788, 2015.

Sutton, A., Ades, A., Cooper, N., and Abrams, K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics*, 26(9): 753–767, 2008.

Sutton, A. J. and Abrams, K. R. Bayesian methods in meta-analysis and evidence synthesis. *Statistical methods in medical research*, 10(4):277–303, 2001.

Sutton, A. J., Abrams, K. R., Jones, D. R., Jones, D. R., Sheldon, T. A., and Song, F. *Methods for meta-analysis in medical research*, volume 1. Wiley Chichester, 2000.

Sutton, A. J., Cooper, N. J., Abrams, K. R., Lambert, P. C., and Jones, D. R. A bayesian approach to evaluating net clinical benefit allowed for parameter uncertainty. *Journal of clinical epidemiology*, 58(1):26–40, 2005.

Swami, K., Feneley, R., Hammonds, J., and Abrams, P. Detrusor myectomy for detrusor overactivity: a minimum 1-year follow-up. *British journal of urology*, 81:68–72, 1998.

Sweeney, P., Mutambirwa, S., Van An, N., Sharma, J., and Vanamail, P. Flavoxate in the symptomatic treatment of overactive bladder: a meta-analysis. *European review for medical and pharmacological sciences*, 20 (17):3703, 2016.

Takwoingi, Y., Guo, B., Riley, R. D., and Deeks, J. J. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Statistical methods in medical research*, page 0962280215592269, 2015.

Tang, H., Chen, J., Wang, Y., Yu, T., Guo, C., and Liao, X. Combination of sacral neuromodulation and tolterodine for treatment of idiopathic overactive bladder in women: a clinical trial. *Urology journal*, 11(4):1800, 2014.

Tanner, M. A. and Wong, W. H. The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398): 528–540, 1987.

Thiagamoorthy, G., Giarenis, I., and Cardozo, L. Early investigational $\beta 3$ adreno-receptor agonists for the management of the overactive bladder syndrome. *Expert opinion on investigational drugs*, 24(10):1299–1306, 2015.

Thokala, P. and Duenas, A. Multiple criteria decision analysis for health technology assessment. *Value in Health*, 15(8):1172–1181, 2012.

Thompson, S. G., Smith, T. C., and Sharp, S. J. Investigating underlying risk as a source of heterogeneity in meta-analysis. *Statistics in medicine*, 16(23): 2741–2758, 1997.

Tikkinen, K. A., Auvinen, A., Tiitinen, A., Valpas, A., Johnson, T. M., and Tammela, T. L. Reproductive factors associated with nocturia and urinary urgency in women: a population-based study in finland. *American journal of obstetrics and gynecology*, 199(2):153–e1, 2008.

Tincello, D., Owen, R., Slack, M., and Abrams, K. Validation of the patient global impression scales for use in detrusor overactivity: secondary analysis of the relax study. *BJOG: An International Journal of Obstetrics & Gynaecology*, 120(2):212–216, 2013.

Tincello, D. G., Williams, K. S., Joshi, M., Assassa, R. P., and Abrams, K. R. Urinary diaries: a comparison of data collected for three days versus seven days. *Obstetrics & Gynecology*, 109(2, Part 1):277–280, 2007.

Tincello, D. G., Kenyon, S., Abrams, K. R., Mayne, C., Toozs-Hobson, P., Taylor, D., and Slack, M. Botulinum toxin a versus placebo for refractory detrusor overactivity in women: a randomised blinded placebo-controlled trial of 240 women (the relax study). *European urology*, 62(3):507–514, 2012.

Tubaro, A. Defining overactive bladder: epidemiology and burden of disease. *Urology*, 64(6):2–6, 2004.

Turner, D., Shaw, C., McGrother, C., Dallosso, H., Cooper, N., and Team, T. M. I. The cost of clinically significant urinary storage symptoms for community dwelling adults in the uk. *BJU international*, 93(9):1246–1252, 2004.

Turner, R. M., Spiegelhalter, D. J., Smith, G., and Thompson, S. G. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1):21–47, 2009.

van Haasteren, R. Marginal likelihood calculation with mcmc methods. In *Gravitational Wave Detection and Data Analysis for Pulsar Timing Arrays*, pages 99–120. Springer, 2014.

Van Houwelingen, H. C., Arends, L. R., and Stijnen, T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in medicine*, 21(4):589–624, 2002.

van Kerrebroeck, P., Abrams, P., Chaikin, D., Donovan, J., Fonda, D., Jackson, S., Jennum, P., Johnson, T., Lose, G., Mattiasson, A., et al. The standardisation of terminology in nocturia: Report from the standardisation sub-committee of the international continence society. *Neurourology and urodynamics*, 21(2):179–183, 2002.

Vandoninck, V., van Balken, M. R., Agró, E. F., Petta, F., Caltagirone, C., Heesakkers, J. P., Kiemeney, L. A., Debruyne, F. M., and Bemelmans, B. L. Posterior tibial nerve stimulation in the treatment of urge incontinence. *Neurourology and urodynamics*, 22(1):17–23, 2003.

Wagg, A., Franks, B., Ramos, B., and Berner, T. Persistence and adherence with the new beta-3 receptor agonist, mirabegron, versus antimuscarinics in overactive bladder: early experience in canada. *Canadian Urological Association Journal*, 9(9-10):343, 2015.

Wagg, A., Nitti, V. W., Kelleher, C., Castro-Diaz, D., Siddiqui, E., and Berner, T. Oral pharmacotherapy for overactive bladder in older patients: mirabegron as a potential alternative to antimuscarinics. *Current medical research and opinion*, 32(4):621–638, 2016.

Wagner, T., Patrick, D. L., Bavendam, T., Martin, M., and Buesching, D. Quality of life of persons with urinary incontinence: development of a new measure. *Urology*, 47(1):67–71, 1996.

Walsh, C. D. and Mengersen, K. L. Ordering of hierarchies in hierarchical models: Bone mineral density estimation. *Case Studies in Bayesian Statistical Modelling and Analysis*, pages 159–170, 2012.

Wang, A. C., Wang, Y.-Y., and Chen, M.-C. Single-blind, randomized trial of pelvic floor muscle training, biofeedback-assisted pelvic floor muscle training, and electrical stimulation in the management of overactive bladder. *Urology*, 63(1):61–66, Jan 2004.

Wang, A. C., Chih, S.-Y., and Chen, M.-C. Comparison of electric stimulation and oxybutynin chloride in management of overactive bladder with special reference to urinary urgency: a randomized placebo-controlled trial. *Urology*, 68(5):999–1004, 2006.

Wang, A. C., Chen, M.-C., Kuo, W.-Y., Lin, Y.-H., Wang, Y.-C., and Lo, T.-S. Urgency-free time interval as primary endpoint for evaluating the outcome of a randomized oab treatment. *International Urogynecology Journal*, 20(7): 819–825, 2009.

Warren, F. C., Abrams, K. R., and Sutton, A. J. Hierarchical network meta-analysis models to address sparsity of events and differing treatment classifications with regard to adverse outcomes. *Statistics in medicine*, 33(14): 2449–2466, 2014.

Watkins, D. S. Fundamentals of matrix computations john wiley & sons. *Inc., New York, NY*, 1991.

Wei, Y. and Higgins, J. Bayesian multivariate meta-analysis with multiple outcomes. *Statistics in medicine*, 32(17):2911–2934, 2013.

Welton, N. J., Caldwell, D., Adamopoulos, E., and Vedhara, K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions

in coronary heart disease. *American Journal of Epidemiology*, 169(9): 1158–1165, 2009.

Welton, N. J., Sutton, A. J., Ades, A., Cooper, N., and Abrams, K. R. *Evidence synthesis for decision making in healthcare*, volume 132. John Wiley & Sons, 2012.

White, I. R., Barrett, J. K., Jackson, D., and Higgins, J. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis Methods*, 3(2):111–125, 2012.

Whitehead, A. *Meta-analysis of controlled clinical trials*, volume 7. John Wiley & Sons, 2002.

Williamson, P. R., Altman, D. G., Blazeby, J. M., Clarke, M., and Gargon, E. The comet (core outcome measures in effectiveness trials) initiative. *Trials*, 12 (Suppl 1):A70, 2011.

Wong, M., Lam, K., and Lo, E. Bayesian analysis of clustered interval-censored data. *Journal of dental research*, 84(9):817–821, 2005.

Wu, T., Duan, X., Cao, C.-X., Peng, C.-D., Bu, S.-Y., and Wang, K.-J. The role of mirabegron in overactive bladder: a systematic review and meta-analysis. *Urologia internationalis*, 93(3):326–337, 2014.

Wyman, J. F., Fantl, J. A., McClish, D. K., Bump, R. C., et al. Comparative efficacy of behavioral interventions in the management of female urinary incontinence. *American journal of obstetrics and gynecology*, 179(4):999–1007, 1998.

Wyman, J., Burgio, K., and Newman, D. Practical aspects of lifestyle modifications and behavioural interventions in the treatment of overactive bladder and urgency urinary incontinence. *International journal of clinical practice*, 63(8):1177–1191, 2009.

Zimmern, P., Litman, H. J., Mueller, E., Norton, P., Goode, P., and , U. I. T. N. Effect of fluid management on fluid intake and urge incontinence in a trial for overactive bladder in women. *BJU Int*, 105(12):1680–1685, Jun 2010.

Zinner, N., Gittelman, M., Harris, R., Susset, J., Kanelos, A., and Auerbach, T. S. G., Stephen and. Trospium chloride improves overactive bladder symptoms: a multicenter phase iii trial. *J Urol*, 171(6 Pt 1):2311–5, Jun 2004.

Appendix A

Bayesian Methods for Clinical Trial Evaluation

A.1 Visual inspection of basic parametric distributions

FIGURE A.1: Visual inspection of the exponential model

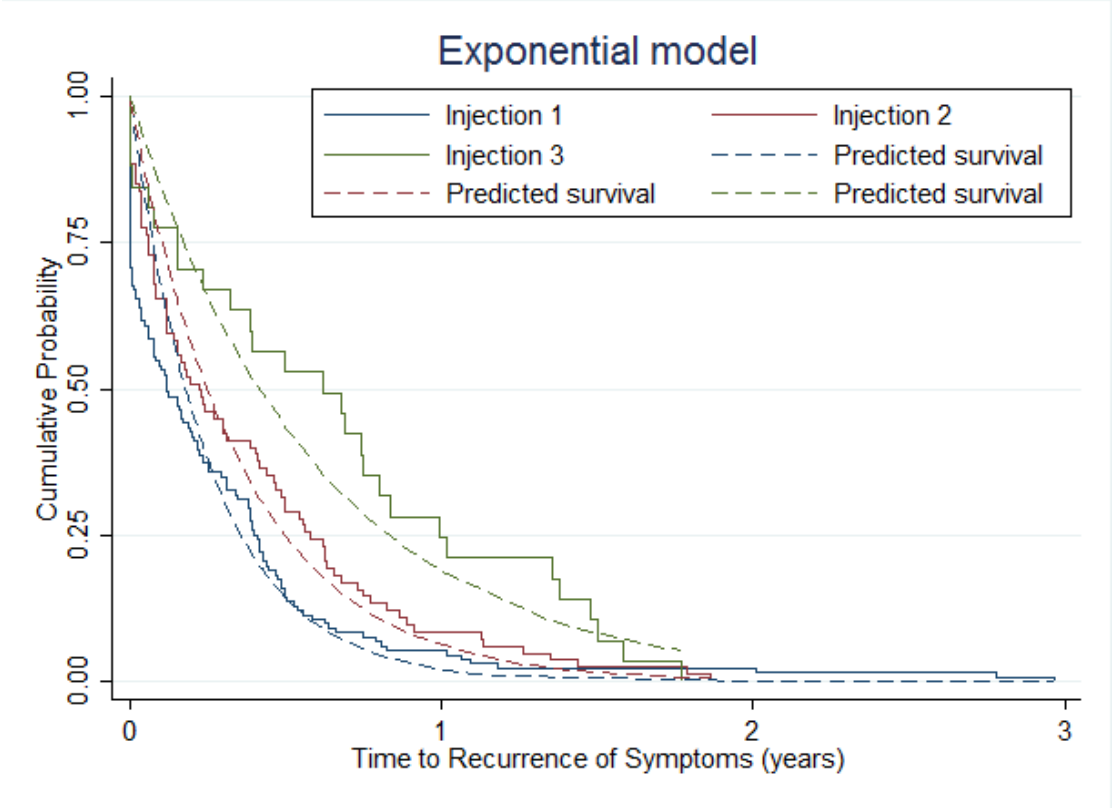


FIGURE A.2: Visual inspection of the generalised gamma model

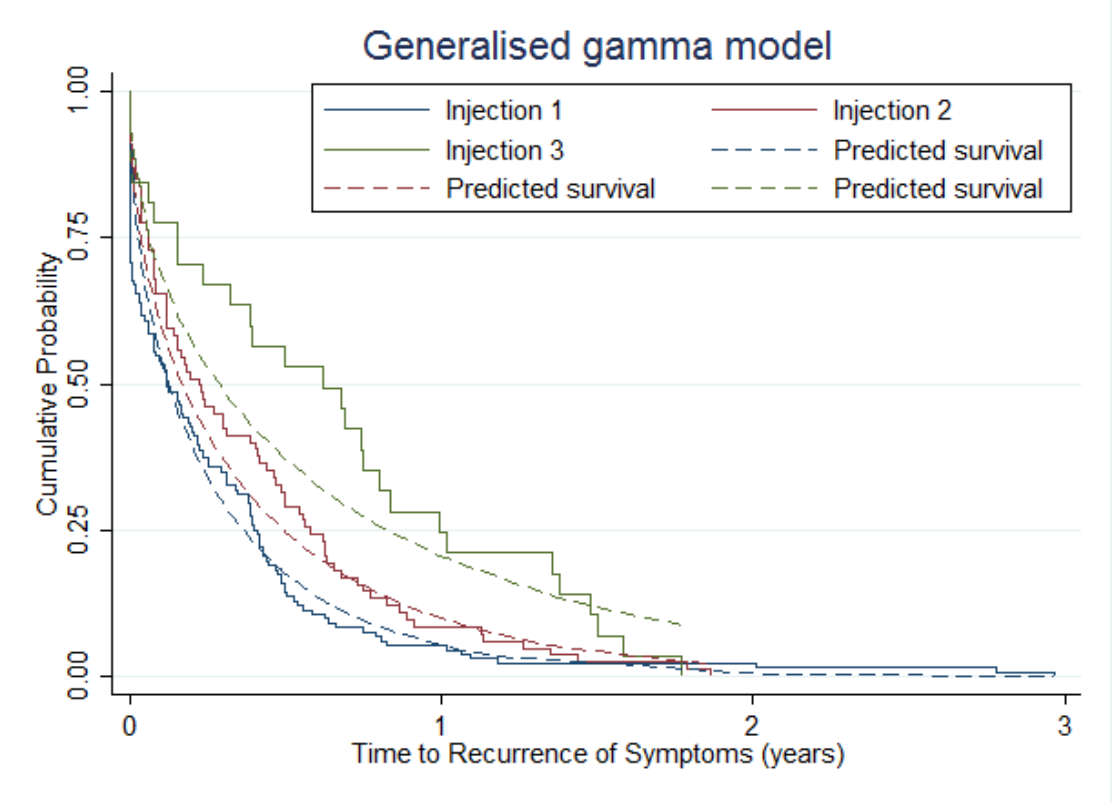


FIGURE A.3: Visual inspection of the log-normal model

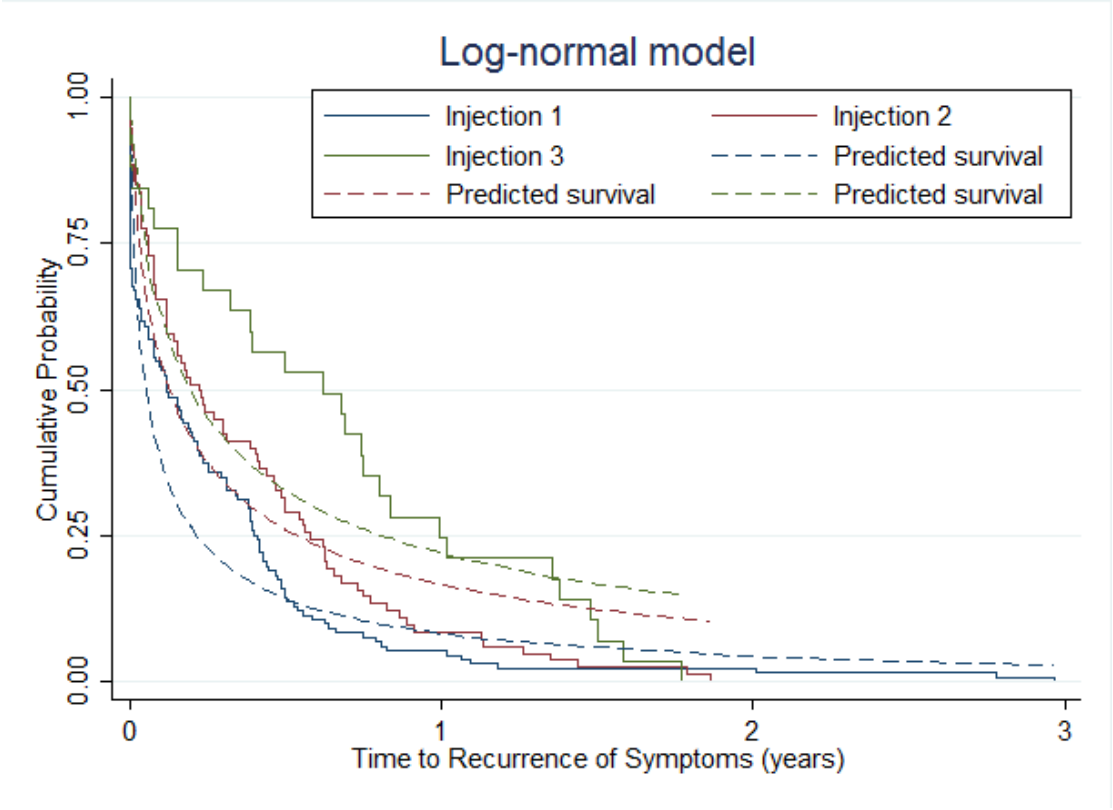


FIGURE A.4: Visual inspection of the Weibull model

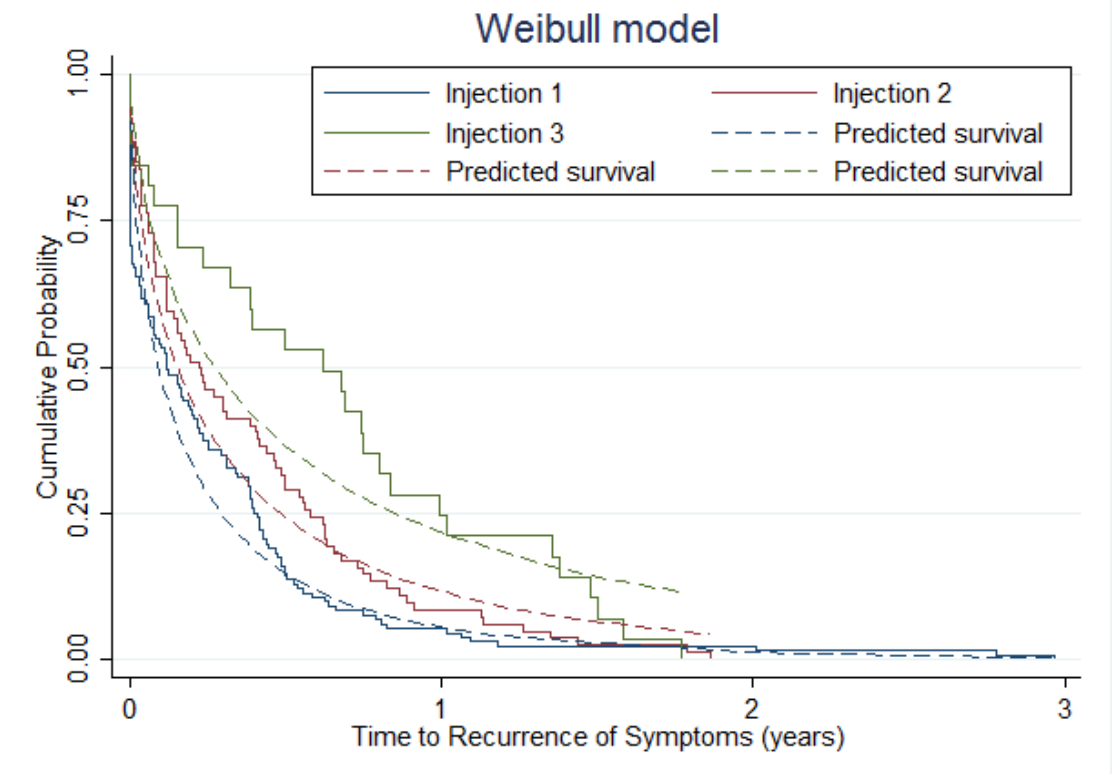
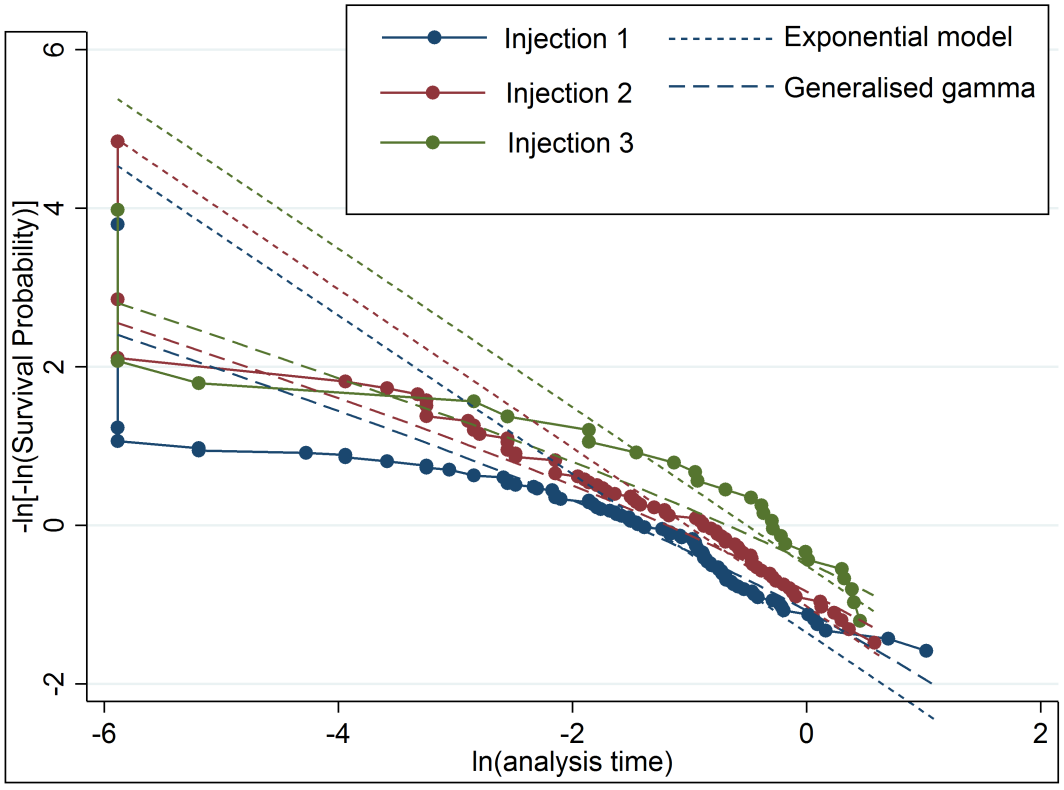


FIGURE A.5: Log-log plot of survival for the exponential and generalised gamma models



A.2 Example WinBUGS code - Model 0

```
#MODEL

model{
# loop through number of observations
        for (i in 1:N) {

# Truncated Weibull distribution
                t[i] ~ dweib(r,mu[i]) I(t.cen[i], nextinj[i])

# Exponent of the hazard model
                log(mu[i]) <- alpha + beta.strata[strata[i]] + b[id[i]]


# Frailty
        for (k in 1:M){
                b[k] ~ dnorm(0.0, tau)
                        }

# Reference group set to 0
        beta.strata[1]<-0


# Priors
        alpha ~ dnorm(0.0, 0.01)

        for(j in 2:3){
                beta.strata[j] ~ dnorm(0, 0.01)
                }


# Variance parameters on frailty
        sigma~ dunif(0,2)
        tau<-1/(sigma*sigma)

# Shape parameter of hazard distribution
        r ~ dgamma(1.0E-2, 1.0E-2)

# Calculating predictive survival times
        for (i in 1:31){
        for (j in 1:3){
        surv.pred[i,j] <- exp(-exp(alpha + beta.strata[j])*pow(t.surv[i], r))
        }
        }
}

#DATA list(N=442, M=228,
t.surv=c(0,0.001,0.005,0.01,0.02,0.05,0.1,0.2,0.4,0.6,0.8,1,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2
,2.1,2.2,2.3,2.4,2.5,2.6,2.7,2.8,2.9,3),
t=c( 0.006, 0.461, 1.772, 0.381, NA, NA, NA, NA, 0.300, 0.078,NA, NA, 0.117, 1.581, NA,
NA, 0.003, 0.003, NA, NA,NA, 0.167, NA, 0.003, NA, NA, NA, 0.833, 0.444, NA,1.353,
```

0.003, NA, NA, NA, NA, NA, 0.083, NA, 0.003,0.019, 0.747, NA, 0.792, 0.889, NA, 0.114,
0.194, NA, 0.122, NA, 0.500, 1.861, 0.389, 0.769, 1.014, 0.003, NA, 0.003, NA,NA, 0.039,
0.028, NA, 0.117, 0.156, NA, NA, NA, 0.058, NA, 0.003, NA, NA, 0.019, 0.003, 0.003,
0.047, 0.625, 0.006, NA, 0.822, 1.500, 0.003, 0.653, 0.003, NA, NA, 0.006, NA,NA, 0.003,
0.003, 0.003, NA, 0.058, 0.003, NA, 0.500, 0.197,0.417, NA, 0.181, 1.475, NA, 0.417, 0.175,
0.058, NA, 0.058,NA, NA, NA, 0.028, 0.631, 0.322, NA, 0.117, 0.075, 0.078,0.003, NA,
0.003, NA, NA, 0.403, 0.386, 0.161, 0.117, 0.417,0.681, NA, 0.236, 0.039, 1.014, NA, 0.003,
0.078, 1.378, 0.003, 0.039, NA, 0.036, 0.750, 1.436, 0.167, 0.117, NA, NA, NA, NA, NA,
NA, NA, NA, NA, 0.003, 0.003, NA, NA, 0.428, 0.500, NA, 0.003, 0.039,NA, NA, NA,
0.394, 0.003, NA, 0.078, NA, NA, 0.028,0.003, 0.639, 0.272, 0.003, 0.003, NA, NA,
0.803,0.156, NA, 0.233, 0.100, 0.003, 0.692, 0.292, 0.442, NA, 0.219, 0.056, NA, NA, NA,
0.467, 0.622, 0.003, NA,1.094, NA, NA, 0.003, NA, NA, 0.525, 0.003,0.003, 0.003, 0.003,
NA, 0.039, NA, NA, 0.386, 0.392, NA,0.747, 0.078, NA, 0.822, 2.011, 0.731, 1.178, NA,NA,
0.658, 0.564, NA, 0.439, 0.222, 0.156, NA, 0.019, 0.003,NA, 0.425, NA, NA, 0.003, 0.003,
NA, 0.228, 0.003, 0.003, 0.003,NA, 0.003, 0.019, 0.156, NA, 0.911, 0.500, 0.472, NA,
0.097,0.003, NA, 0.003, NA, NA, 0.003, 0.003, 0.222, NA, NA,NA, 0.003, 0.003, NA, NA,
NA, NA,0.039, 0.864, NA, 0.542, 0.311, 0.117, 0.389, 0.006, 2.778, NA,NA, 0.078, 0.233,
0.219, 0.411, NA, 0.586, 1.125, NA,0.414, 0.411, 1.133, 0.003, 0.142, NA, 0.619, NA, 1.264,
0.117,NA, 0.003, 0.003, NA, NA, 1.786, 0.003, NA, 0.242, NA, 0.628, NA, 0.172, NA,0.003,
0.003, NA, NA, 0.536, 0.483, 0.003, 0.556, 0.003, NA,0.156, NA),

⋮

id=c(1.000, 1.000, 1.000, 2.000, 2.000, 3.000, 4.000, 5.000, 5.000, 6.000,6.000, 7.000, 7.000,
7.000, 8.000, 8.000, 9.000, 10.000, 11.000, 11.000, 12.000, 12.000, 13.000, 14.000, 14.000,
15.000, 15.000, 15.000, 16.000, 16.000, 16.000, 17.000, 17.000, 18.000, 19.000, 19.000,
20.000, 20.000, 20.000, 21.000, 22.000, 22.000, 23.000, 24.000, 24.000, 25.000, 26.000,
26.000, 27.000, 28.000, 28.000, 29.000, 29.000, 30.000, 30.000, 30.000, 31.000, 31.000,
32.000, 32.000, 32.000, 33.000, 34.000, 35.000, 36.000, 36.000, 36.000, 37.000, 38.000,
38.000, 39.000, 39.000, 40.000, 41.000, 41.000, 42.000, 42.000, 43.000, 43.000, 43.000,
44.000, 44.000, 44.000, 45.000, 45.000, 46.000, 46.000, 47.000, 48.000, 48.000, 48.000,
49.000, 49.000, 49.000, 50.000, 50.000, 51.000, 52.000, 52.000, 53.000, 53.000, 54.000,
54.000, 54.000, 55.000, 56.000, 56.000, 57.000, 58.000, 59.000, 59.000, 59.000, 60.000,
61.000, 61.000, 61.000, 62.000, 62.000, 63.000, 63.000, 64.000, 64.000, 64.000, 65.000,
66.000, 66.000, 66.000, 67.000, 67.000, 68.000, 68.000, 68.000, 69.000, 69.000, 70.000,
71.000, 72.000, 73.000, 73.000, 74.000, 74.000, 75.000, 75.000, 75.000, 76.000, 76.000,
76.000, 77.000, 78.000, 78.000, 79.000, 80.000, 80.000, 80.000, 81.000, 81.000, 81.000,
82.000, 82.000, 82.000, 83.000, 83.000, 83.000, 84.000, 84.000, 84.000, 85.000, 85.000,
85.000, 86.000, 86.000, 87.000, 87.000, 88.000, 88.000, 89.000, 89.000, 89.000, 90.000,
90.000, 91.000, 91.000, 92.000, 93.000, 93.000, 93.000, 94.000, 94.000, 94.000, 95.000,
95.000, 96.000, 96.000, 97.000, 97.000, 98.000, 98.000, 99.000, 100.000, 100.000, 101.000,
101.000, 102.000, 103.000, 103.000, 104.000, 105.000, 105.000, 106.000, 106.000, 107.000,
108.000, 108.000, 108.000, 109.000, 109.000, 109.000, 110.000, 111.000, 111.000, 111.000,
112.000, 112.000, 113.000, 113.000, 114.000, 114.000, 115.000, 115.000, 115.000,116.000,
116.000, 117.000, 117.000, 117.000, 118.000, 118.000, 118.000, 119.000, 119.000,120.000,
120.000, 121.000, 121.000, 122.000, 123.000, 123.000, 124.000, 125.000, 125.000,126.000,
127.000, 127.000, 128.000, 128.000, 128.000, 129.000, 130.000, 130.000, 131.000,132.000,
133.000, 133.000, 134.000, 134.000, 135.000, 135.000, 136.000, 136.000, 137.000,137.000,
138.000, 138.000, 139.000, 140.000, 141.000, 141.000, 142.000, 142.000, 142.000, 143.000,
144.000, 144.000, 144.000, 145.000, 145.000, 145.000, 146.000, 146.000, 147.000,147.000,
147.000, 148.000, 148.000, 148.000, 149.000, 149.000, 150.000, 150.000, 150.000,151.000,
152.000, 153.000, 153.000, 154.000, 155.000, 156.000, 157.000, 158.000, 159.000,160.000,
160.000, 161.000, 161.000, 161.000, 162.000, 162.000, 163.000, 163.000, 163.000, 164.000,
165.000, 165.000, 166.000, 167.000, 168.000, 168.000, 169.000, 169.000, 170.000, 170.000,
170.000, 171.000, 172.000, 173.000, 174.000, 174.000, 175.000, 175.000, 176.000, 176.000,
177.000, 177.000, 178.000, 179.000, 179.000, 180.000, 180.000, 181.000, 182.000, 183.000,
183.000, 183.000, 184.000, 185.000, 185.000, 186.000, 186.000, 186.000, 187.000, 187.000,

188.000, 189.000, 189.000, 190.000, 191.000, 191.000, 191.000, 192.000, 192.000, 192.000, 193.000, 194.000, 194.000, 195.000, 195.000, 195.000, 196.000, 197.000, 198.000, 198.000, 199.000, 199.000, 199.000, 200.000, 200.000, 201.000, 201.000, 201.000, 202.000, 202.000, 203.000, 204.000, 204.000, 204.000, 205.000, 205.000, 206.000, 207.000, 207.000, 207.000, 208.000, 209.000, 210.000, 210.000, 210.000, 211.000, 212.000, 212.000, 212.000, 213.000, 213.000, 214.000, 214.000, 215.000, 215.000, 215.000, 216.000, 216.000, 217.000, 217.000, 218.000, 218.000, 219.000, 219.000, 219.000, 220.000, 220.000, 221.000, 221.000, 221.000, 222.000, 222.000, 223.000, 224.000, 224.000, 225.000, 226.000, 227.000, 227.000, 228.000, 228.000))

#INITS list(alpha = 0, beta.strata =c(NA,0,0), r=1, sigma = 1,
t=c(NA, NA, NA, NA, 0.003, 0.003, 0.003, 0.003, NA, NA, 0.003, 0.003, NA, NA, 0.003, 0.003, NA, NA, 0.003, 0.003,0.003, NA, 0.003, NA, 0.003, 0.003, 0.003, NA, NA, 0.003,NA, NA, 0.003, 0.003, 0.003, 0.003, 0.003, NA, 0.003, NA,NA, NA, 0.003, NA, NA, 0.003, NA, NA, 0.003, NA, 0.003, 0.003, 0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA, NA, NA, 0.003, NA,0.003, 0.003, 0.003, NA, NA, NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, 0.003, NA, 0.003, NA, 0.003, 0.003, NA, NA, NA, NA, NA, NA, 0.003, NA, NA, NA, 0.003, 0.003, NA, 0.003,0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA, NA, NA, NA, 0.003, NA, NA, NA, NA,NA, 0.003, NA, 0.003, 0.003, 0.003, NA, NA, NA, 0.003,NA, 0.003, 0.003, NA, 0.003, 0.003, NA, NA, NA, NA, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003,0.003, NA, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA,NA, 0.003, NA, 0.003, 0.003, NA, NA, NA, NA, NA,NA, 0.003, NA, NA, NA, NA, NA, NA,0.003, 0.003, 0.003, NA, NA, 0.003, NA, 0.003, 0.003, NA,NA, NA, NA, NA, NA, 0.003, 0.003, NA,NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, NA, NA, 0.003, NA, NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003, 0.003, NA, 0.003, NA, 0.003, NA, NA, NA, 0.003, NA, NA, NA, NA, NA, 0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, NA, 0.003, 0.003,0.003, NA, NA, 0.003, 0.003, 0.003, NA, NA, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003,NA, NA, 0.003, 0.003, NA, 0.003, 0.003, 0.003, 0.003, 0.003,0.003, NA, NA, 0.003, 0.003, 0.003, NA, 0.003, NA, 0.003,NA, NA, 0.003, 0.003, NA, NA, NA, NA, NA, 0.003,NA, 0.003),
b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))

A.3 Example WinBUGS code - Model 1

```
#MODEL

model{
# loop through number of observations
        for (i in 1:N) {

# Truncated poly-Weibull distribution
        t[i] ~ nd.dpolyweib(3,r[],mu[,i]) I(t.cen[i], nextinj[i])

# Exponent of the hazard model
        for(j in 1:3){
        log(mu[j,i]) <- beta0[j] + beta.strata[j,strata[i]] +b[id[i]]
        }
        }

for(j in 1:3){
        beta.strata[j,1]<-0

        # Priors
        beta0[j] ~ dnorm(0.0, 0.01)

        for(k in 2:3){
                beta.strata[j,k] ~ dnorm(0, 0.01)
                }
        }

for (k in 1:M){
        b[k] ~ dnorm(0.0, tau)
        }

sigma~ dunif(0,2)
tau<-1/(sigma*sigma)

# Shape parameter of hazard distribution

log(r[1]) <- log.r1
log.r1 ~ dnorm(0,1)

log(r[2]) <- log.r2
log.r2~dnorm(0,0.01)

log(r[3]) <- log.r3
log.r3~dnorm(0,0.01)

# Constraints on shape parameters
d<-1
d~dbern(gamma)
gamma<-step(r[2]-r[1])

d1<-1
d1~dbern(gamma1)
gamma1<-step(r[3]-r[2])


# Calculating predictive survival
for (i in 2:31){
```

```
for (j in 1:3){
        surv.pred[i,j] <- exp(-(exp(alpha[1] +  beta.strata[1,j] )*pow(t.surv[i], r[1]) + exp(alpha[2] +
beta.strata[2,j] )*pow(t.surv[i], r[2] )+ exp(alpha[3] +  beta.strata[3,j] )*pow(t.surv[i], r[3] )))
                }
S1[i] <- surv.pred[i,1]
S2[i] <- surv.pred[i,2]
S3[i] <- surv.pred[i,3]
}
S1[1] <-0
S2[1] <-0
S3[1] <-0

for(i in 1:30){
dS1[i]<- (S1[i] + S1[i+1])*(0.5)*(0.1)
dS2[i]<- (S2[i] + S2[i+1])*(0.5)*(0.1)
dS3[i]<- (S3[i] + S3[i+1])*(0.5)*(0.1)
}

AUC1 <-sum(dS1[])
AUC2<- sum(dS2[])
AUC3<-sum(dS3[])

diff21 <-AUC2-AUC1
diff32<-AUC3-AUC2

prob21 <- step(diff21) # 1 if AUC2 >= AUC1
prob32 <- step(diff32)
}
```

#DATA

```
list(N=442, M=228,
t.surv=c(0,0.001,0.005,0.01,0.02,0.05,0.1,0.2,0.4,0.6,0.8,1,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2,2.1,2.2,2.3,2.4,2.5,2.6,2.7,2.8,2.9,3),
t=c( 0.006,  0.461,  1.772,  0.381,  NA,  NA,  NA,  NA,  0.300,  0.078, NA,  NA,  0.117,  1.581,  NA,  NA,
0.003,  0.003,  NA,  NA, NA,  0.167,  NA,  0.003,  NA,  NA,  NA,  0.833,  0.444,  NA, 1.353,  0.003,  NA,
NA,  NA,  NA,  NA,  0.083,  NA,  0.003, 0.019,  0.747,  NA,  0.792,  0.889,  NA,  0.114,  0.194,  NA,  0.122,
NA,  0.500,  1.861,  0.389,  0.769,  1.014,  0.003,  NA,  0.003,  NA, NA,  0.039,  0.028,  NA,  0.117,  0.156,  NA,
NA,  NA,  0.058, NA,  0.003,  NA,  NA,  0.019,  0.003,  0.003,  0.047,  0.625,  0.006, NA,  0.822,  1.500,  0.003, 0.653,
0.003,  NA,  NA,  0.006,  NA, NA,  0.003,  0.003,  0.003,  NA,  0.058,  0.003,  NA,  0.500,  0.197, 0.417,  NA,
0.181,  1.475,  NA,  0.417,  0.175,  0.058,  NA,  0.058, NA,  NA,  NA,  0.028,  0.631,  0.322,  NA,  0.117,  0.075,
0.078, 0.003,  NA,  0.003,  NA,  NA,  0.403,  0.386,  0.161,  0.117,  0.417, 0.681,  NA,  0.236,  0.039,  1.014,  NA,
0.003,  0.078,  0.003,  0.039, NA,  NA,  NA,  0.394,  0.003,  NA,  0.078,  NA,  NA,  0.028, 0.003,  0.639,  0.272,
1.378,  0.003,  0.003,  0.003,  NA,  NA,  0.803, 0.156,  NA,  0.233,  0.100,  0.003,  0.692,  0.292,  0.442,  NA,  0.219,
0.039,  NA,  0.056,  NA,  NA,  NA,  0.467,  0.622,  0.003,  NA, 1.094,  NA,  NA,  0.003,  NA,  NA,  0.525,
0.036,  0.750,  0.003, 0.003,  0.003,  0.003,  NA,  0.039,  NA,  NA,  0.386,  0.392,  NA, 0.747,  0.078,  NA,  0.822,
1.436,  0.167,  2.011,  0.731,  1.178,  NA, NA,  0.658,  0.564,  NA,  0.439,  0.222,  0.156,  NA,  0.019,  0.003, NA,
0.117,  NA,  NA,  NA,  NA,  NA,  NA,  NA,  NA, NA,  0.003,  0.003,  NA,  NA,  0.428,  0.500,  NA,
0.425,  NA, 0.039,  0.864,  NA,  0.542,  0.311,  0.117,  0.389,  0.006,  2.778,  NA, NA,  0.078,  0.233,  0.219,  0.411,
NA,  NA,  0.586,  1.125,  NA, 0.414,  0.411,  1.133,  0.003,  0.142,  NA,  0.619,  NA,  1.264,  0.117, NA,  0.003,
NA,  0.003,  0.003,  NA,  0.228,  0.003,  0.003,  0.003, NA,  0.003,  0.019,  0.156,  NA,  0.911,  0.500,  0.472,  NA,
0.097, 0.003,  NA,  0.003,  NA,  NA,  0.003,  0.003,  0.222,  NA,  NA, NA,  0.003,  0.003,  NA,  NA,  NA,
0.003,  NA,  0.311,  0.003, 0.378,  NA,  0.058,  NA,  NA,  0.003,  0.078,  NA,  NA,  NA, NA,  0.500,  NA,
NA,  NA,  0.389,  NA,  0.003,  0.019,  0.003, 0.150,  0.003,  NA,  2.964,  NA,  0.006,  0.003,  NA,  NA,  0.344,
NA,  NA,  NA,  NA,  0.400,  0.625,  0.003,  NA,  0.250,  0.014, 0.083,  0.117,  NA,  0.167,  0.503,  1.347,  NA,
1.786,  NA,  NA, 0.083,  0.003,  NA,  0.308,  NA,  0.233,  NA,  0.058,  0.339,  NA, 0.744,  0.003,  0.003,  NA,
0.003,  NA,  NA,  1.064,  NA,  0.311, 0.039,  NA,  NA,  0.992,  NA,  NA,  0.483,  0.061,  NA,  0.486, 0.297,
NA,  NA,  0.117,  NA,  0.194,  NA,  0.808,  0.003,  0.581, 0.681,  0.003,  0.250,  0.469,  NA,  NA,  0.003,  0.003,
0.242,  NA, 0.394,  0.556,  NA,  NA,  0.206,  NA,  NA,  NA,  NA,  NA, NA,  0.186,  0.078,  NA,  NA,
NA,  0.628,  NA,  0.172,  NA, 0.003,  0.003,  NA,  NA,  0.536,  0.483,  0.003,  0.556,  0.003,  NA, 0.156,  NA),
                                                                    ⋮
id=c( 1.000,  1.000,  1.000,  2.000,  2.000,  3.000,  4.000,  5.000,  5.000,  6.000, 6.000,  7.000,  7.000,  7.000,  8.000,  8.000,
9.000,  10.000,  11.000,  11.000, 12.000,  12.000,  13.000,  14.000,  14.000,  15.000,  15.000,  15.000,  16.000,  16.000, 16.000,
17.000,  17.000,  18.000,  19.000,  19.000,  20.000,  20.000,  20.000,  21.000, 22.000,  22.000,  23.000,  24.000,  24.000,  25.000,
```

```
26.000, 26.000, 27.000, 28.000, 28.000, 29.000, 29.000, 30.000, 30.000, 30.000, 31.000, 31.000, 32.000, 32.000, 32.000,
33.000, 34.000, 35.000, 36.000, 36.000, 36.000, 37.000, 38.000, 38.000, 39.000, 39.000, 40.000, 41.000, 41.000, 42.000,
42.000, 43.000, 43.000, 43.000, 44.000, 44.000, 44.000, 45.000, 45.000, 46.000, 46.000, 47.000, 48.000, 48.000, 48.000,
49.000, 49.000, 49.000, 50.000, 50.000, 51.000, 52.000, 52.000, 53.000, 53.000, 54.000, 54.000, 54.000, 55.000, 56.000,
56.000, 57.000, 58.000, 59.000, 59.000, 59.000, 60.000, 61.000, 61.000, 61.000, 62.000, 62.000, 63.000, 63.000, 64.000,
64.000, 64.000, 65.000, 66.000, 66.000, 66.000, 67.000, 67.000, 68.000, 68.000, 68.000, 69.000, 69.000, 70.000, 71.000,
72.000, 73.000, 73.000, 74.000, 74.000, 75.000, 75.000, 75.000, 76.000, 76.000, 76.000, 77.000, 78.000, 78.000, 79.000,
80.000, 80.000, 80.000, 81.000, 81.000, 81.000, 82.000, 82.000, 82.000, 83.000, 83.000, 83.000, 84.000, 84.000, 84.000,
85.000, 85.000, 85.000, 86.000, 86.000, 87.000, 87.000, 88.000, 88.000, 89.000, 89.000, 89.000, 90.000, 90.000, 91.000,
91.000, 92.000, 93.000, 93.000, 93.000, 94.000, 94.000, 94.000, 95.000, 95.000, 96.000, 96.000, 97.000, 97.000, 98.000,
98.000, 99.000, 100.000, 100.000, 101.000, 101.000, 102.000, 103.000, 103.000, 104.000, 105.000, 105.000, 106.000, 106.000,
107.000, 108.000, 108.000, 108.000, 109.000, 109.000, 109.000, 110.000, 111.000, 111.000, 111.000, 112.000, 112.000, 113.000,
113.000, 114.000, 114.000, 115.000, 115.000, 115.000,116.000, 116.000, 117.000, 117.000, 117.000, 118.000, 118.000, 118.000,
119.000, 119.000,120.000, 120.000, 121.000, 121.000, 122.000, 123.000, 123.000, 124.000, 125.000, 125.000,126.000, 127.000,
127.000, 128.000, 128.000, 128.000, 129.000, 130.000, 130.000, 131.000,132.000, 133.000, 133.000, 134.000, 134.000, 135.000,
135.000, 136.000, 136.000, 137.000,137.000, 138.000, 138.000, 139.000, 140.000, 141.000, 141.000, 142.000, 142.000, 142.000,
143.000, 144.000, 144.000, 144.000, 145.000, 145.000, 145.000, 146.000, 146.000, 147.000,147.000, 147.000, 148.000, 148.000,
148.000, 149.000, 149.000, 150.000, 150.000, 150.000,151.000, 152.000, 153.000, 153.000, 154.000, 155.000, 156.000, 157.000,
158.000, 159.000,160.000, 160.000, 161.000, 161.000, 161.000, 162.000, 162.000, 163.000, 163.000, 163.000, 164.000, 165.000,
165.000, 166.000, 167.000, 168.000, 168.000, 169.000, 169.000, 170.000, 170.000, 170.000, 171.000, 172.000, 173.000, 174.000,
174.000, 175.000, 175.000, 176.000, 176.000, 177.000, 177.000, 178.000, 179.000, 179.000, 180.000, 180.000, 181.000, 182.000,
183.000, 183.000, 183.000, 184.000, 185.000, 185.000, 186.000, 186.000, 186.000, 187.000, 187.000, 188.000, 189.000, 189.000,
190.000, 191.000, 191.000, 191.000, 192.000, 192.000, 192.000, 193.000, 194.000, 194.000, 195.000, 195.000, 195.000, 196.000,
197.000, 198.000, 198.000, 199.000, 199.000, 199.000, 200.000, 200.000, 201.000, 201.000, 201.000, 202.000, 202.000, 203.000,
204.000, 204.000, 204.000, 205.000, 205.000, 206.000, 207.000, 207.000, 207.000, 208.000, 209.000, 210.000, 210.000, 210.000,
211.000, 212.000, 212.000, 212.000, 213.000, 213.000, 214.000, 214.000, 215.000, 215.000, 215.000, 216.000, 216.000, 217.000,
217.000, 218.000, 218.000, 219.000, 219.000, 219.000, 220.000, 220.000, 221.000, 221.000, 221.000, 222.000, 222.000, 223.000,
224.000, 224.000, 225.000, 226.000, 227.000, 227.000, 228.000, 228.000)
)
```

#INITS

```
list( beta0 = c(0,0,0), beta.strata = structure(.Data = c(NA,0,0,NA,0,0, NA,0,0),.Dim = c(3,3)), log.r1 = 0, log.r2 = 0, log.r3 = 0,
sigma = 1, t=c( NA, NA, NA, NA, 0.003, 0.003, 0.003, 0.003, NA, NA, 0.003, 0.003, NA, NA, 0.003,
0.003, NA, NA, 0.003, 0.003, 0.003, NA, 0.003, NA, 0.003, 0.003, 0.003, NA, NA, 0.003, NA, NA,
0.003, 0.003, 0.003, 0.003, 0.003, NA, 0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA, NA, 0.003,
NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA,
0.003, 0.003, 0.003, NA, 0.003, NA, 0.003, 0.003, NA, NA, NA, NA, NA, NA, 0.003, NA, NA,
NA, NA, NA, 0.003, 0.003, NA, 0.003, 0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA, NA,
NA, 0.003, NA, NA, 0.003, NA, NA, NA, 0.003, NA, 0.003, 0.003, 0.003, NA, NA, NA, 0.003,
NA, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, NA, NA, NA, NA, 0.003, NA, NA, NA,
0.003, NA, NA, NA, NA, 0.003, 0.003, 0.003, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, NA,
NA, NA, NA, NA, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003,
NA, NA, 0.003, NA, 0.003, 0.003, 0.003, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, 0.003, 0.003, NA,
NA, NA, NA, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, 0.003, NA, NA,
NA, NA, NA, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, NA, 0.003, NA, NA, 0.003, NA,
0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, NA, NA, 0.003, 0.003, NA, NA, 0.003, NA,
0.003, NA, NA, 0.003, NA, NA, NA, NA, NA, NA, 0.003, 0.003, NA, NA, NA, NA, 0.003,
0.003, NA, NA, 0.003, NA, NA, NA, NA, NA, 0.003, NA, 0.003, NA, NA, 0.003, NA, 0.003,
NA, NA, 0.003, NA, NA, NA, NA, 0.003, NA, NA, NA, 0.003, NA, NA, NA, 0.003, NA,
NA, 0.003, NA, 0.003, 0.003, NA, NA, NA, 0.003, 0.003, 0.003, NA, NA, 0.003, 0.003, 0.003, NA,
0.003, NA, NA, NA, 0.003, NA, 0.003, 0.003, NA, NA, 0.003, 0.003, 0.003, 0.003, NA, 0.003, 0.003,
0.003, NA, 0.003, NA, NA, NA, NA, NA, 0.003, NA, 0.003, NA, NA, 0.003, 0.003, NA, 0.003,
0.003, 0.003, 0.003, NA, NA, NA, 0.003, NA, NA, NA, NA, 0.003, NA, NA, NA, 0.003, NA,
0.003, 0.003, NA, NA, 0.003, NA, 0.003, NA, 0.003, NA, NA, 0.003, NA, NA, NA, 0.003, NA,
0.003, 0.003, NA, 0.003, NA, NA, 0.003, 0.003, NA, 0.003, 0.003, NA, NA, 0.003, NA, NA, 0.003,
0.003, NA, 0.003, NA, 0.003, NA, NA, NA, NA, NA, NA, NA, 0.003, 0.003, NA, NA, NA,
0.003, NA, NA, 0.003, 0.003, NA, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, NA, NA, 0.003, 0.003, 0.003,
NA, 0.003, NA, 0.003, NA, NA, 0.003, 0.003, NA, NA, NA, NA, NA, 0.003, NA, 0.003),
b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))
```

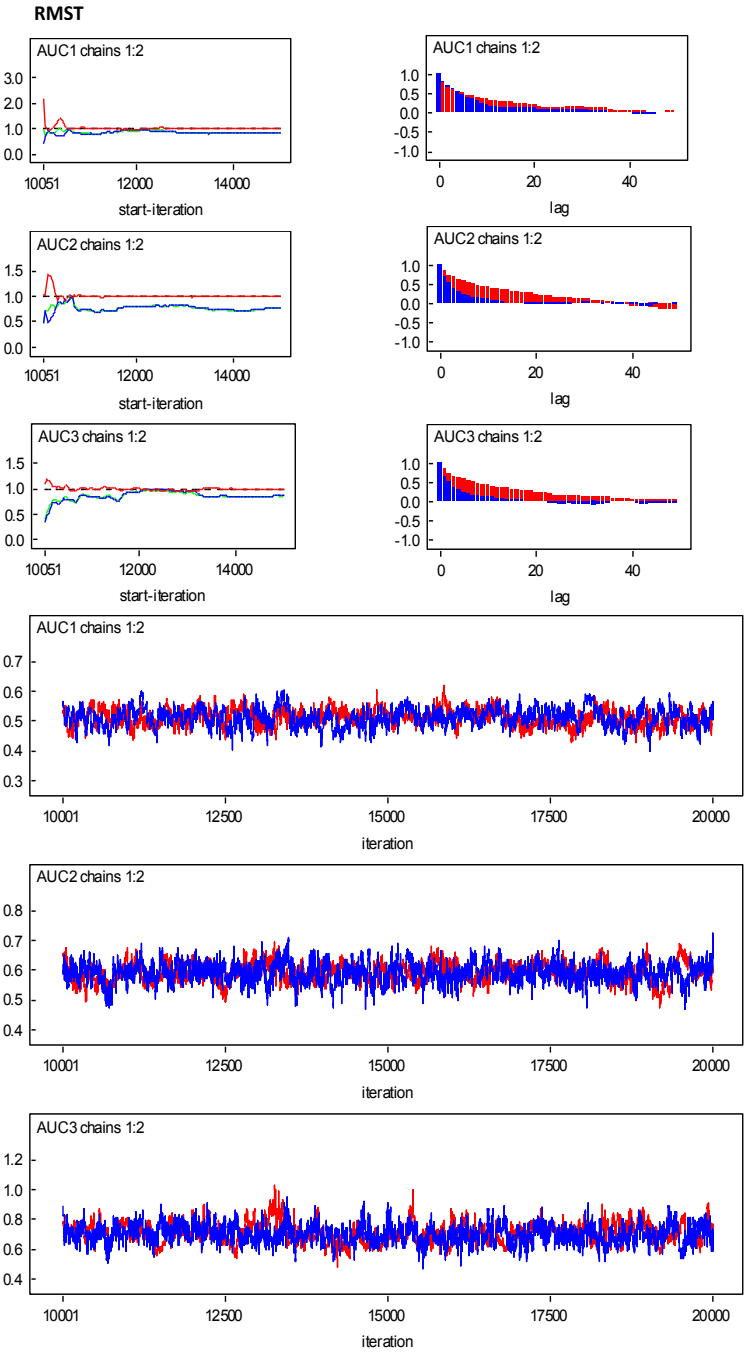
A.4 Example WinBUGS code - Model 2

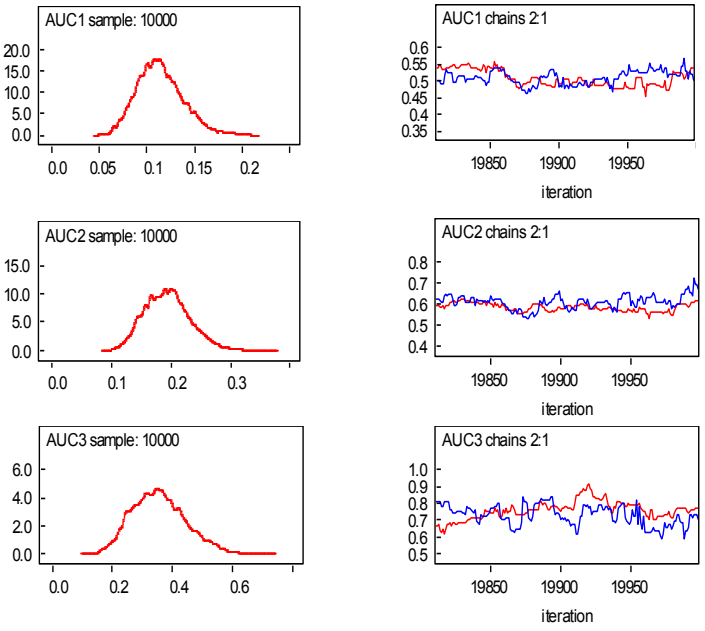
```
        model
        {
# loop through number of observations
                for (i in 1:N) {


# Truncated poly-Weibull distribution
                t[i] ~ nd.dpolyweib(3,r[,i],mu[,i]) I(t.cen[i], nextinj[i])

# Exponent of the hazard model
for(j in 1:3){
        log(mu[j,i]) <- beta0[j] + beta.strata[j,strata[i]] + b[id[i]]
        log(r[j,i]) <- alpha0[j] + alpha1[j]*X1[i] + alpha2[j]*X2[i]
}
}

for(j in 1:3){
                beta.strata[j,1]<-0

        # Priors
                beta0[j] ~ dnorm(0.0, 0.01)

                alpha0[j] ~ dnorm(0.0, 0.01)
                alpha1[j] ~ dnorm(0.0, 0.01)
                alpha2[j] ~ dnorm(0.0, 0.01)

                for(k in 2:3){
                        beta.strata[j,k] ~ dnorm(0, 0.01)
                        }
        }

for(k in 1:M){
        b[k] ~ dnorm(0.0, tau)
        }

sigma~ dunif(0, 2)
tau<-1/(sigma*sigma)


# Constraints on the shape parameters of hazard distribution

v2[1:2] <- minmax(w[2,])
v1[1:2] <- minmax(w[1,])
v3[1:2] <- minmax(w[3,])
```

```
C11 <- step((v2[1])-(v1[1]))
C21 <- step((v2[2])-(v1[2]))

C12 <- step((v3[1])-(v2[1]))
C22 <- step((v3[2])-(v2[2]))


for(j in 1:3){
w[j,1] <- alpha0[j]
w[j,2] <- alpha0[j] + alpha1[j]
w[j,3] <- alpha0[j] + alpha2[j]
}

b<-1
b~ dbern(gamma)
gamma<- C11*C21

b1<-1
b1~ dbern(gamma1)
gamma1<- C12*C22

log(rstar11)<- alpha0[1]
log(rstar12)<- alpha0[1]+alpha1[1]
log(rstar13)<-alpha0[1]+alpha1[1]+alpha2[1]

log(rstar21)<- alpha0[2]
log(rstar22)<- alpha0[2]+alpha1[2]
log(rstar23)<-alpha0[2]+alpha1[2]+alpha2[2]

log(rstar31)<- alpha0[3]
log(rstar32)<- alpha0[3]+alpha1[3]
log(rstar33)<-alpha0[3]+alpha1[3]+alpha2[3]

# Calculating predictive survival times
                for (i in 2:31){
                surv.pred[i,1] <- exp(-(exp(beta0[1] +  beta.strata[1,1]
)*pow(t.surv[i], rstar11) + exp(beta0[2] +  beta.strata[2,1] )*pow(t.surv[i], rstar21 )
+ exp(beta0[3] +  beta.strata[3,1] )*pow(t.surv[i], rstar31 )))
                surv.pred[i,2] <- exp(-(exp(beta0[1] +  beta.strata[1,2]
)*pow(t.surv[i], rstar12) + exp(beta0[2] +  beta.strata[2,2] )*pow(t.surv[i], rstar22 )
+ exp(beta0[3] +  beta.strata[3,1] )*pow(t.surv[i], rstar31 )))
                surv.pred[i,3] <- exp(-(exp(beta0[1] +  beta.strata[1,3]
)*pow(t.surv[i], rstar13) + exp(beta0[2] +  beta.strata[2,3] )*pow(t.surv[i], rstar23 )
+ exp(beta0[3] +  beta.strata[3,1] )*pow(t.surv[i], rstar31 )))

S1[i] <- surv.pred[i,1]
```

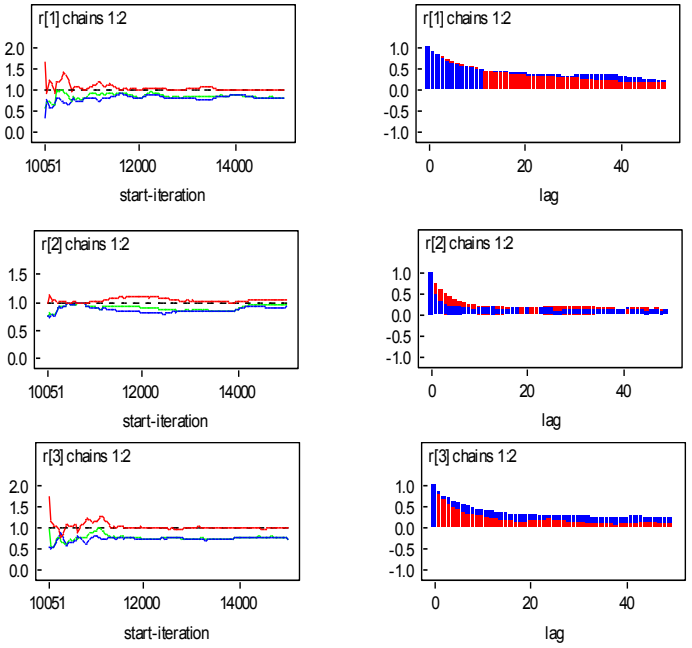
```
S2[i] <- surv.pred[i,2]
S3[i] <- surv.pred[i,3]
        }
S1[1] <-0
S2[1] <-0
S3[1] <-0

for(i in 1:30){
dS1[i]<- (S1[i] + S1[i+1])*(0.5)*(0.1)
dS2[i]<- (S2[i] + S2[i+1])*(0.5)*(0.1)
dS3[i]<- (S3[i] + S3[i+1])*(0.5)*(0.1)
}

AUC1 <-sum(dS1[])
AUC2<- sum(dS2[])
AUC3<-sum(dS3[])

diff21 <-AUC2-AUC1
diff32<-AUC3-AUC2

prob21<-step(diff21)
prob32<-step(diff32)
}
```

A.5 Convergence diagnostics





Shape parameters (γ)



Appendix B

Systematic Review of Clinical Trials in Overactive Bladder

B.1 Search Strategy

Search strategy (Medline)

1. exp controlled clinical trial/ or exp randomized controlled trial/
2. randomized.ab.
3. randomly.ab.
4. placebo.ab.
5. randomised.ab.
6. trial.ab.
7. 1 or 2 or 3 or 4 or 5 or 6
8. exp Urinary Incontinence/
9. exp Urinary Bladder, Overactive/
10. exp Urinary Incontinence, Urge/
11. urinary incontinence.tw.
12. urge incontinence.tw.
13. overactive bladder.tw.
14. detrusor overactivity.tw.
15. detrusor instability.tw.
16. detrusor hyperactivity.tw.
17. unstable detrusor contraction.tw.
18. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17
19. exp Cholinergic Antagonists/
20. exp Muscarinic Antagonists/
21. anticholinergic*.tw.
22. antimuscarinic*.tw.
23. exp Life Style/ or lifestyle change*.tw. or lifestyle modification*.tw.
24. exp Physical Therapy Modalities/ or physiotherap*.tw. or physcial therap*.tw. or pelvic floor exercise*.tw. or pelvic floor muscle training.tw. or pelvic floor*.tw.
25. exp Exercise Therapy/ or exercise therap*.tw.
26. exp Behavior Therapy/ or behav* therapy.tw.
27. exp Electric Stimulation/ or exp Electric Stimulation Therapy/ or electric* stimulation.tw. or neuromodulation.tw. or nerve stimulation.tw.
28. exp Drug Therapy/ or drug therap*.tw. or drug*.tw. or pharmacotherap*.tw.
29. exp botulinum toxins/ or exp botulinum toxins, type a/ or botulinum toxin*.tw. or bonta.tw. or BoNT*.tw. or BTX*.tw. or onabotulinum toxin*.tw.
30. exp Receptors, Adrenergic, beta/ or beta* adren* receptor*.tw.
31. exp Adrenergic beta-Agonists/
32. mirabegron.tw.
33. solabegron.tw.
34. desmopressin.tw.
35. duloxetine.tw.
36. oestrogen.tw.
37. 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36
38. 7 and 18 and 37
39. limit 38 to (humans and "all adult (19 plus years)" and humans and randomized controlled trial)

Search strategy (EMBASE)

1. exp controlled clinical trial/ or exp randomized controlled trial/
2. randomized.ab.
3. randomly.ab.
4. placebo.ab.
5. trial.ab.
6. randomised.ab.
7. 1 or 2 or 3 or 4 or 5 or 6
8. exp urinary bladder, overactive/ or exp urinary incontinence/ or exp urinary incontinence, urge/
9. urinary incontinence.tw.
10. urge incontinence.tw.
11. overactive bladder.tw.
12. detrusor overactivity.tw.
13. detrusor instability.tw.
14. detrusor hyperactivity.tw.
15. 8 or 9 or 10 or 11 or 12 or 13 or 14
16. exp cholinergic receptor blocking agent/ or anticholinergic*.tw.
17. exp muscarinic receptor blocking agent/ or antimuscarinic*.tw.
18. exp lifestyle modification/ or lifestyle change*.tw.
19. exp physiotherapy/ or physiotherap*.tw. or physical therap*.tw. or pelvic floor exercise*.tw. or pelvic floor muscle training.tw. or pelvic floor*.tw.
20. exp behavior therapy/ or behav* therapy.tw.
21. exp nerve stimulation/ or neuromodulation.tw. or electric* stimulation.tw. or nerve stimulation.tw.
22. exp drug therapy/ or drug therap*.tw. or drug*.tw. or pharmacotherap*.tw.
23. exp botulinum toxin/ or exp botulinum toxin a/ or botulinum toxin*.tw. or bonta.tw. or BoNT*.tw. or BTX*.tw. or onabotulinum toxin*.tw.
24. exp beta adrenergic receptor stimulating agent/ or exp mirabegron/ or exp solabegron/ or beta* adren* receptor*.tw. or mirabegron.tw. or solabegron.tw.
25. desmopressin.tw.
26. duloxetine.tw.
27. oestrogen.tw.
28. 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27
29. 7 and 15 and 28
30. limit 29 to humans
31. limit 30 to randomized controlled trial
32. limit 31 to human
33. limit 32 to (adult <18 to 64 years> or aged <65+ years>)
34. remove duplicates from 33

B.2 Study references

Table B.1: Study references

A1	Abdelbary AM, El-Dessoukey AA, Massoud AM, et al. Combined vaginal pelvic floor electrical stimulation (PFS) and local vaginal estrogen for treatment of overactive bladder (OAB) in perimenopausal females. randomized controlled trial (RCT). *Urology*. 2015;86(3):482-486.
A2	Abrams P, Freeman R, Anderstrom C, Mattiasson A. Tolterodine, a new antimuscarinic agent: As effective but better tolerated than oxybutynin in patients with an overactive bladder. *Br J Urol*. 1998;81(6):801-810.
A3	Anderson RC, Carnes M, Clark A, et al. Effects of terodiline on urinary incontinence among older non- institutionalized women. *J Am Geriatr Soc*. 1993;41(9):915-922.
A4	Anderson RU, Mobley D, Blank B, Saltzstein D, Susset J, Brown JS. Once daily controlled versus immediate release oxybutynin chloride for urge urinary incontinence. OROS oxybutynin study group. *J Urol*. 1999;161(6):1809-1812.
A5	Appell RA, Sand P, Dmochowski R, Anderson R, Zinner N. Prospective randomized controlled trial of extended-release oxybutynin chloride and tolterodine tartrate in the treatment of overactive bladder: Results of the OBJECT study. *Mayo Clin Proc*. 2001;76(4):358-363.
A6	Barkin J, Corcos J, Radomski S, et al. A randomized, double-blind, parallel-group comparison of controlled- and immediate-release oxybutynin chloride in urge urinary incontinence. *Clin Ther*. 2004;26(7):1026-1036.
A7	Batista JE, Kölbl H, Herschorn S3, Rechberger T4, Cambronero J5, Halaska M6, Coppell A7, Kaper M8, Huang M7, Siddiqui E7; BEYOND study group. The efficacy and safety of mirabegron compared with solifenacin in overactive bladder patients dissatisfied with previous antimuscarinic treatment due to lack of efficacy: Results of a noninferiority, randomized, phase IIIb trial. *Therapeutic Advances in Urology*. 2015;7(4):167-179.
A8	Bent AE, Gousse AE, Hendrix SL, et al. Duloxetine compared with placebo for the treatment of women with mixed urinary incontinence. *Neurourology & Urodynamics*. 2008;27(3):212-221.
A9	Burgio KL, Kraus SR, Menefee S, et al. Behavioral therapy to enable women with urge incontinence to discontinue drug treatment: A randomized trial. *Ann Intern Med*. 2008;149(3):161-169.
A10	Burgio KL, Goode PS, Richter HE, Markland AD, Johnson II TM, Redden DT. Combined behavioral and individualized drug therapy versus individualized drug therapy alone for urge urinary incontinence in women. *J Urol*. 2010;184(2):598-603.
A11	Burgio KL, Goode PS, Johnson TM, et al. Behavioral versus drug treatment for overactive bladder in men: The male overactive bladder treatment in veterans (MOTIVE) trial. *J Am Geriatr Soc*. 2011;59(12):2209-2216.
A12	But I, Goldstajn MS, Oreskovic S. Comparison of two selective muscarinic receptor antagonists (solifenacin and darifenacin) in women with overactive bladder--the SOLIDAR study. *Coll Antropol*. 2012;36(4):1347-1353.
A13	Cardozo L, Lisec M, Millard R, et al. Randomized, double-blind placebo controlled trial of the once daily antimuscarinic agent solifenacin succinate in patients with overactive bladder. *J Urol*. 2004;172(5 I):1919-1924.
A14	Cardozo L, Dixon A. Increased warning time with darifenacin: A new concept in the management of urinary urgency. *J Urol*. 2005;173(4):1214-1218.
A15	Cardozo L, Hessdorfer E, Milani R, et al. Solifenacin in the treatment of urgency and other symptoms of overactive bladder: Results from a randomized, double-blind, placebo-controlled, rising-dose trial. *BJU Int*. 2008;102(9):1120-1127.
A16	Cartwright R, Srikrishna S, Cardozo L, Robinson D. Patient-selected goals in overactive bladder: A placebo controlled randomized double-blind trial of transdermal oxybutynin for the treatment of urgency and urge incontinence. *BJU Int*. 2011;107(1):70-76.
A17	Chancellor M, Freedman S, Mitcheson HD, Antoci J, Primus G, Wein A. Tolterodine, an effective and well tolerated treatment for urge incontinence and other overactive bladder symptoms. *Clinical Drug Investigation*. 2000;19(2):83-91.
A18	Chancellor MB, Kianifard F, Beamer E, et al. A comparison of the efficacy of darifenacin alone vs. darifenacin plus a behavioural modification programme upon the symptoms of overactive bladder. *Int J Clin Pract*. 2008;62(4):606-613.
A19	Chapple CR, Arano P, Bosch JL, De Ridder D, Kramer AE, Ridder AM. Solifenacin appears effective and well tolerated in patients with symptomatic idiopathic detrusor overactivity in a placebo- and tolterodine controlled phase 2 dose-finding study. *BJU Int*. 2004;93(1):71-77.
A20	Chapple CR, Rechberger T, Al-Shukri S, et al. Randomized, double-blind placebo-and tolterodine-controlled trial of the once-daily antimuscarinic agent solifenacin in patients with symptomatic overactive bladder. *BJU Int*. 2004;93(3):303-310.
A21	Chapple CR, Martinez-Garcia R, Selvaggi L, et al. A comparison of the efficacy and tolerability of solifenacin succinate and extended release tolterodine at treating overactive bladder syndrome: Results of the STAR trial. *Eur Urol*. 2005;48(3):464-470.
A22	Chapple CR, Patroneva A, Raines SR. Effect of an ATP-sensitive potassium channel opener in subjects with overactive bladder: A randomized, double-blind, placebo-controlled study (ZD0947IL/0004). *Eur Urol*. 2006;49(5):879-886.
A23	Chapple C, Van Kerrebroeck P, Tubaro A, et al. Clinical efficacy, safety, and tolerability of once-daily fesoterodine in subjects with overactive bladder. *Eur Urol*. 2007;52(4):1204-1212.
A24	Chapple C, Sievert K-, Macdiarmid S, et al. OnabotulinumtoxinA 100 U significantly improves all idiopathic overactive bladder symptoms and quality of life in patients with overactive bladder and urinary incontinence: A randomised, double-blind, placebo-controlled trial. *Eur Urol*. 2013;64(2):249-256.

TABLE B.1: Study references (cont.)

A25	Chapple CR, Amarenco G, Lopez Aramburu MA, et al. A proof-of-concept study: Mirabegron, a new therapy for overactive bladder. *Neurourol Urodyn*. 2013;32(8):1116-1122.
A26	Chapple CR, Dvorak V, Radziszewski P, et al. A phase II dose-ranging study of mirabegron in patients with overactive bladder. *Int Urogynecol J Pelvic Floor Dysfunct*. 2013;24(9):1447-1458.
A27	Chapple CR, Kaplan SA, Mitcheson D, et al. Randomized double-blind, active-controlled phase 3 study to assess 12-month safety and efficacy of mirabegron, a (3)-adrenoceptor agonist, in overactive bladder. *Eur Urol*. 2013;63(2):296-305.
A28	Chapple C, Schneider T, Haab F, et al. Superiority of fesoterodine 8 mg vs 4 mg in reducing urgency urinary incontinence episodes in patients with overactive bladder: Results of the randomised, double-blind, placebo-controlled EIGHT trial. *BJU Int*. 2014;114(3):418-426.
A29	Chapple CR, Abrams P, Andersson K-, et al. Phase II study on the efficacy and safety of the EP1 receptor antagonist ONO-8539 for nonneurogenic overactive bladder syndrome. *J Urol*. 2014;191(1):253-260.
A30	Choo MS, Lee JZ, Lee JB, et al. Efficacy and safety of solifenacin succinate in korean patients with overactive bladder: A randomised, prospective, double-blind, multicentre study. *Int J Clin Pract*. 2008;62(11):1675-1683.
A31	Chu F, Smith N, Uchida T. Efficacy and safety of solifenacin succinate 10 mg once daily: A multicenter, phase III, randomized, double-blind, placebo-controlled, parallel-group trial in patients with overactive bladder. *Current Therapeutic Research - Clinical and Experimental*. 2009;70(6):405-420.
A32	Chuang Y, Kaufmann JH, Chancellor DD, Chancellor MB, Kuo H. Bladder instillation of liposome encapsulated onabotulinumtoxina improves overactive bladder symptoms: A prospective, multicenter, double-blind, randomized trial. *J Urol*. 2014;192(6):1743-1749.
A33	Davila GW, Daugherty CA, Sanders SW. A short-term, multicenter, randomized double-blind dose titration study of the efficacy and anticholinergic side effects of transdermal compared to immediate release oral oxybutynin treatment of patients with urge urinary incontinence. *J Urol*. 2001;166(1):140-145.
A34	Digesu GA, Verdi E, Cardozo L, Olivieri L, Khullar V, Colli E. Phase IIb, multicenter, double-blind, randomized, placebo-controlled, parallel-group study to determine effects of elocalcitol in women with overactive bladder and idiopathic detrusor overactivity. *Urology*. 2012;80(1):48-54.
A35	Diokno AC, Appell RA, Sand PK, et al. Prospective, randomized, double-blind study of the efficacy and tolerability of the extended-release formulations of oxybutynin and tolterodine for overactive bladder: Results of the OPERA trial. *Mayo Clin Proc*. 2003;78(6):687-695.
A36	Dmochowski RR, Davila GW, Zinner NR, et al. Efficacy and safety of transdermal oxybutynin in patients with urge and mixed urinary incontinence. *J Urol*. 2002;168(2):580-586.
A37	Dmochowski RR, Sand PK, Zinner NR, Gittelman MC, Davila GW, Sanders SW. Comparative efficacy and safety of transdermal oxybutynin and oral tolterodine versus placebo in previously treated patients with urge and mixed urinary incontinence. *Urology*. 2003;62(2):237-242.
A38	Dmochowski RR, Sand PK, Zinner NR, Staskin DR. Trospium 60 mg once daily (QD) for overactive bladder syndrome: Results from a placebo-controlled interventional study. *Urology*. 2008;71(3):449-454.
A39	Dmochowski R, Chapple C, Nitti VW, et al. Efficacy and safety of onabotulinumtoxinA for idiopathic overactive bladder: A double-blind, placebo controlled, randomized, dose ranging trial. *J Urol*. 2010;184(6):2416-2422.
A40	Dmochowski RR, Peters KM, Morrow JD, et al. Randomized, double-blind, placebo-controlled trial of flexible-dose fesoterodine in subjects with overactive bladder. *Urology*. 2010;75(1):62-68.
A41	Dmochowski RR, Staskin DR, Duchin K, Paborji M, Tremblay TM. Clinical safety, tolerability and efficacy of combination tolterodine/pilocarpine in patients with overactive bladder. *Int J Clin Pract*. 2014;68(8):986-994.
A42	Drutz HP, Appell RA, Gleason D, Klimberg I, Radomski S. Clinical efficacy and safety of tolterodine compared to oxybutynin and placebo in patients with overactive bladder. *Int Urogynecol J Pelvic Floor Dysfunct*. 1999;10(5):283-289.
A43	Dubeau CE, Kraus SR, Griebling TL, et al. Effect of fesoterodine in vulnerable elderly subjects with urgency incontinence: A double-blind, placebo controlled trial. *J Urol*. 2014;191(2):395-404.
A44	Enzelsberger H, Kurz C, Schatten C, Huber J. Effects of intravaginal application of oestriol tablets in women with urge incontinence. *Geburtshilfe Frauenheilkd*. 1991;51(10):834-838.
A45	Enzelsberger H, Schatten C, Kurz C. [Comparison of emepronium bromide with intravesical administration of lidocaine gel in women with urge incontinence]. *Geburtshilfe Frauenheilkd*. 1991;51(1):54-57.
A46	Enzelberger H, Helmer H, Kurz C. Intravesical instillation of oxybutynin in women with idiopathic detrusor instability: A randomised trial. *Br J Obstet Gynaecol*. 1995;102(11):929-930.
A47	Enzelsberger H, Kurz C, Helmer H, Mittermayer F. [Topical administration of oxybutynin hydrochloride in women with urge incontinence. results of a prospective randomized double-blind study]. *Geburtshilfe Frauenheilkd*. 1995;55(5):240-243.
A48	Finazzi-Agro E, Petta F, Sciobica F, Pasqualetti P, Musco S, Bove P. Percutaneous tibial nerve stimulation effects on detrusor overactivity incontinence are not due to a placebo effect: A randomized, double-blind, placebo controlled trial. *J Urol*. 2010;184(5):2001-2006.
A49	FitzGerald MP, Lemack G, Wheeler T, Litman HJ. Nocturia, nocturnal incontinence prevalence, and response to anticholinergic and behavioral therapy. *Int Urogynecol J Pelvic Floor Dysfunct*. 2008;19(11):1545 1550.
A50	Franzen K, Johansson JE, Lauridsen I, Canelid J, Heiwall B, Nilsson K. Electrical stimulation compared with tolterodine for treatment of urge/urge incontinence amongst women--a randomized controlled trial. *International Urogynecology Journal*. 2010;21(12):1517-1524.
A51	Frenkl TL, Zhu H, Reiss T, Seltzer O, Rosenberg E, Green S. A multicenter, double-blind, randomized, placebo controlled trial of a neurokinin-1 receptor antagonist for overactive bladder. *J Urol*. 2010;184(2):616-622.

TABLE B.1: Study references (cont.)

A52 Fukuda T, Yamanishi T, Uchiyama T, Kamai T. Randomized, single-blind, parallel study of the effectiveness and safety of solifenacin versus propiverine in the treatment of overactive bladder. *LUTS: Lower Urinary Tract Symptoms*. 2013;5(1):11-16.

A53 Giannitsas K, Perimenis P, Athanasopoulos A, Gyftopoulos K, Nikiforidis G, Barbalias G. Comparison of the efficacy of tolterodine and oxybutynin in different urodynamic severity grades of idiopathic detrusor overactivity. *Eur Urol*. 2004;46(6):776-782.

A54 Gittelman M, Weiss H, Seidman L. A phase 2, randomized, double-blind, efficacy and safety study of oxybutynin vaginal ring for alleviation of overactive bladder symptoms in women. *J Urol*. 2014;191(4):1014-1021.

A55 Goldfischer ER, Sand PK, Thomas H, Peters-Gee J. Efficacy and safety of oxybutynin topical gel 3% in patients with urgency and/or mixed urinary incontinence: A randomized, double-blind, placebo-controlled study. *Neurourology & Urodynamics*. 2015;34(1):37-43.

A56 Gotoh M, Yokoyama O, Nishizawa O. Propiverine hydrochloride in japanese patients with overactive bladder: A randomized, double-blind, placebo-controlled trial. *International Journal of Urology*. 2011;18(5):365-373.

A57 Haab F, Braticevici B, Krivoborodov G, Palmas M, Zufferli Russo M, Pietra C. Efficacy and safety of repeated dosing of netupitant, a neurokinin-1 receptor antagonist, in treating overactive bladder. *Neurourol Urodyn*. 2014;33(3):335-340.

A58 Halaska M, Ralph G, Wiedemann A, et al. Controlled, double-blind, multicentre clinical trial to investigate long-term tolerability and efficacy of trospium chloride in patients with detrusor instability. *World J Urol*. 2003;20(6):392-399.

A59 Hassouna MM, Siegel SW, Lycklama A Nyeholt AAB, et al. Sacral neuromodulation in the treatment of urgency-frequency symptoms: A multicenter study on efficacy and safety. *J Urol*. 2000;163(6):1849-1854.

A60 Herschorn S, Becker D, Miller E, Thompson M, Forte L. Impact of a health education intervention in overactive bladder patients. *Can J Urol*. 2004;11(6):2430-2437.

A61 Herschorn S, Heesakkers J, Castro-Diaz D, Wang JT, Brodsky M, Guan Z. Effects of tolterodine extended release on patient perception of bladder condition and overactive bladder symptoms*. *Curr Med Res Opin*. 2008;24(12):3513-3521.

A62 Herschorn S, Stothers L, Carlson K, et al. Tolerability of 5 mg solifenacin once daily versus 5 mg oxybutynin immediate release 3 times daily: Results of the VECTOR trial. *J Urol*. 2010;183(5):1892-1898.

A63 Herschorn S, Swift S, Guan Z, et al. Comparison of fesoterodine and tolterodine extended release for the treatment of overactive bladder: A head-to-head placebo-controlled trial. *BJU Int*. 2010;105(1):58-66.

A64 Herschorn S, Barkin J, Castro-Diaz D, et al. A phase III, randomized, double-blind, parallel-group, placebo-controlled, multicentre study to assess the efficacy and safety of the beta3 adrenoceptor agonist, mirabegron, in patients with symptoms of overactive bladder. *Urology*. 2013;82(2):313-320.

A65 Hill S, Khullar V, Wyndaele J-, Lheritier K. Dose response with darifenacin, a novel once-daily M3 selective receptor antagonist for the treatment of overactive bladder: Results of a fixed dose study. *Int Urogynecol J Pelvic Floor Dysfunct*. 2006;17(3):239-247.

A66 Ho CH, Chang TC, Lin HH, Liu SP, Huang KH, Yu HJ. Solifenacin and tolterodine are equally effective in the treatment of overactive bladder symptoms. *J Formosan Med Assoc*. 2010;109(10):702-708.

A67 Holmes DM, Montz FJ, Stanton SL. Oxybutinin versus propantheline in the management of detrusor instability. A patient-regulated variable dose trial. *British Journal of Obstetrics & Gynaecology*. 1989;96(5):607-612.

A68 Homma Y, Paick JS, Lee JG, Kawabe K. Clinical efficacy and tolerability of extended-release tolterodine and immediate-release oxybutynin in japanese and korean patients with an overactive bladder: A randomized, placebo-controlled trial. *BJU Int*. 2003;92(7):741-747.

A69 Homma Y, Yamaguchi T, Yamaguchi O. A randomized, double-blind, placebo-controlled phase II dose-finding study of the novel anti-muscarinic agent imidafenacin in japanese patients with overactive bladder. *International Journal of Urology*. 2008;15(9):809-815.

A70 Homma Y, Yamaguchi O. A randomized, double-blind, placebo- and propiverine-controlled trial of the novel antimuscarinic agent imidafenacin in japanese patients with overactive bladder. *International Journal of Urology*. 2009;16(5):499-506.

A71 Hsiao SM, Chang TC, Wu WY, Chen CH, Yu HJ, Lin HH. Comparisons of urodynamic effects, therapeutic efficacy and safety of solifenacin versus tolterodine for female overactive bladder syndrome. *Journal of Obstetrics & Gynaecology Research*. 2011;37(8):1084-1091.

A72 Huang AJ, Hess R, Arya LA, et al. Pharmacologic treatment for urgency-predominant urinary incontinence in women diagnosed using a simplified algorithm: A randomized trial. *American Journal of Obstetrics & Gynecology*. 2012;206(5):444.e1-444.11.

A73 Huo L-, Jing H-, Wang T-, et al. A combination of solifenacin succinate and naftopidil in the treatment of female overactive bladder. *National Medical Journal of China*. 2013;93(44):3526-3528.

A74 Jabs C, Carleton E. Efficacy of botulinum toxin a intradetrusor injections for non-neurogenic urinary urge incontinence: A randomized double-blind controlled trial. *Journal of Obstetrics & Gynaecology Canada: JOGC*. 2013;35(1):53-60.

A75 Jacquetin B, Wyndaele J. Tolterodine reduces the number of urge incontinence episodes in patients with an overactive bladder. *European Journal of Obstetrics, Gynecology, & Reproductive Biology*. 2001;98(1):97-102.

A76 Johnson II TM, Burgio KL, Redden DT, Wright KC, Goode PS. Effects of behavioral and drug therapy on nocturia in older incontinent women. *J Am Geriatr Soc*. 2005;53(5):846-850.

TABLE B.1: Study references (cont.)

A77	Junemann KP, Halaska M, Rittstein T, et al. Propiverine versus tolterodine: Efficacy and tolerability in patients with overactive bladder. *Eur Urol*. 2005;48(3):478-482.
A78	Junemann KP, Hessdorfer E, Unamba-Oparah I, et al. Propiverine hydrochloride immediate and extended release: Comparison of efficacy and tolerability in patients with overactive bladder. *Urol Int*. 2006;77(4):334-339.
A79	Kafri R, Deutscher D, Shames J, Golombp J, Melzer I. Randomized trial of a comparison of rehabilitation or drug therapy for urgency urinary incontinence: 1-year follow-up. *International Urogynecology Journal*. 2013;24(7):1181-1189.
A80	Kaplan SA, Schneider T, Foote JE, Guan Z, Carlsson M, Gong J. Superior efficacy of fesoterodine over tolterodine extended release with rapid onset: A prospective, head-to-head, placebo-controlled trial. *BJU Int*. 2011;107(9):1432-1440.
A81	Kaplan SA, Cardozo L, Herschorn S, Grenabo L, Carlsson M, Arumi D, Crook TJ, Whelan L, Scholfield D, Ntanios F; Assessment of Fesoterodine after Tolterodine ER (AFTER) Study Group. Efficacy and safety of fesoterodine 8 mg in subjects with overactive bladder after a suboptimal response to tolterodine ER. *Int J Clin Pract*. 2014;68(9):1065-1073.
A82	Karram MM, Toglia MR, Serels SR, Andoh M, Fakhoury A, Forero-Schwanhaeuser S. Treatment with solifenacin increases warning time and improves symptoms of overactive bladder: Results from VENUS, a randomized, double-blind, placebo-controlled trial. *Urology*. 2009;73(1):14-18.
A83	Kaya S, Akbayrak T, Beksac S. Comparison of different treatment protocols in the treatment of idiopathic detrusor overactivity: A randomized controlled trial. *Clin Rehabil*. 2011;25(4):327-338.
A84	Khullar V, Hill S, Laval KU, Schiotz HA, Jonas U, Versi E. Treatment of urge-predominant mixed urinary incontinence with tolterodine extended release: A randomized, placebo-controlled trial. *Urology*. 2004;64(2):269-274.
A85	Khullar V, Amarenco G, Angulo JC, et al. Efficacy and tolerability of mirabegron, a (3)-adrenoceptor agonist, in patients with overactive bladder: Results from a randomised european-australian phase 3 trial. *Eur Urol*. 2013;63(2):283-295.
A86	Randomized controlled trial of cyclic and continuous therapy with trospium and solifenacin combination for severe overactive bladder in elderly patients with regard to patient compliance. *Therapeutic Advances in Urology*. 2014;6(6):215-223.
A87	Kuo HC. Bladder base/trigone injection is safe and as effective as bladder body injection of onabotulinumtoxinA for idiopathic detrusor overactivity refractory to antimuscarinics. *Neurourology & Urodynamics*. 2011;30(7):1242-1248.
A88	Kuo H-, Liu H-, Chuang Y-, Birder LA, Chancellor MB. Pilot study of liposome-encapsulated onabotulinumtoxinA for patients with overactive bladder: A single-center study. *Eur Urol*. 2014;65(6):1117-1124.
A89	Kuo HC, Lee KS, Na Y, Sood R, Nakaji S, Kubota Y, Kuroishi K. Results of a randomized, double-blind, parallel-group, placebo- and active-controlled, multicenter study of mirabegron, a beta3-adrenoceptor agonist, in patients with overactive bladder in asia. *Neurourol Urodyn*. 2015;34(7):685-692.
A90	Kuo HC, Lin HH, Yu HJ, Cheng CL, Hung MJ, Tong-Long A. Results of a randomized, double-blind, placebo-controlled study of mirabegron in a taiwanese population with overactive bladder and comparison with other clinical trials. *Urological Science*. 2015;26(1):41-48.
A91	Kurz C, Nagele F, Sevelda P, Enzelsberger H. [Intravesical administration of estriol in sensory urge incontinence--a prospective study]. *Geburtshilfe Frauenheilkd*. 1993;53(8):535-538.
A92	Lauti M, Herbison P, Hay-Smith J, Ellis G, Wilson D. Anticholinergic drugs, bladder retraining and their combination for urge urinary incontinence: A pilot randomised trial. *International Urogynecology Journal*. 2008;19(11):1533-1543.
A93	Lee JG, Hong JY, Choo MS, et al. Tolterodine: As effective but better tolerated than oxybutynin in asian patients with symptoms of overactive bladder. *International Journal of Urology*. 2002;9(5):247-252.
A94	Lee KS, Lee HW, Choo MS, et al. Urinary urgency outcomes after propiverine treatment for an overactive bladder: The 'propiverine study on overactive bladder including urgency data'. *BJU Int*. 2010;105(11):1565-1570.
A95	Lee KS, Park B, Kim JH, et al. A randomised, double-blind, parallel design, multi-institutional, non-inferiority phase IV trial of imidafenacin versus fesoterodine for overactive bladder. *Int J Clin Pract*. 2013;67(12):1317-1326.
A96	Lehtoranta K, Tainio H, Lukkari-Lax E, Hakonen T, Tammela TLJ. Pharmacokinetics, efficacy, and safety of intravesical formulation of oxybutynin in patients with detrusor overactivity. *Scand J Urol Nephrol*. 2002;36(1):18-24.
A97	Madersbacher H, Halaska M, Voigt R, Alloussi S, Hofner K. A placebo-controlled, multicentre study comparing the tolerability and efficacy of propiverine and oxybutynin in patients with urgency and urge incontinence. *BJU Int*. 1999;84(6):646-651.
A98	Mak HL, Cheon WC, Wong T, Liu YS, Tong WM. Randomized controlled trial of foot reflexology for patients with symptomatic idiopathic detrusor overactivity. *International Urogynecology Journal*. 2007;18(6):653-658.
A99	Malone-Lee JG, Bernard Walsh J, Maugourd M-. Tolterodine: A safe and effective treatment for older patients with overactive bladder. *J Am Geriatr Soc*. 2001;49(6):700-705.
A100	Malone-Lee J, Shaffu B, Anand C, Powell C. Tolterodine: Superior tolerability than and comparable efficacy to oxybutynin in individuals 50 years old or older with overactive bladder: A randomized controlled trial. *J Urol*. 2001;165(5 I):1452-1456.
A101	Malone-Lee JG, Al-Buheissi S. Does urodynamic verification of overactive bladder determine treatment success? results from a randomized placebo-controlled study. *BJU Int*. 2009;103(7):931-937.

TABLE B.1: Study references (cont.)

A102	Marencak J, Cossons NH, Darekar A, Mills IW. Investigation of the clinical efficacy and safety of pregabalin alone or combined with tolterodine in female subjects with idiopathic overactive bladder. *Neurourology & Urodynamics*. 2011;30(1):75-82.
A103	Martinez-Garcia R, Abadias M, Arano P, et al. Cizolirtine citrate, an effective treatment for symptomatic patients with urinary incontinence secondary to overactive bladder: A pilot dose-finding study. *Eur Urol*. 2009;56(1):184-192.
A104	Mattiasson A, Blaakaer J, Hoye K, Wein AJ, Tolterodine Scandinavian Study G. Simplified bladder training augments the effectiveness of tolterodine in patients with an overactive bladder. *BJU Int*. 2003;91(1):54-60.
A105	Mattiasson A, Masala A, Morton R, Bolodeoku J. Efficacy of simplified bladder training in patients with overactive bladder receiving a solifenacin flexible-dose regimen: Results from a randomized study. *BJU Int*. 2010;105(8):1126-1135.
A106	Mazur D, Wehnert J, Dorschner W, Schubert G, Herfurth G, Alken RG. Clinical and urodynamic effects of propiverine in patients suffering from urgency and urge incontinence. A multicentre dose-optimizing study. *Scandinavian Journal of Urology & Nephrology*. 1995;29(3):289-294.
A107	Meyhoff HH, Gerstenberg TC, Nordling J. Placebo--the drug of choice in female motor urge incontinence?. *Br J Urol*. 1983;55(1):34-37.
A108	Millard R, Tuttle J, Moore K, et al. Clinical efficacy and safety of tolterodine compared to placebo in detrusor overactivity. *J Urol*. 1999;161(5):1551-1555.
A109	Millard RJ, Asia Pacific Tolterodine Study G. Clinical efficacy of tolterodine with or without a simplified pelvic floor exercise regimen. *Neurourology & Urodynamics*. 2004;23(1):48-53.
A110	Nitti VW, Dmochowski R, Sand PK, et al. Efficacy, safety and tolerability of fesoterodine for overactive bladder syndrome. *J Urol*. 2007;178(6):2488-2494.
A111	Nitti VW, Rovner ES, Bavendam T. Response to fesoterodine in patients with an overactive bladder and urgency urinary incontinence is independent of the urodynamic finding of detrusor overactivity. *BJU Int*. 2010;105(9):1268-1275.
A112	Nitti VW, Auerbach S, Martin N, Calhoun A, Lee M, Herschorn S. Results of a randomized phase III trial of mirabegron in patients with overactive bladder. *J Urol*. 2013;189(4):1388-1395.
A113	Nitti VW, Dmochowski R, Herschorn S, et al. OnabotulinumtoxinA for the treatment of patients with overactive bladder and urinary incontinence: Results of a phase 3, randomized, placebo controlled trial. *J Urol*. 2013;189(6):2186-2193.
A114	Norton P, Karram M, Wall LL, Rosenzweig B, Benson JT, Fantl JA. Randomized double-blind trial of terodiline in the treatment of urge incontinence in women. *Obstetrics & Gynecology*. 1994;84(3):386-391.
A115	O'Reilly BA, Fynes M, Achtari C, et al. A prospective randomised double-blind controlled trial evaluating the effect of trans-sacral magnetic stimulation in women with overactive bladder. *International Urogynecology Journal*. 2008;19(4):497-502.
A116	Ohlstein EH, von Keitz A, Michel MC. A multicenter, double-blind, randomized, placebo-controlled trial of the 3-adrenoceptor agonist solabegron for overactive bladder. *Eur Urol*. 2012;62(5):834-840.
A117	Olmo Carmona MV, Molleja AMG, Rios IL, Torronteras AR, Tamajon VMC, Obreroc IG. Percutaneous tibial nerve stimulation versus neurostimulation of SP 6 (sanyinjiao) in urge incontinence. *Revista Internacional de Acupuntura*. 2013;7(4):124-130.
A118	Oreskovic S, But I, Banovic M, Goldstajn MS. The efficacy and safety of solifenacin in patients with overactive bladder syndrome. *Coll Antropol*. 2012;36(1):243-248.
A119	Orri M, Lipset CH, Jacobs BP, Costello AJ, Cummings SR. Web-based trial to evaluate the efficacy and safety of tolterodine ER 4mg in participants with overactive bladder: REMOTE trial. *Contemporary Clinical Trials*. 2014;38(2):190-197.
A120	Terodiline in the Elderly American Multicenter Study Group. Effects of terodiline on urinary incontinence among older non-institutionalized women. terodiline in the elderly american multicenter study group. *J Am Geriatr Soc*. 1993;41(9):915-922.
A121	Ozdedeli S, Karapolat H, Akkoc Y. Comparison of intravaginal electrical stimulation and trospium hydrochloride in women with overactive bladder syndrome: A randomized controlled study. *Clin Rehabil*. 2010;24(4):342-351.
A122	Park C, Park J, Choo M-, et al. A randomised, prospective double-blind, propiverine-controlled trial of imidafenacin in patients with overactive bladder. *Int J Clin Pract*. 2014;68(2):188-196.
A123	Peters KM, Macdiarmid SA, Wooldridge LS, et al. Randomized trial of percutaneous tibial nerve stimulation versus extended-release tolterodine: Results from the overactive bladder innovative therapy trial. *J Urol*. 2009;182(3):1055-1061.
A124	Peters KM, Carrico DJ, Perez-Marrero RA, et al. Randomized trial of percutaneous tibial nerve stimulation versus sham efficacy in the treatment of overactive bladder syndrome: Results from the SUmiT trial. *J Urol*. 2010;183(4):1438-1443.
A125	Preik M, Albrecht D, O'Connell M, Hampel C, Anderson R. Effect of controlled-release delivery on the pharmacokinetics of oxybutynin at different dosages: Severity-dependent treatment of the overactive bladder. *BJU Int*. 2004;94(6):821-827.
A126	Preyer O, Umek W, Laml T, Bjelic-Radisic V, Gabriel B, Mittlboeck M, Hanzal E. Percutaneous tibial nerve stimulation versus tolterodine for overactive bladder in women: A randomised controlled trial. *European Journal of Obstetrics Gynecology and Reproductive Biology*. 2015;191:51-56.
A127	Rentzhog L, Stanton SL, Cardozo L, Nelson E, Fall M, Abrams P. Efficacy and safety of tolterodine in patients with detrusor instability: A dose-ranging study. *Br J Urol*. 1998;81(1):42-48.
A128	Rios LAS, Panhoca R, Mattos J, D, Srugi M, Bruschini H. Intravesical resiniferatoxin for the treatment of women with idiopathic detrusor overactivity and urgency incontinence: A single dose, 4 weeks, double-blind, randomized, placebo controlled trial. *Neurourol Urodyn*. 2007;26(6):773-778.

Table B.1: Study references (cont.)

A129	Rogers R, Bachmann G, Jumadilova Z, et al. Efficacy of tolterodine on overactive bladder symptoms and sexual and emotional quality of life in sexually active women. *International Urogynecology Journal*. 2008;19(11):1551-1557.
A130	Rudy D, Cline K, Harris R, Goldberg K, Dmochowski R. Multicenter phase III trial studying trospium chloride in patients with overactive bladder. *Urology*. 2006;67(2):275-280.
A131	Rufford J, Hextall A, Cardozo L, Khullar V. A double-blind placebo-controlled trial on the effects of 25 mg estradiol implants on the urge syndrome in postmenopausal women. *International Urogynecology Journal*. 2003;14(2):78-83.
A132	Sahai A, Khan MS, Dasgupta P. Efficacy of botulinum toxin-A for treating idiopathic detrusor overactivity: Results from a single center, randomized, double-blind, placebo controlled trial. *J Urol*. 2007;177(6):2231-2236.
A133	Sancaktar M, Ceyhan ST, Akyol I, et al. The outcome of adding peripheral neuromodulation (stoller afferent neuro-stimulation) to anti-muscarinic therapy in women with severe overactive bladder. *Gynecological Endocrinology*. 2010;26(10):729-732.
A134	Schmidt RA, Jonas U, Oleson KA, et al. Sacral nerve stimulation for treatment of refractory urinary urge incontinence. sacral nerve stimulation study group. *J Urol*. 1999;162(2):352-357.
A135	Schreiner L, dos Santos TG, Knorst MR, da Silva Filho IG. Randomized trial of transcutaneous tibial nerve stimulation to treat urge urinary incontinence in older women. *International Urogynecology Journal*. 2010;21(9):1065-1070.
A136	Song C, Park JT, Heo KO, Lee KS, Choo MS. Effects of bladder training and/or tolterodine in female patients with overactive bladder syndrome: A prospective, randomized study. *J Korean Med Sci*. 2006;21(6):1060-1063.
A137	Song M, Kim JH, Lee KS, Lee JZ, Oh SJ, Seo JT, Choi JB, Kim SW, Rhee SJ, Choo MS. The efficacy and tolerability of tarafenacin, a new muscarinic acetylcholine receptor M3 antagonist in patients with overactive bladder; randomised, double-blind, placebo-controlled phase 2 study. *Int J Clin Pract*. 2015;69(2):242-250.
A138	Soomro NA, Khadra MH, Robson W, Neal DE. A crossover randomized trial of transcutaneous electrical nerve stimulation and oxybutynin in patients with detrusor instability. *J Urol*. 2001;166(1):146-149.
A139	Staskin D, Sand P, Zinner N, Dmochowski R, Trospium Study G. Once daily trospium chloride is effective and well tolerated for the treatment of overactive bladder: Results from a multicenter phase III trial. *J Urol*. 2007;178(3 Pt 1):978-983.
A140	Staskin DR, Dmochowski RR, Sand PK, et al. Efficacy and safety of oxybutynin chloride topical gel for overactive bladder: A randomized, double-blind, placebo controlled, multicenter study. *J Urol*. 2009;181(4):1764-1772.
A141	Steers W, Corcos J, Foote J, Kralidis G. An investigation of dose titration with darifenacin, an M 3-selective receptor antagonist. *BJU Int*. 2005;95(4):580-586.
A142	Steers WD, Herschorn S, Kreder KJ, et al. Duloxetine compared with placebo for treating women with symptoms of overactive bladder. *BJU Int*. 2007;100(2):337-345.
A143	Subak LL, Quesenberry CP, Posner SF, Cattolica E, Soghikian K. The effect of behavioral therapy on urinary incontinence: A randomized controlled trial. *Obstetrics & Gynecology*. 2002;100(1):72-78.
A144	Swift S, Garely A, Dimpfl T, Payne C. A new once-daily formulation of tolterodine provides superior efficacy and is well tolerated in women with overactive bladder. *Int Urogynecol J Pelvic Floor Dysfunct*. 2003;14(1):50-55.
A145	Tang H, Chen J, Wang Y, Yu T, Guo C, Liao X. Combination of sacral neuromodulation and tolterodine for treatment of idiopathic overactive bladder in women: A clinical trial. *Urology Journal*. 2014;11(4):1800-1805.
A146	Tapp A, Fall M, Norgaard J, et al. Terodiline: A dose titrated, multicenter study of the treatment of idiopathic detrusor instability in women. *J Urol*. 1989;142(4):1027-1031.
A147	Tincello DG, Adams EJ, Sutherst JR, Richmond DH. Oxybutynin for detrusor instability with adjuvant salivary stimulant pastilles to improve compliance: Results of a multicentre, randomized controlled trial. *BJU Int*. 2000;85(4):416-420.
A148	Tincello DG, Kenyon S, Abrams KR, et al. Botulinum toxin A versus placebo for refractory detrusor overactivity in women: A randomised blinded placebo-controlled trial of 240 women (the RELAX study). *Eur Urol*. 2012;62(3):507-514.
A149	Tseng LH, Wang AC, Chang YL, Soong YK, Lloyd LK, Ko YJ. Randomized comparison of tolterodine with vaginal estrogen cream versus tolterodine alone for the treatment of postmenopausal women with overactive bladder syndrome. *Neurourology & Urodynamics*. 2009;28(1):47-51.
A150	Ulshofer B, Bihr A-, Bodeker R-, Schwantes U, Jahn H-. Randomised, double-blind, placebo-controlled study on the efficacy and tolerance of trospium chloride in patients with motor urge incontinence. *Clinical Drug Investigation*. 2001;21(8):563-569.
A151	Van Kerrebroeck P, Kreder K, Jonas U, Zinner N, Wein A, Tolterodine Study G. Tolterodine once-daily: Superior efficacy and tolerability in the treatment of the overactive bladder. *Urology*. 2001;57(3):414-421.
A152	Van Kerrebroeck PEV, Kelleher CJ, Coyne KS, Kopp Z, Brodsky M, Wang JT. Correlations among improvements in urgency urinary incontinence, health-related quality of life, and perception of bladder-related problems in incontinent subjects with overactive bladder treated with tolterodine or placebo. *Health and Quality of Life Outcomes*. 2009;18:7-13.
A153	Vardy MD, Mitcheson HD, Samuels T-, et al. Effects of solifenacin on overactive bladder symptoms, symptom bother and other patient-reported outcomes: Results from VIBRANT - A double-blind, placebo-controlled trial. *Int J Clin Pract*. 2009;63(12):1702-1714.

Table B.1: Study references (cont.)

A154	Versi E, Appell R, Mobley D, Patton W, Saltzstein D. Dry mouth with conventional and controlled-release oxybutynin in urinary incontinence. the ditropan XL study group. *Obstetrics & Gynecology*. 2000;95(5):718-721.
A155	Visco AG, Brubaker L, Richter HE, et al. Anticholinergic therapy vs. onabotulinumtoxina for urgency urinary incontinence. *N Engl J Med*. 2012;367(19):1803-1813.
A156	Wagg A, Khullar V, Marschall-Kehrel D, et al. Flexible-dose fesoterodine in elderly adults with overactive bladder: Results of the randomized, double-blind, placebo-controlled study of fesoterodine in an aging population trial. *J Am Geriatr Soc*. 2013;61(2):185-193.
A157	Wang AC, Chih SY, Chen MC. Comparison of electric stimulation and oxybutynin chloride in management of overactive bladder with special reference to urinary urgency: A randomized placebo-controlled trial. *Urology*. 2006;68(5):999-1004.
A158	Wang AC, Chen MC, Kuo WY, Lin YH, Wang YC, Lo TS. Urgency-free time interval as primary endpoint for evaluating the outcome of a randomized OAB treatment. *International Urogynecology Journal*. 2009;20(7):819-825.
A159	Weiss JP, Jumadilova Z, Johnson TM,2nd, et al. Efficacy and safety of flexible dose fesoterodine in men and women with overactive bladder symptoms including nocturnal urinary urgency. *J Urol*. 2013;189(4):1396-1401.
A160	Yamaguchi O, Marui E, Kakizaki H, et al. Randomized, double-blind, placebo- and propiverine-controlled trial of the once-daily antimuscarinic agent solifenacin in japanese patients with overactive bladder. *BJU Int*. 2007;100(3):579-587.
A161	Yamaguchi O, Nishizawa O, Takeda M, et al. Efficacy, safety and tolerability of fesoterodine in asian patients with overactive bladder. *LUTS: Lower Urinary Tract Symptoms*. 2011;3(1):43-50.
A162	Yamaguchi O, Marui E, Kakizaki H, et al. Phase III, randomised, double-blind, placebo-controlled study of the beta3-adrenoceptor agonist mirabegron, 50=mg once daily, in japanese patients with overactive bladder. *BJU Int*. 2014;113(6):951-960.
A163	Yamaguchi O, Uchida E, Higo N, et al. Efficacy and safety of once-daily oxybutynin patch versus placebo and propiverine in japanese patients with overactive bladder: A randomized double-blind trial. *International Journal of Urology*. 2014;21(6):586-593.
A164	Yamaguchi O, Marui E, Igawa Y, Takeda M, Nishizawa O, Ikeda Y, Ohkawa S. Efficacy and safety of the selective beta3-adrenoceptor agonist mirabegron in japanese patients with overactive bladder: A randomized, double-blind, placebo-controlled, dose-finding study. *LUTS: Lower Urinary Tract Symptoms*. 2015;7(2):84-92.
A165	Yamanishi T, Yasuda K, Sakakibara R, Hattori T, Suda S. Randomized, double-blind study of electrical stimulation for urinary incontinence due to detrusor overactivity. *Urology*. 2000;55(3):353-357.
A166	Yokoyama T, Koide T, Hara R, Fukumoto K, Miyaji Y, Nagai A. Long-term safety and efficacy of two different antimuscarinics, imidafenacin and solifenacin, for treatment of overactive bladder: A prospective randomized controlled study. *Urol Int*. 2013;90(2):161-167.
A167	Yokoyama O, Hiro S, Hotta S, Mogami S, Yamagami H. Efficacy of fesoterodine on nocturia and quality of sleep in asian patients with overactive bladder. *Urology*. 2014;83(4):750-755.
A168	Yoon HS, Song HH, Ro YJ. A comparison of effectiveness of bladder training and pelvic muscle exercise on female urinary incontinence. *Int J Nurs Stud*. 2003;40(1):45-50.
A169	Zat'ura F, Vsetica J, Abadias M, et al. Cizolirtine citrate is safe and effective for treating urinary incontinence secondary to overactive bladder: A phase 2 proof-of-concept study. *Eur Urol*. 2010;57(1):145-153.
A170	Zellner M, Madersbacher H, Palmtag H, Stohrer M, Bodeker R-. Trospium chloride and oxybutynin hydrochloride in a german study of adults with urinary urge incontinence: Results of a 12-week, multicenter, randomized, double-blind, parallel-group, flexible-dose noninferiority trial. *Clin Ther*. 2009;31(11):2519-2539.
A171	Zimmern P, Litman HJ, Mueller E, Norton P, Goode P, Urinary Incontinence Treatment N. Effect of fluid management on fluid intake and urge incontinence in a trial for overactive bladder in women. *BJU Int*. 2010;105(12):1680-1685.
A172	Zinner NR, Mattiasson A, Stanton SL. Efficacy, safety, and tolerability of extended-release once-daily tolterodine treatment for overactive bladder in older versus younger patients. *J Am Geriatr Soc*. 2002;50(5):799-807.
A173	Zinner N, Tuttle J, Marks L. Efficacy and tolerability of darifenacin, a muscarinic M3 selective receptor antagonist (M3 SRA), compared with oxybutynin in the treatment of patients with overactive bladder. *World J Urol*. 2005;23(4):248-252.
A174	Zinner N, Susset J, Gittelman M, Arguinzoniz M, Rekeda L, Haab F. Efficacy, tolerability and safety of darifenacin, an M(3) selective receptor antagonist: An investigation of warning time in patients with OAB. *Int J Clin Pract*. 2006;60(1):119-126.
A175	Amaro JL, Gameiro MO, Padovani CR. Effect of intravaginal electrical stimulation on pelvic floor muscle strength. *International Urogynecology Journal*. 2005;16(5):355-358.
A176	Bellette PO, Rodrigues-Palma PC, Hermann V, Riccetto C, Bigozzi M, Olivares JM. [Posterior tibial nerve stimulation in the management of overactive bladder: A prospective and controlled study]. *Actas Urol Esp*. 2009;33(1):58-63.
A177	Burgio KL, Goode PS, Locher JL, et al. Behavioral training with and without biofeedback in the treatment of urge incontinence in older women: A randomized controlled trial. *JAMA*. 2002;288(18):2293-2299.
A178	Dowson C, Sahai A, Watkins J, Dasgupta P, Khan MS. The safety and efficacy of botulinum toxin-A in the management of bladder oversensitivity: A randomised double-blind placebo-controlled trial. *Int J Clin Pract*. 2011;65(6):698-704.

Table B.1: Study references (cont.)

A179	Homma Y, Kawabe K. Health-related quality of life of japanese patients with overactive bladder treated with extended-release tolterodine or immediate-release oxybutynin: A randomized, placebo-controlled trial. *World J Urol*. 2004;22(4):251-256.
A180	Karademir K, Baykal K, Sen B, Senkul T, Iseri C, Erden D. A peripheric neuromodulation technique for curing detrusor overactivity: Stoller afferent neurostimulation. *Scandinavian Journal of Urology & Nephrology*. 2005;39(3):230-233.
A181	Kelleher CJ, Reese PR, Pleil AM, Okano GJ. Health-related quality of life of patients receiving extended-release tolterodine for overactive bladder. *Am J Manag Care*. 2002;8(19):S608-15.
A182	Nascimento-Correia G, Santos-Pereira V, Tahara N, Driusso P. [Effects of pelvic floor muscle training on quality of life of a group of women with urinary incontinence: Randomized controlled trial]. *Actas Urol Esp*. 2012;36(4):216-221.
A183	Newman DK, Hanno PM, Dmochowski RR, et al. Effects of oxybutynin chloride topical gel on health-related quality of life in adults with overactive bladder: A randomized, double-blind, placebo-controlled study. *Clinical Medicine Insights: Therapeutics*. 2010;2:889-896.
A184	Peters KM, Feber KM, Bennett RC. Sacral versus pudendal nerve stimulation for voiding dysfunction: A prospective, single-blinded, randomized, crossover trial. *Neurourol Urodyn*. 2005;24(7):643-647.
A185	Pleil AM, Reese PR, Kelleher CJ, Okano GJ. Health-related quality of life of patients with overactive bladder receiving immediate-release tolterodine. *HEPAC Health Economics in Prevention and Care*. 2001;2(2):69-75.
A186	Robinson D, Cardozo L, Terpstra G, Bolodeoku J. A randomized double-blind placebo-controlled multicentre study to explore the efficacy and safety of tamsulosin and tolterodine in women with overactive bladder syndrome. *BJU Int*. 2007;100(4):840-845.
A187	Rogers RG, Bachmann G, Scarpero H, et al. Effects of tolterodine ER on patient-reported outcomes in sexually active women with overactive bladder and urgency urinary incontinence. *Current Medical Research & Opinion*. 2009;25(9):2159-2165.
A188	Sahai A, Dowson C, Khan MS, Dasgupta P. Improvement in quality of life after botulinum toxin - A injections for idiopathic detrusor overactivity: Results from a randomized double-blind placebo-controlled trial. *BJU Int*. 2009;103(11):1509-1515.
A189	Souto SC, Reis LO, Palma T, Palma P, Denardi F. Prospective and randomized comparison of electrical stimulation of the posterior tibial nerve versus oxybutynin versus their combination for treatment of women with overactive bladder syndrome. *World J Urol*. 2014;32(1):179-184.
A190	Wang AC, Wang YY, Chen MC. Single-blind, randomized trial of pelvic floor muscle training, biofeedback-assisted pelvic floor muscle training, and electrical stimulation in the management of overactive bladder. *Urology*. 2004;63(1):61-66.
A191	Zinner N, Gittelman M, Harris R, Susset J, Kanellos A, Auerbach S. Trospium chloride improves overactive bladder symptoms: A multicenter phase III trial. *J Urol*. 2004;171(6 I):2311-2315.
A192	Abrams P, Kelleher C, Staskin D, et al. Combination treatment with mirabegron and solifenacin in patients with overactive bladder: Efficacy and safety results from a randomised, double-blind, dose-ranging, phase 2 study (symphony). *Eur Urol*. 2015;67(3):577-588.
A193	Kosilov KV, Loparev SA, Ivanovskaya MA, Kosilova LV. Comparative effectiveness of combined low- and standard-dose trospium and solifenacin for moderate overactive bladder symptoms in elderly men and women. *Urol Int*. 2014;93(4):470-473.
A194	Kosilov K, Loparev S, Ivanovskaya M, Kosilova L. A randomized, controlled trial of effectiveness and safety of management of OAB symptoms in elderly men and women with standard-dosed combination of solifenacin and mirabegron. *Arch Gerontol Geriatr*. 2015;61(2):212-216.

B.3 Efficacy data

Table B.2: Efficacy data

Ref	First Author	Year	Treatment		Incontinence N	Incontinence Mean change	Incontinence SD	Voiding N	Voiding Mean change	Voiding SD	Urgency N	Urgency Mean change	Urgency SD	Nocturia N	Nocturia Mean change	Nocturia SD
A1	Abdelbary	2015	Electrostimulation	[80]	100	-0.40		100	-1.80		100	-2.60		100	-1.30	
			Vaginal oestrogen cream 1.25mg/day	[132]	98	-0.10		98	-1.30		98	-1.10		98	-0.60	
			Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	102	-0.41		102	-1.80		102	-4.70		102	-1.30	
A2	Abrams	1998	Placebo	[1]	40	-0.90	1.50	56	-1.60	3.60						
			Tolterodine IR 2mg b.i.d	[5]	92	-1.30	3.20	118	-2.70	3.80						
			Oxybutynin IR 5mg t.i.d	[7]	88	-1.70	3.10	117	-2.30	2.70						
A3	Anderson	1993	Terodiline 25mg b.i.d	[28]	40	-1.16	1.94	40	-0.76	1.73				40	-0.04	0.84
			Placebo	[1]	41	-0.60	1.96	41	-0.18	1.47				41	-0.26	0.56
A4	Anderson	1999	Oxybutynin ER 5mg-30mg	[22]	53	-3.33										
			Oxybutynin IR 5mg-20mg	[23]	52	-3.21										
A5	Appell	2001	Oxybutynin ER 10mg q.d	[8]	185	-3.07		185	-3.53							
			Tolterodine IR 2mg b.i.d	[5]	193	-2.53		193	-2.87							
A6	Barkin	2004	Oxybutynin ER 15mg q.d	[9]	53	-1.99		53	-1.80							
			Oxybutynin IR 5mg t.i.d	[7]	41	-2.41		41	-2.40							
A8	Bent	2008	Placebo	[1]	75	-0.52	1.20									
			Duloxetine 40mg b.i.d	[65]	81	-0.78	1.20									
A9	Burgio	2008	Tolterodine ER 4mg q.d + BT	[87]	154	-2.91	2.53	154	-0.50							
			Tolterodine ER 4mg q.d	[4]	153	-2.64	2.53	153	0.40							
A10	Burgio	2010	Oxybutynin ER 5 - 30mg/day	[22]	32	-2.29		32	-3.20	3.80						
			Oxybutynin ER 5-30mg q.d + BT	[86]	32	-1.46		32	-3.20	2.70						
A11	Burgio	2011	Bladder Training (BT)/Behaviour Therapy	[85]	22	-0.22		64	-2.20					64	-0.70	0.72
			Oxybutynin ER 5 - 30mg/day	[22]	24	-0.15		60	-2.00					60	-0.32	1.28
A12	But	2012	Solifenacin ER 5mg q.d	[29]				32	-2.17	2.50	32	-1.00	1.95	32	-1.20	1.10
			Darifenacin ER 7.5mg q.d	[39]				29	-2.10	3.40	29	-1.60	1.80	29	-0.80	1.30
A13	Cardozo	2004	Placebo	[1]	281	-1.25		281	-1.59		281	-1.98		281	-0.52	
			Solifenacin ER 5mg q.d	[29]	286	-1.63		286	-2.37	4.23	286	-2.84	5.00	286	-0.58	
			Solifenacin ER 10mg q.d	[30]	290	-1.57		290	-2.81	4.30	290	-2.90	4.95	290	-0.71	1.61
A14	Cardozo	2005	Darifenacin ER 30mg q.d	[38]							36	-0.80				
			Placebo	[1]							36	-0.30				
A15	Cardozo	2008	Solifenacin ER 5 - 10mg q.d	[31]	505	-1.70	2.20	505	-2.10	2.60	505	-1.70	2.20			
			Placebo	[1]	223	-1.40	2.00	223	-1.30	2.70	223	-1.30	2.00			
A16	Cartwright	2011	Oxybutynin trandermal 3.9mg/day	[10]	48	-0.47	0.81	48	-0.69	1.49	48	-1.23	1.40	48	-0.09	0.58
			Placebo	[1]	48	-0.23	0.61	48	0.05	1.49	48	-0.21	1.58	48	0.02	0.82
A17	Chancellor	2000	Placebo	[1]	507	-0.99	2.20	507	-1.20	2.90						
			Tolterodine IR 2mg b.i.d	[5]	514	-1.51	2.41	514	-1.70	3.30						
A18	Chancellor	2008	Darifenacin ER 7.5 - 15mg q.d	[104]	190	-1.89	2.29	190	-2.96	2.91	190	-2.87	3.59	190	-0.65	1.26
			Darifenacin 7.5 - 15mg q.d + BT	[88]	205	-2.10	2.32	205	-2.82	2.87	205	-2.68	3.54	205	-0.67	1.21
A19	Chapple	2004	Tolterodine IR 2mg b.i.d	[5]	37	-0.41		37	-1.79		37	-1.62				
			Placebo	[1]	36	-0.29		36	-1.03		36	-1.03				
			Solifenacin ER 2.5mg q.d	[32]	40	-0.66		40	-1.45		40	-1.07				
			Solifenacin ER 5mg q.d	[29]	37	-0.83		37	-2.21		37	-2.35				
			Solifenacin ER 10mg q.d	[30]	33	-0.79		33	-2.47		33	-2.46				
			Solifenacin ER 20mg q.d	[33]	34	-0.58		34	-2.75		34	-2.24				
A20	Chapple	2004	Placebo	[1]	253	-0.76	2.26	253	-1.20	3.26	253	-1.41	3.67			
			Solifenacin ER 5mg q.d	[29]	266	-1.42	1.82	266	-2.19	2.87	266	-2.85	3.74			
			Solifenacin ER 10mg q.d	[30]	264	-1.45	2.24	264	-2.61	3.24	264	-3.07	3.90			
			Tolterodine IR 2mg b.i.d	[5]	250	-1.14	2.15	250	-1.88	3.00	250	-2.05	3.58			
A21	Chapple	2005	Solifenacin ER 5 - 10mg q.d	[31]	578	-1.60		578	-2.45		578	-2.85				
			Tolterodine ER 4mg q.d	[4]	599	-1.11		599	-2.24		599	-2.42				
A22	Chapple	2006	ZD0947IL 25mg/day	[58]	92	-0.40		92	-0.90		92	-1.00				
			Placebo	[1]	99	-0.30		99	-1.60		99	-0.90				
A23	Chapple	2007	Placebo	[1]	211	-1.14	2.32	279	-0.95	2.67	279	-1.07	3.17	279	-0.32	1.00
			Tolterodine ER 4mg q.d	[4]	223	-1.74	2.39	283	-1.73	2.69	283	-2.03	3.20	283	-0.40	1.01
			Fesoterodine ER 4mg q.d	[25]	199	-1.95	2.40	265	-1.76	2.77	265	-1.88	3.26	265	-0.39	0.98
			Fesoterodine ER 8mg q.d	[26]	223	-2.22	2.39	276	-1.88	2.66	276	-2.36	3.32	276	-0.39	1.00
A24	Chapple	2013	Placebo	[1]	271	-1.03	3.02	271	-0.83	2.56	271	-1.24	3.86	271	-0.25	1.09
			OnaBoNT-A 100u trigone sparing	[72]	277	-2.95	3.57	277	-2.56	3.44	277	-3.67	4.42	277	-0.54	1.36
A25	Chapple	2013	Placebo	[1]	41	-1.01		64	-1.18		64	-1.03		57	-0.22	
			Mirabegron IR 100mg b.i.d	[48]	37	-2.17		65	-2.19		65	-2.30		58	-0.61	
			Mirabegron IR 150mg b.i.d	[49]	41	-1.58		63	-2.21		63	-2.30		54	-0.39	
			Tolterodine ER 4mg q.d	[4]	41	-1.65		63	-1.49		63	-2.09		58	-0.41	
A26	Chapple	2013	Placebo	[1]	166	-0.53		166	-1.44		166	-1.07		166	-0.38	
			Mirabegron ER 25mg q.d	[50]	167	-1.36	2.84	167	-1.88		167	-1.77		167	-0.52	
			Mirabegron ER 50mg q.d	[51]	167	-1.15	2.79	167	-2.08		167	-1.67		167	-0.60	
			Mirabegron ER 100mg q.d	[52]	168	-1.06	2.75	168	-2.12		168	-2.28		168	-0.42	
			Mirabegron ER 200mg q.d	[53]	166	-1.10	2.72	166	-2.24		166	-2.48		166	-0.59	
			Tolterodine ER 4mg q.d	[4]	85	-0.81	2.95	85	-1.99		85	-1.46		85	-0.59	
A27	Chapple	2013	Mirabegron ER 50mg q.d	[51]	479	-1.09		789	-1.12					789	-0.46	1.12
			Mirabegron ER 100mg q.d	[52]	483	-1.26		802	-1.44					802	-0.39	1.13
			Tolterodine ER 4mg q.d	[4]	488	-1.06		791	-1.26					791	-0.43	1.12
A28	Chapple	2014	Placebo	[1]	386	-2.20		386	-1.60		386	-3.00				
			Fesoterodine ER 4mg q.d	[25]	790	-2.90		790	-2.50		790	-4.20				
			Fesoterodine ER 8mg q.d	[26]	779	-3.10		779	-3.00		779	-5.00				
A29	Chapple	2014	Placebo	[1]	80	-1.86	2.60	80	-1.40	2.74	80	-2.18	3.72			
			ONO-8539 30mg b.i.d	[59]	87	-1.22	2.69	87	-1.02	2.74	87	-1.46	3.69			
			ONO-8539 100mg b.i.d	[60]	83	-1.91	2.58	83	-1.53	2.70	83	-2.27	3.65			
			ONO-8539 300mg b.i.d	[61]	82	-1.28	2.54	82	-1.31	2.73	82	-2.52	3.67			
			Tolterodine ER 4mg q.d	[4]	83	-2.24	2.72	83	-2.18	2.72	83	-3.27	3.67			
A30	Choo	2008	Solifenacin ER 5mg q.d	[29]	107	-1.14		107	-2.18		107	-2.50		107	-0.67	
			Solifenacin ER 10mg q.d	[30]	111	-1.84		111	-2.47		111	-2.35		111	-0.60	
			Tolterodine IR 2mg b.i.d	[5]	111	-1.02		111	-2.14		111	-2.20		111	-0.54	
A31	Chu	2009	Solifenacin ER 10mg q.d	[30]	225	-2.20	2.85	340	-2.90	2.95	305	-4.00	3.67	267	-0.60	
			Placebo	[1]	237	-1.20	2.77	332	-1.50	2.92	306	-2.40	3.85	279	-0.40	
A32	Chuang	2014	Lipo-BoNTA 200U	[138]	29	-0.37	0.95	29	-1.55	2.58	29	-2.48	3.89			
			Placebo	[1]	29	-0.31	2.06	29	-0.06	3.56	29	-1.14	3.54			
A33	Davila	2001	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	38	-4.60										
			Oxybutynin transdermal 1.3mg/day	[11]	38	-4.80										
A34	Digesu	2012	Elocalcitol 150mg	[69]	87	-0.50	1.90	87	-1.60	2.80	87	-1.20	2.90			
			Elocalcitol 75mg	[70]	84	-0.70	1.70	84	-1.40	2.50	84	-1.50	3.00			
			Placebo	[1]	86	-0.10	1.70	86	-1.30	2.10	86	-0.80	2.40			
A35	Diokno	2003	Oxybutynin ER 10mg q.d	[8]	391	-4.44		391	-5.43							
			Tolterodine ER 4mg q.d	[4]	399	-4.09		399	-6.56							
A36	Dmochowski	2002	Oxybutynin transdermal 1.3mg/day	[11]	128	-2.64	2.80	128	-1.80	2.60						
			Oxybutynin transdermal 2.6mg/day	[12]	131	-2.41	2.63	131	-1.80	2.40						
			Oxybutynin transdermal 3.9mg/day	[10]	123	-3.09	2.50	123	-2.30	2.50						
			Placebo	[1]	130	-2.74	3.01	130	-1.70	3.00						
A37	Dmochowski	2003	Oxybutynin transdermal 3.9mg/day	[10]	121	-2.90	3.00	121	-1.90	2.70						
			Tolterodine ER 4mg q.d	[4]	123	-3.20	2.80	123	-2.20	2.60						
			Placebo	[1]	117	-2.10	3.00	117	-1.40	2.70						
A38	Dmochowski	2008	Placebo	[1]	276	-1.60	3.32	276	-1.80	3.32						
			Trospium chloride ER 60mg q.d	[44]	267	-2.40	3.27	267	-2.50	3.27						

TABLE B.2: Efficacy data (cont.)

A39	Dmochowski	2010	Placebo	[1]	43	-2.49											
			OnaBoNT-A 50u trigone sparing	[74]	56	-2.96											
			OnaBoNT-A 100u trigone sparing	[72]	55	-2.63											
			OnaBoNT-A 150u trigone sparing	[75]	50	-3.29											
			OnaBoNT-A 200U trigone sparing	[73]	52	-2.80											
			OnaBoNT-A 300u trigone sparing	[76]	55	-2.77											
A40	Dmochowski	2010	Placebo	[1]	445	-2.20		445	-2.10	2.11	445	-2.95					
			Fesoterodine ER 4 - 8mg q.d	[27]	438	-2.00		438	-2.90	2.09	438	-4.54					
A41	Dmochowski	2014	Placebo	[1]	130	0.52		130	0.59								
			Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	130	0.05		130	-0.29								
			Tolterodine IR 2mg b.i.d	[5]	130	0.09		130	-0.24								
A42	Drutz	1999	Placebo	[1]	33	-1.00	2.20	36	-1.10	2.90							
			Tolterodine IR 2mg b.i.d	[5]	60	-1.70	2.00	70	-2.00	2.50							
			Oxybutynin IR 5mg t.i.d	[7]	39	-1.70	1.70	41	-2.00	2.30							
A43	Dubeau	2014	Placebo	[1]	281	-2.20		281	-1.50		281	-2.75		281	-0.68		
			Fesoterodine ER 4 - 8mg q.d	[27]	281	-2.84		281	-2.34		281	-4.15		281	-0.97		
A44	Enzelsberger	1991	Estradiol 1mg intravaginally	[127]				15	-1.40					15	-2.10		
			Estradiol 3mg intravaginally	[128]				15	-3.80					15	-3.60		
			Placebo	[1]				10	-1.90					10	-1.80		
A45	Enzelsberger	1991	Lidocaine gel 2x6ml	[129]				15	-5.00					15	-2.80		
			Emepronium Bromide IR 200mg t.i.d	[130]				15	-3.50					15	-3.70		
A46	Enzelsberger	1995	Placebo	[1]				20	-1.30					20	-1.10		
			Oxybutynin 20mg intravesically q.d	[106]				23	-3.50					23	-3.30		
A47	Enzelsberger	1995	Oxybutynin 20mg intravesically q.d	[106]				21	-3.60					21	-3.40		
			Placebo	[1]				18	-3.30					18	-0.80		
A48	Finazzi-Agro	2010	Sham Therapy	[3]	17	-0.13		17	-1.10								
			Percutaneous tibial nerve stimulation	[83]	18	-0.77		18	-4.10								
A49	FitzGerald	2008	Tolterodine ER 4mg q.d + BT	[87]										134	-0.51	1.01	
			Tolterodine ER 4mg q.d	[4]										137	-0.28	1.15	
A50	Franzen	2010	Electrostimulation	[80]				32	-2.60	2.45							
			Tolterodine ER 4mg q.d	[4]				31	-2.90	1.99							
A51	Frenkl	2010	Serlopitant 0.25mg q.d	[107]				110	-1.10	1.87							
			Serlopitant 1mg q.d	[108]				110	-0.80	1.87							
			Serlopitant 4mg q.d	[109]				114	-1.10	1.91							
			Tolterodine ER 4mg q.d	[4]				114	-1.50	1.91							
			Placebo	[1]				109	-0.50	1.86							
A52	Fukuda	2013	Solifenacin ER 5mg q.d	[29]	13	-0.90		22	-2.20		22	-1.70		21	-0.90		
			Propiverine ER 20mg q.d	[41]	13	-0.60		22	-1.70		20	-1.20		20	-0.10		
A53	Giannitsas	2004	Oxybutynin IR 5mg t.i.d	[7]				107	-0.80								
			Tolterodine IR 2mg b.i.d	[5]				107	-0.90								
A54	Gittelman	2014	Placebo	[1]	112	-1.97	2.07	112	-1.10								
			Oxybutynin vaginal ring 4mg q.d	[16]	115	-2.40	2.35	115	-1.80								
			Oxybutynin vaginal ring 6mg q.d	[17]	96	-2.39	2.04	96	-2.10								
A55	Goldfischer	2015	Oxybutynin gel 84mg/day	[134]	211	-2.90	3.48	211	-2.60	2.66							
			Oxybutynin gel 56mg/day	[135]	198	-3.30	4.04	198	-2.20	2.88							
			Placebo	[1]	192	-2.60	4.12	192	-1.90	3.34							
A56	Gotoh	2011	Propiverine ER 20mg q.d	[41]	291	-1.18	1.87	291	-1.86	1.87	291	-2.84	2.57	291	-0.29	0.65	
			Placebo	[1]	274	-0.68	1.14	274	-1.36	1.69	274	-1.99	2.62	274	-0.25	0.72	
A57	Haab	2014	Placebo	[1]				59	-1.84	3.05	59	-0.83	1.73				
			Netupitant 50mg q.d	[110]				60	-1.93	2.85	60	-1.00	1.71				
			Netupitant 100mg q.d	[111]				59	-2.04	2.97	59	-1.37	1.77				
			Netupitant 200mg q.d	[112]				55	-2.17	3.04	55	-0.91	1.74				
A58	Halaska	2003	Trospium chloride IR 20mg b.i.d	[45]	267	-1.00		267	-1.20		267	-1.60					
			Oxybutynin IR 5mg b.i.d	[18]	90	-1.00		90	-1.50		90	-1.70					
A59	Hassouna	2000	Sacral nerve stimulation	[81]				25	-7.60								
			Control	[2]				26	0.50								
A60	Herschorn	2004	Tolterodine	[105]	17	-1.46	2.79	28	-2.18	4.89				30	-0.07	0.91	
			Tolterodine + BT	[94]	18	-1.10	3.02	33	-1.82	3.41				34	-0.44	1.13	
A61	Herschorn	2008	Placebo	[1]	201	-1.50	1.42	201	-1.70	2.84	201	-1.90	2.84				
			Tolterodine ER 4mg q.d	[4]	402	-1.90	2.00	402	-2.30	2.00	402	-2.80	4.01				
A63	Herschorn	2010	Placebo	[1]	334	-2.13		334	-1.44		334	-1.65		334	-0.60		
			Tolterodine ER 4mg q.d	[4]	684	-2.50		684	-1.89		684	-2.86		684	-0.60		
			Fesoterodine ER 8mg q.d.	[26]	679	-2.40		679	-2.21		679	-3.52		679	-0.60		
A64	Herschorn	2013	Mirabegron ER 25mg q.d	[50]	254	-1.36	1.91	410	-1.65	2.63	410	-1.68	3.24				
			Mirabegron ER 50mg q.d	[51]	257	-1.38	1.92	426	-1.60	2.48	426	-1.94	3.10				
A65	Hill	2006	Darifenacin ER 7.5mg q.d	[39]	108	-1.16		108	-1.70		108	-1.80		108	-0.27		
			Darifenacin ER 15mg q.d	[40]	107	-1.49		107	-1.90		107	-2.30		107	-0.24		
			Darifenacin ER 30mg q.d	[38]	115	-1.63		115	-2.20		115	-3.00		115	-0.29		
			Placebo	[1]	109	-0.84		109	-1.10		109	-1.20		109	-0.06		
A66	Ho	2010	Solifenacin ER 5mg q.d	[29]	39	-2.79	2.82	39	-2.56		39	-1.70	3.07				
			Tolterodine ER 4mg q.d	[4]	36	-4.67	9.29	36	-2.44		36	-1.15	2.68				
A67	Holmes	1989	Oxybutynin IR 5mg t.i.d	[7]				23	-1.97					23	-0.47		
			Propantheline Bromide 15mg t.i.d	[113]				23	-1.00					23	-0.20		
A68	Homma	2003	Tolterodine ER 4mg q.d	[4]	239	-2.28	2.57	239	-2.00	2.32							
			Oxybutynin IR 3mg t.i.d	[19]	244	-2.38	1.02	244	-2.10	3.13							
			Placebo	[1]	122	-1.26	1.47	122	-1.10	3.19							
A69	Homma	2008	Placebo	[1]	95	-1.21		95	-1.08		95	-2.02					
			Imidafenacin IR 0.05mg b.i.d	[35]	91	-1.42		91	-1.71		91	-2.61					
			Imidafenacin IR 0.1mg b.i.d	[36]	93	-1.55		93	-1.59		93	-2.55					
			Imidafenacin IR 0.25mg b.i.d	[37]	76	-2.00		76	-2.34		76	-3.40					
A70	Homma	2009	Placebo	[1]	143	-1.24		143	-1.08	1.62	143	-1.94					
			Imidafenacin IR 0.1mg b.i.d	[36]	318	-1.67		318	-1.52	1.70	318	-2.35					
			Propiverine ER 20mg q.d	[41]	305	-1.81		305	-1.80	1.86	305	-2.79					
A71	Hsiao	2011	Solifenacin ER 5mg q.d	[29]	26	-0.90	2.35	26	-5.80	10.54	26	-5.70	10.28	26	-1.10	2.72	
			Tolterodine ER 4mg q.d	[4]	22	-4.10	17.83	22	-6.60	14.72	22	-3.00	8.97	22	-2.00	3.83	
A72	Huang	2012	Fesoterodine ER 4 - 8mg q.d	[27]	303	-2.90	2.70	303	-1.30	1.70				303	-0.50	1.10	
			Placebo	[1]	301	-2.10	2.90	301	-0.70	1.75				301	-0.20	1.20	
A73	Huo	2013	Solifenacin ER 5mg q.d	[29]				24	-9.00								
			Naftopidil 25mg q.d	[114]				22	-5.00								
			Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]				21	-3.00								
A74	Jabs	2013	OnaBoNT-A 100u trigone sparing	[72]	11	-4.10		11	-2.70					11	-1.40		
			Placebo	[1]	10	0.40		10	0.50					10	-0.30		
A75	Jacquetin	2001	Placebo	[1]	51	-0.40	1.90	51	-1.20	2.70							
			Tolterodine IR 1mg b.i.d	[6]	97	-1.10	2.20	97	-1.40	2.80							
			Tolterodine IR 2mg b.i.d	[5]	103	-1.30	1.80	103	-1.40	4.30							
A76	Johnson	2005	Bladder Training (BT)/Behaviour Therapy	[85]										47	-0.50	0.60	
			Oxybutynin ER 2.5mg q.d	[20]										46	-0.20	0.50	
			Placebo	[1]										38	0.10	0.70	
A77	Junemann	2005	Propiverine IR 15mg b.i.d	[43]	100	-1.20	2.01	100	-2.75	3.10	100	-3.26	3.50				
			Tolterodine IR 2mg b.i.d	[5]	101	-0.91	1.48	101	-3.07	2.91	101	-3.04	4.13				
A78	Junemann	2006	Propiverine IR 15mg b.i.d	[43]	395	-2.21											
			Propiverine ER 30mg q.d	[42]	391	-2.47											
			Placebo	[1]	202	-1.78											
A79	Kafri	2013	Tolterodine ER 4mg q.d	[4]	42	-0.80		42	-1.30								
			Bladder Training (BT)/Behaviour Therapy	[85]	41	-0.26		41	-2.00								
			Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	40	-0.36		40	-1.90								
			PFMT + BT	[89]	41	-0.53		41	-3.40								

TABLE B.2: Efficacy data (cont.)

A80	Kaplan	2011	Placebo	[1]	462	-2.40		462	-2.13		462	-2.95		462	-0.57	
			Tolterodine ER 4mg q.d	[4]	942	-2.60		942	-2.48		942	-3.64		942	-0.77	
			Fesoterodine ER 4mg q.d	[25]	930	-2.60		930	-2.75		930	-4.41		930	-0.73	
A82	Karram	2009	Solifenacin ER 5 - 10mg q.d	[31]	229	-2.10	2.39	348	-2.67	3.31	348	-3.91	3.54			
			Placebo	[1]	224	-1.24	2.30	337	-1.94	3.30	336	-2.73	3.84			
A83	Kaya	2011	Trospium chloride IR 15mg t.i.d	[46]	15	-0.30	1.30	15	0.30	3.40				15	-0.80	1.00
			Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	15	-0.90	0.90	15	-5.20	5.50				15	-2.30	2.60
			Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	16	-1.60	1.50	16	-4.70	5.70				16	-1.70	1.00
A84	Khullar	2004	Placebo	[1]	285	-1.14	2.09	285	-1.30	2.30	285	-0.90	2.70			
			Tolterodine ER 4mg q.d	[4]	569	-1.76	2.07	569	-2.10	2.40	569	-2.00	3.00			
A85	Khullar	2013	Placebo	[1]	291	-1.13	2.15	494	-1.37	2.56	494	-1.65	3.33			
			Mirabegron ER 50mg q.d	[51]	293	-1.62	2.35	493	-1.94	2.58	493	-2.25	3.33			
			Mirabegron ER 100mg q.d	[52]	281	-1.51	2.15	496	-1.75	2.45	496	-1.96	3.34			
			Tolterodine ER 4mg q.d	[4]	300	-1.21	2.37	495	-1.57	2.74	495	-2.07	3.34			
A87	Kuo	2011	OnaBoNT-A 100u trigone sparing	[72]	37	-1.32	2.31	37	-1.80	3.16	37	0.06	2.81			
			OnaBoNT-A 100u bladder body + trigone	[78]	35	-1.09	1.69	35	-1.80	3.16	35	-0.55	3.47			
			OnaBoNT-A 100u bladder base + trigone	[79]	33	-0.89	6.79	33	-2.09	3.76	33	-0.42	5.47			
A88	Kuo	2014	OnaBoNT-A 200u trigone sparing	[73]	12	-0.17		12	-3.16		12	-3.34				
			Placebo	[1]	12	-0.66		12	-0.67		12	-1.00				
A90	Kuo	2015	Placebo	[1]	24	-0.58	2.51	68	-1.28	3.49	68	-1.49	4.84	65	-0.55	1.54
			Mirabegron ER 50mg q.d	[51]	33	-1.37	3.72	76	-2.12	2.91	76	-1.97	3.49	73	-0.50	1.14
			Tolterodine ER 4mg q.d	[4]	26	-0.79	2.02	74	-0.98	3.18	74	-1.96	4.33	71	-0.45	1.25
A91	Kurz	1993	Estriol 1mg intravesically	[131]				21	-2.10					21	-1.40	
			Placebo	[1]				21	-0.60					21	0.90	
A92	Lauti	2008	Bladder Training (BT)/Behaviour Therapy	[85]	18	-1.40		18	-1.10		18	-1.60		18	-0.10	
			Oxybutynin ER 2.5mg q.d	[20]	16	-0.90		16	-1.70		16	-1.60		16	-0.60	
			Oxybutynin ER 2.5mg q.d + BT	[92]	12	-1.20		12	-1.70		12	-1.80		12	-0.20	
A93	Lee	2002	Tolterodine IR 2mg b.i.d	[5]	112	-2.20	2.30	112	-2.60	2.90						
			Oxybutynin IR 5mg b.i.d	[18]	116	-1.40	1.80	116	-1.80	4.20						
A94	Lee	2010	Propiverine ER 20mg q.d	[41]				142	-3.56	3.22				142	-0.52	1.01
			Placebo	[1]				79	-2.58	4.18				79	-0.42	1.04
A95	Lee	2013	Imidafenacin 0.1mg b.i.d	[36]	77	-0.96	2.27	96	-3.06	3.69	77	-4.21	4.20	77	-0.67	1.09
			Fesoterodine ER 4mg q.d	[25]	82	-0.82	1.68	64	-2.43	3.61	82	-3.71	3.21	82	-0.47	1.04
A96	Lehtoranta	2002	Oxybutynin intravesically 5mg t.i.d	[14]	9	-0.70		9	-1.20							
			Placebo	[1]	9	0.50		9	0.50							
A98	Mak	2007	Reflexology	[71]	54	0.00	0.74	54	-2.18	2.80	54	-0.27	0.94	54	-1.00	0.74
			Control	[2]	43	0.00	0.74	43	-1.04	2.92	43	-0.48	1.12	43	-1.00	0.74
A99	Malone-Lee	2001	Placebo	[1]	33	0.00	1.03	42	0.00	1.16						
			Tolterodine IR 1mg b.i.d	[6]	44	-0.30	1.18	59	-0.70	3.72						
			Tolterodine IR 2mg b.i.d	[5]	51	-0.70	2.00	73	-0.70	1.74						
A100	Malone-Lee	2009	Tolterodine IR 2mg b.i.d	[5]	104	-1.30		190	-1.70							
			Oxybutynin IR 2.5 - 5mg b.i.d	[24]	102	-1.80		188	-1.70							
A101	Malone-Lee	2009	Placebo	[1]				155	-1.61	2.61						
			Tolterodine ER 4mg q.d	[4]				130	-2.48	2.28						
A102	Marencak	2011	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	37	-1.10	3.04	96	-1.30	2.94	96	-1.30	3.92			
			Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	47	-0.50	1.37	97	-0.70	1.97	97	-0.90	2.95			
			Pregabalin 150mg b.i.d	[62]	41	-0.50	2.56	94	-1.00	1.94	94	-1.30	1.94			
			Tolterodine ER 4mg q.d	[4]	51	-1.10	1.43	101	-0.50	2.01	101	-1.00	3.01			
			Placebo	[1]	50	-0.40	1.41	98	-0.50	1.98	98	-0.80	2.97			
A103	Martinez-Garcia	2009	Placebo	[1]	20	-0.61	1.56	20	-0.75	2.21	20	-0.68	1.42			
			Cizilirtine citrate 200mg b.i.d	[56]	20	-1.07	3.04	20	-0.88	2.66	20	-1.34	2.68			
			Cizilirtine citrate 400mg b.i.d	[57]	16	-1.40	1.30	16	-2.08	3.14	16	-3.32	3.09			
A104	Mattiasson	2003	Tolterodine IR 2mg b.i.d + BT	[93]	141	-1.70										
			Tolterodine IR 2mg b.i.d	[5]	160	-2.00										
A105	Mattiasson	2010	Solifenacin ER 5 - 10mg q.d	[31]	305	-0.95		305	-2.35		305	-2.40				
			Solifenacin ER 5 - 10mg q.d + BT	[77]	297	-1.00		297	-3.05		297	-2.50				
A106	Mazur	1995	Propiverine IR 15mg b.i.d	[43]				45	-3.40							
			Propiverine IR 30mg b.i.d	[117]				45	-4.60							
			Propiverine IR 45mg t.i.d	[118]				48	-3.50							
			Propiverine ER 60mg q.d	[119]				39	-2.60							
A107	Meyhoff	1983	Emepronium bromide ER 200mg q.d	[63]	19	-0.33		19	-0.17					19	-0.17	
			Flavoxate chloride 200mg q.d	[64]	19	-0.34		19	0.33					19	0.00	
			Placebo	[1]	19	-0.67		19	-0.33					19	-0.67	
A108	Millard	1999	Placebo	[1]	55	-1.30	2.50	64	-1.40	2.30						
			Tolterodine IR 1mg b.i.d	[6]	109	-1.70	2.80	123	-2.30	3.00						
			Tolterodine IR 2mg b.i.d	[5]	117	-1.70	2.50	129	-2.30	2.10						
A109	Millard	2004	Tolterodine IR 2mg b.i.d	[5]	252	-2.15	2.70				252	-2.20	3.60			
			Tolterodine IR 2mg b.i.d + PFMT	[95]	223	-2.15	3.00				223	-1.90	4.00			
A110	Nitti	2007	Placebo	[1]	271	-0.96	2.80	271	-1.08	2.96	271	-0.79	3.29	271	-0.39	
			Fesoterodine ER 4mg q.d	[25]	282	-1.65	2.69	282	-1.61	3.02	282	-1.91	3.36	282	-0.58	
			Fesoterodine ER 8mg q.d	[26]	279	-2.28	2.67	279	-2.09	3.01	279	-2.30	3.34	279	-0.55	
A112	Nitti	2013	Placebo	[1]	453	-1.13	2.34	453	-1.05	2.77	453	-0.82	3.41	453	-0.38	
			Mirabegron ER 50mg q.d	[51]	442	-1.47	2.31	442	-1.66	2.73	442	-1.57	3.36	442	-0.57	
			Mirabegron ER 100mg q.d	[52]	433	-1.63	2.50	433	-1.75	2.91	433	-1.76	3.54	433	-0.57	
A113	Nitti	2013	Placebo	[1]	272	-0.87		272	-0.91	2.65	272	-1.21	3.83	272	-0.24	1.09
			OnaBoNT-A 100u trigone sparing	[72]	278	-2.65		278	-2.15	3.02	278	-2.93	4.21	278	-0.45	1.28
A114	Norton	1994	Terodiline 25mg b.i.d	[28]	46	-1.51	3.43	46	-0.70	1.90				46	0.00	0.80
			Placebo	[1]	46	-0.54	1.30	46	-1.10	2.50				46	0.00	0.50
A115	O'Reilly	2008	Electromagnetic stimulation	[125]				33	-1.00							
			Sham therapy	[3]				30	0.00							
A116	Ohlstein	2012	Placebo	[1]	85	-2.00	0.26	85	-2.10	0.28						
			Solabegron IR 50mg b.i.d	[54]	88	-2.20	0.27	88	-2.60	0.29						
			Solabegron IR 125mg b.i.d	[55]	85	-2.60	0.26	85	-3.00	0.28						
A117	Olmo	2013	Electrostimulation	[80]				11	-4.00	8.21	11	-2.00	3.19	11	-0.91	0.94
			Percutaneous tibial nerve stimulation	[83]				11	-3.54	3.91	11	-3.27	2.83	11	-2.27	2.37
A118	Oreskovic	2012	Placebo	[1]				86	-0.11		86	-0.01		86	0.50	
			Solifenacin ER 5mg q.d	[29]				85	-0.49		85	-0.13		85	-0.45	
A119	Orri	2014	Placebo	[1]				6	-0.80							
			Tolterodine ER 4mg q.d	[4]				12	-2.40							
A120	Ouslander	1993	Terodiline IR 25mg b.i.d	[28]	40	-1.16	1.94	40	-0.76	1.73				40	-0.04	0.84
			Placebo	[1]	41	-0.60	1.83	41	-0.18	1.47				41	-0.26	0.56
A121	Ozdedeli	2010	Trospium chloride IR 15mg t.i.d	[46]	15	-0.30		15	0.30					15	-0.80	
			Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	15	-0.90		15	-5.20					15	-2.30	
			Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	16	-1.60		16	-4.70					16	-1.70	
A122	Park	2014	Imidafenacin 0.1mg b.i.d	[36]	75	-2.16										
			Propiverine ER 20mg q.d	[41]	76	-2.04										
A123	Peters	2009	Percutaneous tibial nerve stimulation	[83]	41	-1.00	2.20	41	-2.40	4.00	41	-2.20	4.30	41	-0.70	1.00
			Tolterodine ER 4mg q.d	[4]	43	-1.70	3.80	43	-2.50	3.90	43	-2.90	4.80	43	-0.60	1.70
A124	Peters	2010	Sham therapy	[3]	110	-0.30		110	-1.50	2.40	110	-2.00		110	-0.30	1.40
			Percutaneous tibial nerve stimulation	[83]	110	-1.30		110	-2.40	2.50	110	-3.70		110	-0.70	1.20
A125	Preik	2004	Oxybutynin ER 5 - 30mg/day	[22]	53	-3.26										
			Oxybutynin IR 5 - 20mg	[23]	52	-2.89										
A126	Preyer	2015	Percutaneous tibial nerve stimulation	[83]	18	-1.50		18	-0.90							
			Tolterodine IR 2mg b.i.d	[5]	18	-1.00		18	-2.20							
A128	Rios	2007	Resiniferatoxin 50nM	[67]	34	-0.28		34	-0.63					34	-0.44	
			Placebo	[1]	24	-0.86		24	-0.69					24	-0.39	

TABLE B.2: Efficacy data (cont.)

A129	Rogers	2008	Placebo	[1]	189	-1.40	1.37	189	-2.30	2.75						
			Tolterodine ER 4mg q.d	[4]	182	-1.80	1.35	182	-3.30	2.70						
A130	Rudy	2006	Placebo	[1]	325	-1.29		325	-1.76					325	-0.29	
			Trospium chloride IR 20mg b.i.d	[45]	323	-1.86		322	-2.67					322	-0.57	
A131	Rufford	2003	Estradiol 25mg	[68]	20	0.00		20	0.10							
			Placebo	[1]	20	0.00		20	0.50							
A132	Sahai	2007	OnaBoNT-A 200u trigone sparing	[73]	16	-3.50		16	-6.19		16	-8.19				
			Placebo	[1]	18	-0.71		18	-1.14		18	-0.92				
A133	Sancaktar	2010	Tolterodine ER 4mg q.d	[4]	18	-1.50		18	-6.40		18	-5.10				
			Tolerodine ER 4mg q.d + Neurostimulation	[96]	20	-2.29		20	-7.70		20	-6.70				
A134	Schmidt	1999	Control	[2]	42	2.00										
			Sacral nerve stimulation	[81]	34	-7.10										
A135	Schreiner	2010	Electrostimulation + PFMT + BT	[97]	25	-2.10	1.77	25	-1.40	2.00				25	-1.60	1.10
			PFMT + BT	[89]	26	-0.43	0.53	26	-0.20	0.90				26	-0.40	1.10
A136	Song	2006	Bladder Training (BT)/Behaviour Therapy	[85]				26	-2.80		26	-1.20		26	-0.90	
			Tolterodine IR 2mg b.i.d	[5]				32	-3.50		32	-1.70		32	-1.10	
			Tolterodine IR 2mg b.i.d + BT	[93]				31	-4.00		31	-1.80		31	-1.40	
A137	Song	2015	Tarafenacin 0.2mg q.d	[90]	77	-0.42	1.91	77	-1.92	2.45	77	-3.08	3.71	77	-0.16	0.82
			Tarafenacin 0.4mg q.d	[82]	76	-0.66	1.71	76	-2.43	2.21	76	-3.42	3.93	76	-0.42	0.92
			Placebo	[1]	72	-0.53	2.20	72	-1.77	2.95	72	-2.68	4.00	72	-0.30	1.06
A138	Soomro	2001	Oxybutynin IR 2.5 - 5mg b.i.d	[24]				43	-2.00							
			Electrostimulation	[80]				43	-2.00							
A139	Staskin	2007	Placebo	[1]	303	-1.93	2.79	303	-1.99	2.79						
			Trospium chloride ER 60mg q.d	[44]	298	-2.48	2.93	298	-2.81	2.59						
A140	Staskin	2009	Oxybutynin chloride topical gel 1g/day	[13]	389	-3.00	2.70	389	-2.70	3.20				389	-0.75	1.40
			Placebo	[1]	400	-2.50	3.10	400	-2.00	2.80				400	-0.65	1.30
A141	Steers	2005	Darifenacin ER 7.5mg q.d	[39]	261	-0.43		261	-1.90		261	-2.30		261	-0.29	
			Placebo	[1]	123	-0.26		123	-1.00		123	-0.90		123	-0.14	
A142	Steers	2007	Placebo	[1]	129	-0.18		129	-0.82					129	-0.21	
			Duloxetine 60mg b.i.d	[66]	99	-0.87		99	-1.90					99	0.01	
A143	Subak	2002	Bladder Training (BT)/Behaviour Therapy	[85]	66	-0.63		66	-0.08							
			Control	[2]	57	-0.32		57	0.08							
A144	Swift	2003	Tolterodine ER 4mg q.d	[4]	417	-1.69	2.57	417	-1.90	3.40						
			Tolterodine IR 2mg b.i.d	[5]	408	-1.44	2.23	408	-1.70	2.90						
			Placebo	[1]	410	-1.03	2.17	410	-1.20	2.90						
A145	Tang	2014	Sacral Nerve Stimulation + Tolterodine ER 2mg q.d	[136]				120	-12.50							
			Tolterodine ER 2mg q.d	[137]				120	-4.00							
A146	Tapp	1989	Placebo	[1]	36	-1.00		36	-0.10					36	-0.10	
			Terodiline 25mg b.i.d	[28]	34	-2.20		34	-2.60					34	-0.40	
A147	Tincello	2000	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	10	-1.00		10	-3.50		10	-2.00				
			Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	8	-2.50		8	0.00		8	-1.00				
A148	Tincello	2012	OnaBoNT-A 200U trigone sparing	[73]	86	-3.11	3.99	86	-3.57	5.07	86	-3.64	3.99			
			Placebo	[1]	86	-0.52	3.73	86	-0.63	3.24	86	-1.38	4.19			
A149	Tseng	2009	Tolterodine IR 2mg b.i.d	[5]	40	-0.30										
			Tolterodine ER 2mg b.i.d + oestrogen 0.625mg 2xwk	[99]	40	-0.60										
A151	Van Kerrebroeck	2001	Tolterodine ER 4mg q.d	[4]	507	-1.69	2.54	507	-3.50	4.90						
			Tolterodine IR 2mg b.i.d	[5]	514	-1.51	2.41	514	-3.30	4.40						
			Placebo	[1]	508	-0.99	2.20	508	-2.20	4.00						
A152	Van Kerrebroeck	2009	Placebo	[1]	487	-0.99	2.81									
			Tolterodine ER 4mg q.d	[4]	500	-1.69	2.81									
A153	Vardy	2009	Solifenacin ER 5mg - 15mg q.d	[34]	377	-1.85	2.70	377	-2.23		377	-3.05		377	-0.63	
			Placebo	[1]	374	-1.24	2.70	374	-1.36		374	-1.84		374	-0.48	
A154	Versi	2000	Oxybutynin ER 5 - 30mg/day	[22]	111	-2.25										
			Oxybutynin IR 5 - 20mg	[23]	115	-2.20										
A155	Visco	2012	Solifenacin/trospium + placebo injection	[100]	118	-3.40	3.16									
			OnaBoNT-A 100U trigone sparing	[72]	113	-3.30	3.16									
A156	Wagg	2013	Placebo	[1]	176	-0.91		382	-1.15	2.60	382	-2.50	4.54	382	-0.30	1.13
			Fesoterodine ER 4mg q.d	[25]	172	-1.00		374	-2.12	2.30	374	-3.85	4.08	374	-0.57	1.13
A157	Wang	2006	Electrostimulation	[80]	24	0.00	0.61	24	-3.00	2.20	24	-10.15	2.43	24	-0.80	1.05
			Oxybutynin IR 2.5mg t.i.d	[21]	23	0.00	0.30	23	-2.15	2.29	23	-3.00	1.81	23	0.00	0.46
			Placebo	[1]	21	0.00	0.46	21	-0.75	1.34	21	-1.30	1.90	21	0.00	0.53
A158	Wang	2009	Electrostimulation	[80]	26	0.00	0.93	26	-3.60	1.41	26	-2.80	1.60	26	0.00	1.02
			Oxybutynin IR 2.5mg t.i.d	[21]	24	-0.40	0.53	24	-5.30	2.19	24	-2.35	1.41	24	0.45	1.28
			Placebo	[1]	23	-0.20	0.71	23	-1.60	1.96	23	0.30	1.66	23	0.00	1.03
A159	Weiss	2013	Placebo	[1]	474	-1.28	2.02	474	-1.86	2.86	474	-2.81	3.81	474	-0.85	1.13
			Fesoterodine ER 4 - 8mg q.d	[27]	463	-1.44	2.42	463	-2.43	2.70	463	-3.53	3.98	463	-1.02	1.15
A160	Yamaguchi	2007	Placebo	[1]	283	-0.72	1.95	395	-0.94	2.29	395	-1.28	2.90	361	-0.30	0.91
			Solifenacin ER 5mg q.d	[29]	274	-1.60	1.81	383	-1.93	1.97	383	-2.41	2.88	344	-0.41	0.96
			Solifenacin ER 10mg q.d	[30]	270	-1.59	2.12	371	-2.19	2.09	371	-2.78	2.82	334	-0.46	0.90
			Propiverine ER 20mg q.d	[41]	295	-1.25	2.79	384	-1.87	2.70	384	-2.30	3.08	348	-0.43	1.21
A161	Yamaguchi	2011	Placebo	[1]	309	-1.01	2.69	309	-0.59	4.39	309	-1.00	5.38	243	-0.18	1.87
			Fesoterodine ER 4mg q.d	[25]	314	-1.35	2.71	314	-1.15	4.38	314	-1.65	5.42	256	-0.21	1.88
			Fesoterodine ER 8mg q.d	[26]	306	-1.40	2.72	306	-1.25	4.42	306	-1.66	5.44	257	-0.29	1.92
A162	Yamaguchi	2014	Placebo	[1]	264	-0.66	1.86	368	-0.86	2.35	368	-1.37	3.19	322	-0.36	1.06
			Mirabegron ER 50mg q.d	[51]	266	-1.12	1.48	369	-1.67	2.21	369	-1.85	2.56	323	-0.44	0.93
			Tolterodine ER 4mg q.d	[4]	240	-0.97	1.61	368	-1.40	2.18	368	-1.66	2.56	332	-0.42	0.85
A163	Yamaguchi	2014	Oxbutynin patch 73.5mg	[15]	391	-1.10	1.40	555	-1.89	2.04	555	-1.92	2.21	481	-0.52	0.79
			Propiverine ER 20mg q.d	[41]	401	-1.07	1.44	559	-1.85	2.10	559	-1.94	2.45	489	-0.49	0.83
			Placebo	[1]	259	-0.95	1.57	373	-1.44	2.23	373	-1.51	2.33	329	-0.42	0.83
A164	Yamaguchi	2014	Placebo	[1]	140	-0.64		211	-1.18	2.16	211	-1.83		168	-0.24	
			Mirabegron ER 25mg q.d	[50]	134	-1.29		209	-1.94	2.16	208	-2.15		179	-0.49	
			Mirabgeron ER 50mg q.d	[51]	144	-1.20		208	-2.12	2.38	208	-2.24		176	-0.38	
			Mirabegron ER 100mg q.d	[52]	150	-1.28		207	-1.97	1.97	207	-2.48		180	-0.39	
A167	Yokoyama	2014	Placebo	[1]										174	-0.56	1.00
			Fesoterodine ER 4mg q.d	[25]										180	-0.63	0.90
			Fesoterodine ER 8mg q.d	[26]										201	-0.77	1.00
A168	Yoon	2003	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]				13	-0.80							
			Bladder Training (BT)/Behaviour Therapy	[85]				19	-6.90							
			Control	[2]				12	1.10							
A169	Zat'ura	2010	Cizolirtine citrate 400mg b.i.d	[57]	52	-1.20	1.40	52	-5.50	5.50						
			Placebo	[1]	54	-0.60	1.90	54	-2.50	5.60						
			Oxybutynin IR 5mg t.i.d	[7]	26	-1.40	1.80	26	-6.70	6.10						
A170	Zellner	2009	Trospium chloride IR 45mg t.i.d	[47]	615	-1.57		615	-2.22	2.00						
			Oxybutynin IR 2.5 - 5mg b.i.d	[24]	611	-1.57		611	-2.35	2.10						
A171	Zimmern	2010	Tolterodine	[105]				153	0.50							
			Tolterodine + BT	[94]				154	-0.20							
A172	Zinner	2002	Tolterodine ER 4mg q.d	[4]	506	-1.68	2.55	506	-1.70	3.37						
			Placebo	[1]	507	-0.98	2.19	507	-1.15	2.89						
A173	Zinner	2005	Darifenacin ER 15mg q.d.	[40]	58	-1.44		58	-1.14		58	-1.27				
			Darifenacin ER 30mg q.d	[38]	58	-1.74		58	-1.62		58	-1.63				
			Oxybutynin IR 5mg t.i.d	[7]	58	-1.65		58	-1.23		58	-1.10				
			Placebo	[1]	58	-0.91		58	-0.85		58	-0.51				
A174	Zinner	2006	Darifenacin ER 15mg q.d	[40]	212	-1.80	2.77	212	-2.20	4.22	212	-2.60	4.08			
			Placebo	[1]	220	-1.40	2.82	220	-1.80	4.30	220	-2.23	4.15			

B.4 Safety and tolerability data

TABLE B.3: Safety and tolerability data

					Adverse events			Discontinuations due to adverse events			Discontinuations due to lack of efficacy		
Ref	**First Author**	**Year**	**Treatment**		**r**	**n**	**%**	**r**	**n**	**%**	**r**	**n**	**%**
A2	Abrams	1998	Placebo	[1]	46	57	80.7%	7	57	12.3%			
			Tolterodine IR 2mg b.i.d	[5]	105	118	89.0%	10	118	8.5%			
			Oxybutynin IR 5mg t.i.d	[7]	114	118	96.6%	20	118	16.9%			
A4	Anderson	1999	Oxybutynin ER 5mg-30mg	[22]	46	53	86.8%						
			Oxybutynin IR 5mg-20mg	[23]	49	52	94.2%						
A5	Appell	2001	Oxybutynin ER 10mg q.d	[8]				14	185	7.6%	3	185	1.6%
			Tolterodine IR 2mg b.i.d	[5]				15	193	7.8%	1	193	0.5%
A6	Barkin	2004	Oxybutynin ER 15mg q.d	[9]				11	65	16.9%			
			Oxybutynin IR 5mg t.i.d	[7]				12	60	20.0%			
A7	Batista	2015	Mirabegron ER 50mg q.d	[51]	274	936	29.3%	14	936	1.5%	5	936	0.5%
			Solifenacin ER 5mg q.d	[29]	282	934	30.2%	16	934	1.7%	6	934	0.6%
A12	But	2012	Solifenacin ER 5mg q.d	[29]				4	40	10.0%	2	40	5.0%
			Darifenacin ER 7.5mg q.d	[39]				4	37	10.8%	2	37	5.4%
A13	Cardozo	2004	Placebo	[1]				10	301	3.3%	2	301	0.7%
			Solifenacin ER 5mg q.d	[29]				7	299	2.3%	2	299	0.7%
			Solifenacin ER 10mg q.d	[30]				12	307	3.9%	2	307	0.7%
A14	Cardozo	2005	Darifenacin ER 30mg q.d	[38]	27	36	75.0%	4	36	11.1%			
			Placebo	[1]	3	36	8.3%	0	36	0.0%			
A15	Cardozo	2008	Solifenacin ER 5 - 10mg q.d	[31]	19	505	3.8%	18	505	3.6%	14	505	2.7%
			Placebo	[1]	5	223	2.2%	6	223	2.7%	4	223	1.7%
A16	Cartwright	2011	Oxybutynin trandermal 3.9mg/day	[10]				4	48	8.3%			
			Placebo	[1]				0	48	0.0%			
A18	Chancellor	2008	Darifenacin ER 7.5 - 15mg q.d	[104]				6	190	3.2%	0	190	0.0%
			Darifenacin 7.5 - 15mg q.d + BT	[88]				21	205	10.2%	2	205	1.0%
A19	Chapple	2004	Tolterodine IR 2mg b.i.d	[5]	12	37	32.4%	1	37	2.7%	1	37	2.7%
			Placebo	[1]	6	38	15.8%	0	38	0.0%	2	38	5.3%
			Solifenacin ER 2.5mg q.d	[32]	6	41	14.6%	1	41	2.4%	1	41	2.4%
			Solifenacin ER 5mg q.d	[29]	12	37	32.4%	1	37	2.7%	2	37	5.4%
			Solifenacin ER 10mg q.d	[30]	12	35	34.3%	3	35	8.6%	1	35	2.9%
			Solifenacin ER 20mg q.d	[33]	21	37	56.8%	5	37	13.5%	0	37	0.0%
A20	Chapple	2004	Placebo	[1]				10	267	3.7%	2	267	0.7%
			Solifenacin ER 5mg q.d	[29]				9	279	3.2%	2	279	0.7%
			Solifenacin ER 10mg q.d	[30]				7	268	2.6%	1	268	0.4%
			Tolterodine IR 2mg b.i.d	[5]				5	263	1.9%	3	263	1.1%
A21	Chapple	2005	Solifenacin ER 5 - 10mg q.d	[31]				20	578	3.5%			
			Tolterodine ER 4mg q.d	[4]				18	599	3.0%			
A22	Chapple	2006	ZD0947IL 25mg/day	[58]	64	90	71.1%	7	90	7.8%	2	90	2.2%
			Placebo	[1]	60	99	60.6%	8	99	8.1%	2	99	2.0%
A23	Chapple	2007	Placebo	[1]	107	283	37.8%	6	283	2.1%			
			Tolterodine ER 4mg q.d	[4]	144	290	49.7%	9	290	3.1%			
			Fesoterodine ER 4mg q.d	[25]	135	272	49.6%	7	272	2.6%			
			Fesoterodine ER 8mg q.d	[26]	167	287	58.2%	14	287	4.9%			
A24	Chapple	2013	Placebo	[1]				2	271	0.7%			
			OnaBoNT-A 100u trigone sparing	[72]				2	277	0.7%			
A25	Chapple	2013	Placebo	[1]	16	66	24.2%	1	66	1.5%	1	66	1.5%
			Mirabegron IR 100mg b.i.d	[48]	12	65	18.5%	3	65	4.6%	0	65	0.0%
			Mirabegron IR 150mg b.i.d	[49]	16	65	24.6%	5	65	7.7%	0	65	0.0%
			Tolterodine ER 4mg q.d	[4]	17	64	26.6%	2	64	3.1%	0	64	0.0%
A26	Chapple	2013	Placebo	[1]	26	169	15.4%	5	169	3.0%	1	169	0.6%
			Mirabegron ER 25mg q.d	[50]	34	169	20.1%	9	169	5.3%	2	169	1.2%
			Mirabegron ER 50mg q.d	[51]	38	169	22.5%	4	169	2.4%	1	169	0.6%
			Mirabegron ER 100mg q.d	[52]	36	168	21.4%	4	168	2.4%	2	168	1.2%
			Mirabegron ER 200mg q.d	[53]	37	167	22.2%	7	167	4.2%	2	167	1.2%
			Tolterodine ER 4mg q.d	[4]	13	85	15.3%	1	85	1.2%	0	85	0.0%
A27	Chapple	2013	Mirabegron ER 50mg q.d	[51]	485	812	59.7%	52	812	6.4%	34	812	4.2%
			Mirabegron ER 100mg q.d	[52]	503	820	61.3%	49	820	6.0%	25	820	3.0%
			Tolterodine ER 4mg q.d	[4]	508	812	62.6%	49	812	6.0%	45	812	5.5%
A28	Chapple	2014	Placebo	[1]				14	386	3.6%	4	386	1.0%
			Fesoterodine ER 4mg q.d	[25]				27	790	3.4%	8	790	1.0%
			Fesoterodine ER 8mg q.d	[26]				45	779	5.8%	2	779	0.3%
A29	Chapple	2014	Placebo	[1]	46	85	54.1%	5	85	5.9%	0	85	0.0%
			ONO-8539 30mg b.i.d	[59]	41	88	46.6%	2	88	2.3%	2	88	2.3%
			ONO-8539 100mg b.i.d	[60]	37	87	42.5%	4	87	4.6%	0	87	0.0%
			ONO-8539 300mg b.i.d	[61]	47	88	53.4%	6	88	6.8%	0	88	0.0%
			Tolterodine ER 4mg q.d	[4]	41	87	47.1%	0	87	0.0%	0	87	0.0%
A30	Choo	2008	Solifenacin ER 5mg q.d	[29]				5	107	4.7%			
			Solifenacin ER 10mg q.d	[30]				7	111	6.3%			
			Tolterodine IR 2mg b.i.d	[5]				2	111	1.8%			
A31	Chu	2009	Solifenacin ER 10mg q.d	[30]	236	340	69.4%	37	340	10.9%	4	340	1.2%
			Placebo	[1]	197	332	59.3%	18	332	5.4%	3	332	0.9%
A32	Chuang	2014	Lipo-BoNTA 200U	[138]	0	29	0.0%						
			Placebo	[1]	0	29	0.0%						
A35	Diokno	2003	Oxybutynin ER 10mg q.d	[8]				20	391	5.1%	0	391	0.0%
			Tolterodine ER 4mg q.d	[4]				19	399	4.8%	3	399	0.8%
A37	Dmochowski	2003	Oxybutynin transdermal 3.9mg/day	[10]				13	121	10.7%			
			Tolterodine ER 4mg q.d	[4]				2	123	1.6%			
A38	Dmochowski	2008	Placebo	[1]	130	284	45.8%	8	284	2.8%			
			Trospium chloride ER 60mg q.d	[44]	154	280	55.0%	18	280	6.4%			
A39	Dmochowski	2010	Placebo	[1]	8	43	18.6%	0	43	0.0%	2	43	4.7%
			OnaBoNT-A 50u trigone sparing	[74]	17	56	30.4%	1	56	1.8%	2	56	3.6%
			OnaBoNT-A 100u trigone sparing	[72]	20	55	36.4%	0	55	0.0%	3	55	5.5%
			OnaBoNT-A 150u trigone sparing	[75]	20	50	40.0%	1	50	2.0%	1	50	2.0%
			OnaBoNT-A 200U trigone sparing	[73]	20	52	38.5%	0	52	0.0%	0	52	0.0%
			OnaBoNT-A 300u trigone sparing	[76]	22	55	40.0%	1	55	1.8%	1	55	1.8%

TABLE B.3: Safety and tolerability data (cont.)

A40	Dmochowski	2010	Placebo	[1]				21	445	4.7%	16	445	3.6%
			Fesoterodine ER 4 - 8mg q.d	[27]				34	438	7.8%	5	438	1.1%
A42	Drutz	1999	Placebo	[1]	42	56	75.0%	4	56	7.1%			
			Tolterodine IR 2mg b.i.d	[5]	85	109	78.0%	7	109	6.4%			
			Oxybutynin IR 5mg t.i.d	[7]	101	112	90.0%	23	112	20.5%			
A43	Dubeau	2014	Placebo	[1]	120	281	42.7%	14	281	5.0%	9	281	3.2%
			Fesoterodine ER 4 - 8mg q.d	[27]	158	281	56.2%	26	281	9.3%	5	281	1.8%
A46	Enzelsberger	1995	Placebo	[1]				0	20	0.0%	6	20	30.0%
			Oxybutynin 20mg intravesically q.d	[106]				3	23	13.0%	0	23	0.0%
A51	Frenkl	2010	Serlopitant 0.25mg q.d	[107]				7	110	6.4%	0	110	0.0%
			Serlopitant 1mg q.d	[108]				4	110	3.6%	2	110	1.8%
			Serlopitant 4mg q.d	[109]				8	114	7.0%	3	114	2.6%
			Tolterodine ER 4mg q.d	[4]				5	114	4.4%	1	114	0.9%
			Placebo	[1]				1	109	0.9%	0	109	0.0%
A52	Fukuda	2013	Solifenacin ER 5mg q.d	[29]	6	29	20.7%						
			Propiverine ER 20mg q.d	[41]	14	26	53.8%						
A54	Gittelman	2014	Placebo	[1]	75	155	48.4%	4	155	2.6%			
			Oxybutynin vaginal ring 4mg q.d	[16]	89	143	62.2%	6	143	4.2%			
			Oxybutynin vaginal ring 6mg q.d	[17]	96	147	65.3%	12	147	8.2%			
A55	Goldfischer	2015	Oxybutynin gel 84mg/day	[134]				19	211	8.9%			
			Oxybutynin gel 56mg/day	[135]				21	198	10.0%			
			Placebo	[1]				10	192	5.0%			
A56	Gotoh	2011	Propiverine ER 20mg q.d	[41]	80	291	27.5%	2	291	0.7%			
			Placebo	[1]	27	274	9.9%	1	274	0.4%			
A57	Haab	2014	Placebo	[1]	11	62	18.3%	1	62	1.6%	1	62	1.6%
			Netupitant 50mg q.d	[110]	9	62	14.5%	5	62	8.1%	0	62	0.0%
			Netupitant 100mg q.d	[111]	10	61	16.4%	3	61	4.9%	2	61	3.3%
			Netupitant 200mg q.d	[112]	13	61	21.7%	4	61	6.6%	0	61	0.0%
A58	Halaska	2003	Trospium chloride IR 20mg b.i.d	[45]	182	267	68.0%	10	267	3.7%	8	267	3.0%
			Oxybutynin IR 5mg b.i.d	[18]	69	90	77.0%	6	90	6.7%	2	90	2.2%
A61	Herschorn	2008	Placebo	[1]	76	201	38.0%	2	201	1.0%	9	201	4.5%
			Tolterodine ER 4mg q.d	[4]	193	402	48.0%	12	402	3.0%	3	402	0.7%
A62	Herschorn	2010	Solifenacin ER 5mg q.d	[29]	49	68	72.1%	9	68	13.2%	2	68	2.9%
			Oxybutynin IR 5mg t.i.d	[18]	59	64	92.2%	19	64	29.7%	1	64	1.6%
A63	Herschorn	2010	Placebo	[1]				6	334	1.8%	5	334	1.5%
			Tolterodine ER 4mg q.d	[4]				28	684	4.1%	5	684	0.7%
			Fesoterodine ER 8mg q.d.	[26]				44	679	6.5%	13	679	1.9%
A64	Herschorn	2013	Mirabegron ER 25mg q.d	[50]	210	432	48.6%	17	432	3.9%	4	432	0.9%
			Mirabegron ER 50mg q.d	[51]	208	440	47.3%	12	440	2.7%	3	440	0.7%
			Placebo	[1]	217	433	50.1%	15	433	3.5%	15	433	3.5%
A65	Hill	2006	Darifenacin ER 7.5mg q.d	[39]	62	108	57.4%	2	108	1.9%	1	108	0.9%
			Darifenacin ER 15mg q.d	[40]	73	107	68.2%	6	107	5.6%	2	107	1.9%
			Darifenacin ER 30mg q.d	[38]	92	115	80.0%	13	115	11.3%	1	115	0.9%
			Placebo	[1]	54	109	49.5%	3	109	2.8%	2	109	1.8%
A66	Ho	2010	Solifenacin ER 5mg q.d	[29]	15	39	38.5%	1	39	2.6%			
			Tolterodine ER 4mg q.d	[4]	9	36	25.0%	1	36	2.8%			
A68	Homma	2003	Tolterodine ER 4mg q.d	[4]				12	239	5.0%			
			Oxybutynin IR 3mg t.i.d	[19]				42	244	17.1%			
			Placebo	[1]				11	122	9.0%			
A69	Homma	2008	Placebo	[1]	20	95	20.8%	0	95	0.0%			
			Imidafenacin IR 0.05mg b.i.d	[35]	28	91	30.3%	5	91	5.5%			
			Imidafenacin IR 0.1mg b.i.d	[36]	34	93	37.0%	3	93	3.2%			
			Imidafenacin IR 0.25mg b.i.d	[37]	47	76	62.4%	17	76	22.4%			
A70	Homma	2009	Placebo	[1]	99	147	67.3%	8	147	5.4%			
			Imidafenacin IR 0.1mg b.i.d	[36]	234	324	72.2%	11	324	3.4%			
			Propiverine ER 20mg q.d	[41]	250	310	80.6%	19	310	6.1%			
A71	Hsiao	2011	Solifenacin ER 5mg q.d	[29]	10	26	38.5%						
			Tolterodine ER 4mg q.d	[4]	5	22	22.7%						
A72	Huang	2012	Fesoterodine ER 4 - 8mg q.d	[27]	187	303	58.1%	9	303	2.8%			
			Placebo	[1]	149	301	46.1%	6	301	1.9%			
A73	Huo	2013	Solifenacin ER 5mg q.d	[29]				0	24	0.0%	0	24	0.0%
			Naftopidil 25mg q.d	[114]				0	22	0.0%	0	22	0.0%
			Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]				0	21	0.0%	0	21	0.0%
A74	Jabs	2013	OnaBoNT-A 100u trigone sparing	[72]	6	11	54.5%						
			Placebo	[1]	2	10	20.0%						
A75	Jacquetin	2001	Placebo	[1]	16	51	31.4%	1	51	2.0%			
			Tolterodine IR 1mg b.i.d	[6]	39	97	40.2%	3	97	3.1%			
			Tolterodine IR 2mg b.i.d	[5]	55	103	53.4%	2	103	1.9%			
A77	Junemann	2005	Propiverine IR 15mg b.i.d	[43]	42	100	42.0%						
			Tolterodine IR 2mg b.i.d	[5]	43	101	42.6%						
A78	Junemann	2006	Propiverine IR 15mg b.i.d	[43]	152	395	38.5%	15	395	3.8%			
			Propiverine ER 30mg q.d	[42]	134	391	34.3%	11	391	2.8%			
			Placebo	[1]	41	202	20.3%	1	202	0.5%			
A80	Kaplan	2011	Placebo	[1]				9	478	1.9%	11	478	2.3%
			Tolterodine ER 4mg q.d	[4]				28	973	2.9%	10	973	1.0%
			Fesoterodine ER 4mg q.d	[25]				45	960	4.7%	4	960	0.4%
A81	Kaplan	2014	Placebo	[1]	30	301	10.0%	10	301	3.3%	4	301	1.3%
			Fesoterodine ER 8mg q.d	[26]	68	308	22.1%	10	308	3.2%	2	308	0.6%
A82	Karram	2009	Solifenacin ER 5 - 10mg q.d	[31]	160	357	44.8%	23	357	6.5%			
			Placebo	[1]	88	350	25.1%	16	350	4.6%			
A84	Khullar	2004	Placebo	[1]	96	285	33.7%	16	285	5.6%	2	285	0.7%
			Tolterodine ER 4mg q.d	[4]	221	569	38.8%	26	569	4.6%	3	569	0.5%
A85	Khullar	2013	Placebo	[1]	214	494	43.3%	13	494	2.6%	5	494	1.0%
			Mirabegron ER 50mg q.d	[51]	211	493	42.8%	25	493	5.1%	6	493	1.2%
			Mirabegron ER 100mg q.d	[52]	199	496	40.1%	16	496	3.2%	2	496	0.4%
			Tolterodine ER 4mg q.d	[4]	231	495	46.7%	24	495	4.8%	3	495	0.6%

TABLE B.3: Safety and tolerability data (cont.)

A86	Kosilov	2014	Trospium 60mg/day + Solifenacin 20mg/day (cyclic)	[139]	7	58	12.1%	1	58	1.7%	0	58	0.0%
			Trospium 30mg/day + Solifenacin 10mg/day (cyclic)	[123]	7	55	12.7%	0	55	0.0%	0	55	0.0%
			Trospium 30mg/day + Solifenacin 10mg/day (continuous)	[126]	12	62	19.4%	3	62	4.8%	0	62	0.0%
			Placebo	[1]				0	64	0.0%	0	64	0.0%
A88	Kuo	2014	OnaBoNT-A 200u trigone sparing	[73]	0	12	0.0%						
			Placebo	[1]	0	12	0.0%						
A89	Kuo	2014	Placebo	[1]	124	366	33.9%	14	366	3.8%	7	366	1.9%
			Mirabegron ER 50mg q.d	[51]	105	366	28.7%	9	366	2.5%	4	366	1.1%
			Tolterodine ER 4mg q.d	[4]	128	371	34.5%	15	371	4.0%	2	371	0.5%
A90	Kuo	2015	Placebo	[1]	33	81	42.9%	2	81	2.5%	1	81	1.2%
			Mirabegron ER 50mg q.d	[51]	36	85	42.4%	2	85	2.4%	3	85	3.5%
			Tolterodine ER 4mg q.d	[4]	40	82	49.4%	3	82	3.7%	0	82	0.0%
A93	Lee	2002	Tolterodine IR 2mg b.i.d	[5]	62	112	55.4%	11	112	10.0%			
			Oxybutynin IR 5mg b.i.d	[18]	94	116	81.0%	19	116	16.0%			
A94	Lee	2010	Propiverine ER 20mg q.d	[41]	32	176	18.2%	6	176	3.4%	4	176	2.3%
			Placebo	[1]	10	88	11.4%	1	88	1.1%	2	88	2.3%
A95	Lee	2013	Imidafenacin 0.1mg b.i.d	[36]	72	104	69.2%						
			Fesoterodine ER 4mg q.d	[25]	68	102	66.7%						
A96	Lehtoranta	2002	Oxybutynin intravesically 5mg t.i.d	[14]	7	9	77.8%						
			Placebo	[1]	5	9	55.6%						
A97	Madersbacher	1999	Propiverine IR 15mg t.i.d	[116]	95	149	64.0%						
			Oxybutynin IR 5mg b.i.d	[18]	104	145	72.0%						
			Placebo	[1]	30	72	42.0%						
A99	Malone-Lee	2001	Placebo	[1]	27	43	63.0%	1	43	2.0%			
			Tolterodine IR 1mg b.i.d	[6]	43	61	70.0%	4	61	7.0%			
			Tolterodine IR 2mg b.i.d	[5]	53	73	73.0%	7	73	10.0%			
A100	Malone-Lee	2009	Tolterodine IR 2mg b.i.d	[5]	132	190	69.5%	22	190	11.6%	1	190	0.5%
			Oxybutynin IR 2.5 - 5mg b.i.d	[24]	153	188	81.4%	28	188	14.9%	4	188	2.1%
A101	Malone-Lee	2009	Placebo	[1]	67	142	47.0%	2	142	1.4%	4	142	2.8%
			Tolterodine ER 4mg q.d	[4]	88	165	53.0%	7	165	4.2%	1	165	0.6%
A102	Marencak	2011	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	33	102	32.4%	3	102	2.9%	0	102	0.0%
			Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	27	105	25.7%	2	105	1.9%	1	105	1.0%
			Pregabalin 150mg b.i.d	[62]	34	105	32.4%	5	105	4.8%	0	105	0.0%
			Tolterodine ER 4mg q.d	[4]	29	104	27.9%	0	104	0.0%	2	104	1.9%
			Placebo	[1]	29	103	28.2%	1	103	1.0%	0	103	0.0%
A103	Martinez-Garcia	2009	Placebo	[1]	10	27	37.0%						
			Cizilirtine citrate 200mg b.i.d	[56]	17	25	68.0%						
			Cizilirtine citrate 400mg b.i.d	[57]	22	27	81.0%						
A104	Mattiasson	2003	Tolterodine IR 2mg b.i.d + BT	[93]	158	244	65.0%						
			Tolterodine IR 2mg b.i.d	[5]	177	257	69.0%						
A105	Mattiasson	2010	Solifenacin ER 5 - 10mg q.d	[31]	150	298	46.4%	20	298	6.7%	2	298	0.7%
			Solifenacin ER 5 - 10mg q.d + BT	[77]	149	304	46.6%	15	304	5.0%	1	304	0.3%
A106	Mazur	1995	Propiverine IR 15mg b.i.d	[43]	11	46	23.9%						
			Propiverine IR 30mg b.i.d	[117]	21	47	44.7%						
			Propiverine IR 45mg t.i.d	[118]	27	49	55.1%						
			Propiverine ER 60mg q.d	[119]	24	43	55.8%						
A108	Millard	1999	Placebo	[1]	50	64	78.1%	0	64	0.0%			
			Tolterodine IR 1mg b.i.d	[6]	90	123	73.2%	2	123	1.6%			
			Tolterodine IR 2mg b.i.d	[5]	93	129	72.1%	8	129	6.2%			
A109	Millard	2004	Tolterodine IR 2mg b.i.d	[5]				23	253	9.1%	3	253	1.2%
			Tolterodine IR 2mg b.i.d + PFMT	[95]				22	227	9.7%	6	227	2.6%
A110	Nitti	2007	Placebo	[1]	149	271	55.0%	11	271	4.0%			
			Fesoterodine ER 4mg q.d	[25]	171	282	60.6%	17	282	6.0%			
			Fesoterodine ER 8mg q.d	[26]	193	279	69.2%	25	279	9.0%			
A111	Nitti	2010	Placebo	[1]	33	43	76.7%	2	43	4.7%			
			Fesoterodine IR 4mg b.i.d	[120]	36	43	83.7%	1	43	2.3%			
			Fesoterodine IR 8mg b.id	[121]	39	47	83.0%	2	47	4.3%			
			Fesoterodine IR 12mg b.i.d	[122]	33	38	86.8%	5	38	13.2%			
A112	Nitti	2013	Placebo	[1]				17	453	3.8%	9	453	2.0%
			Mirabegron ER 50mg q.d	[51]				18	442	4.0%	1	442	0.2%
			Mirabegron ER 100mg q.d	[52]				19	433	4.2%	5	433	1.2%
A113	Nitti	2013	Placebo	[1]				2	272	0.7%	0	272	0.0%
			OnaBoNT-A 100u trigone sparing	[72]				4	278	1.4%	1	278	0.4%
A114	Norton	1994	Terodiline 25mg b.i.d	[28]	36	46	76.6%	1	46	2.1%			
			Placebo	[1]	25	46	54.3%	0	46	0.0%			
A116	Ohlstein	2012	Placebo	[1]	35	85	41.2%	1	85	1.2%			
			Solabegron IR 50mg b.i.d	[54]	34	88	38.6%	10	88	11.4%			
			Solabegron IR 125mg b.i.d	[55]	32	85	37.6%	5	85	5.9%			
A117	Olmo	2013	Electrostimulation	[80]	0	11	0.0%						
			Percutaneous tibial nerve stimulation	[83]	0	11	0.0%						
A119	Orri	2014	Placebo	[1]	3	6	50.0%	0	6	0.0%			
			Tolterodine ER 4mg q.d	[4]	6	12	50.0%	1	12	8.3%			
A121	Ozdedeli	2010	Trospium chloride IR 15mg t.i.d	[46]				1	15	6.7%			
			Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]				0	15	0.0%			
			Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]				0	16	0.0%			
A122	Park	2014	Imidafenacin 0.1mg b.i.d	[36]	44	81	54.3%	4	81	4.9%			
			Propiverine ER 20mg q.d	[41]	49	79	62.0%	5	79	6.3%			
A123	Peters	2009	Percutaneous tibial nerve stimulation	[83]	8	49	16.3%						
			Tolterodine ER 4mg q.d	[4]	7	49	14.3%						
A124	Peters	2010	Sham therapy	[3]	0	110	0.0%						
			Percutaneous tibial nerve stimulation	[83]	6	110	5.5%						
A125	Preik	2004	Oxybutynin ER 5 - 30mg/day	[22]				5	53	9.4%			
			Oxybutynin IR 5 - 20mg	[23]				5	52	9.6%			
A126	Preyer	2015	Percutaneous tibial nerve stimulation	[83]	3	18	17.0%						
			Tolterodine IR 2mg b.i.d	[5]	9	18	50.0%						

Table B.3: Safety and tolerability data (cont.)

A127	Rentzhog	1998	Placebo	[1]	6	13	46.2%						
			Tolterodine IR 0.5mg b.i.d	[124]	8	21	38.1%						
			Tolterodine IR 1mg b.i.d	[6]	6	16	37.5%						
			Tolterodine IR 2mg b.i.d	[5]	7	14	50.0%						
			Tolterodine IR 4mg b.i.d	[140]	12	16	75.0%						
A129	Rogers	2008	Placebo	[1]	111	210	53.0%	6	210	2.9%	1	210	0.5%
			Tolterodine ER 4mg q.d	[4]	114	201	57.0%	9	201	4.5%	0	201	0.0%
A130	Rudy	2006	Placebo	[1]	153	329	46.5%	15	329	4.6%			
			Trospium chloride IR 20mg b.i.d	[45]	196	329	59.6%	24	329	7.3%			
A133	Sancaktar	2010	Tolterodine ER 4mg q.d	[4]				0	18	0.0%			
			Tolerodine ER 4mg q.d + Neurostimulation	[96]				0	20	0.0%			
A135	Schreiner	2010	Electrostimulation + PFMT + BT	[97]	0	25	0.0%						
			PFMT + BT	[89]	0	26	0.0%						
A136	Song	2006	Bladder Training (BT)/Behaviour Therapy	[85]	0	26	0.0%	0	26	0.0%			
			Tolterodine IR 2mg b.i.d	[5]	13	32	40.6%	2	32	6.3%			
			Tolterodine IR 2mg b.i.d + BT	[93]	12	31	38.7%	2	31	6.5%			
A137	Song	2015	Tarafenacin 0.2mg q.d	[90]	42	79	53.2%						
			Tarafenacin 0.4mg q.d	[82]	61	81	75.3%						
			Placebo	[1]	34	75	45.3%						
A139	Staskin	2007	Placebo	[1]	53	303	17.5%	11	303	3.6%			
			Trospium chloride ER 60mg q.d	[44]	80	298	26.8%	12	298	4.0%			
A140	Staskin	2009	Oxybutynin chloride topical gel 1g/day	[13]	73	389	18.8%	19	389	4.9%			
			Placebo	[1]	45	400	11.3%	13	400	3.3%			
A141	Steers	2005	Darifenacin ER 7.5mg q.d	[39]	110	269	40.9%	18	269	6.7%	2	269	0.7%
			Placebo	[1]	26	129	20.2%	4	129	3.1%	1	129	0.8%
A142	Steers	2007	Placebo	[1]	85	153	55.6%	8	153	5.2%	11	153	7.2%
			Duloxetine 60mg b.i.d	[66]	121	153	79.1%	43	153	28.1%	8	153	5.2%
A143	Subak	2002	Bladder Training (BT)/Behaviour Therapy	[85]	0	77	0.0%						
			Control	[2]	0	75	0.0%						
A144	Swift	2003	Tolterodine ER 4mg q.d	[4]				22	417	5.3%			
			Tolterodine IR 2mg b.i.d	[5]				20	408	4.9%			
			Placebo	[1]				26	410	6.3%			
A146	Tapp	1989	Placebo	[1]				4	36	11.1%			
			Terodiline 25mg b.i.d	[28]				5	34	14.7%			
A148	Tincello	2012	OnaBoNT-A 200U trigone sparing	[73]	21	122	17.0%						
			Placebo	[1]	22	118	19.0%						
A150	Ulshofer	2001	Placebo	[1]	11	17	64.7%						
			Trospium chloride IR 15mg t.i.d	[46]	14	22	63.6%						
A151	Van Kerrebroeck	2001	Tolterodine ER 4mg q.d	[4]				27	507	5.3%			
			Tolterodine IR 2mg b.i.d	[5]				28	514	5.4%			
			Placebo	[1]				33	508	6.5%			
A153	Vardy	2009	Solifenacin ER 5mg - 15mg q.d	[34]	100	385	26.0%	12	385	3.0%			
			Placebo	[1]	50	381	13.0%	15	381	4.0%			
A154	Versi	2000	Oxybutynin ER 5 - 30mg/day	[22]				3	111	2.7%	1	111	0.9%
			Oxybutynin IR 5 - 20mg	[23]				7	115	6.1%	1	115	0.9%
A155	Visco	2012	Solifenacin/trospium + placebo injection	[100]	88	127	69.3%						
			OnaBoNT-A 100U trigone sparing	[72]	88	120	73.3%						
A156	Wagg	2013	Placebo	[1]	142	393	36.1%	22	393	5.6%	8	393	2.0%
			Fesoterodine ER 4mg q.d	[25]	244	392	62.2%	46	392	11.7%	12	392	3.1%
A159	Weiss	2013	Placebo	[1]	152	474	32.1%	11	474	2.3%	11	474	2.3%
			Fesoterodine ER 4 - 8mg q.d	[27]	188	463	40.6%	25	463	5.4%	6	463	1.3%
A160	Yamaguchi	2007	Placebo	[1]				12	405	3.0%			
			Solifenacin ER 5mg q.d	[29]				22	398	5.5%			
			Solifenacin ER 10mg q.d	[30]				26	381	6.8%			
			Propiverine ER 20mg q.d	[41]				26	400	6.5%			
A161	Yamaguchi	2011	Placebo	[1]	81	318	25.5%	12	318	3.8%	5	318	1.6%
			Fesoterodine ER 4mg q.d	[25]	150	320	46.9%	15	320	4.7%	3	320	0.9%
			Fesoterodine ER 8mg q.d	[26]	192	313	61.3%	15	313	4.8%	2	313	0.6%
A162	Yamaguchi	2014	Placebo	[1]	91	381	23.9%	9	381	2.4%	3	381	0.8%
			Mirabegron ER 50mg q.d	[51]	93	380	24.5%	15	380	3.9%	4	380	1.1%
			Tolterodine ER 4mg q.d	[4]	131	378	34.7%	13	378	3.4%	2	378	0.5%
A163	Yamaguchi	2014	Oxbutynin patch 73.5mg	[15]	428	572	74.8%	38	572	6.6%			
			Propiverine ER 20mg q.d	[41]	374	576	64.9%	9	576	1.6%			
			Placebo	[1]	215	381	56.4%	6	381	1.6%			
A164	Yamaguchi	2014	Placebo	[1]	157	212	74.1%	6	212	2.8%	4	212	1.9%
			Mirabegron ER 25mg q.d	[50]	169	210	80.5%	6	210	2.9%	0	210	0.0%
			Mirabgeron ER 50mg q.d	[51]	171	208	82.2%	8	208	3.8%	0	208	0.0%
			Mirabegron ER 100mg q.d	[52]	175	208	84.1%	8	208	3.9%	0	208	0.0%
A165	Yamanishi	2000	Electrostimulation	[80]	2	37	5.4%						
			Sham therapy	[3]	2	31	6.5%						
A166	Yokoyama	2013	Imidafenacin 0.1mg b.i.d	[36]				3	55	5.5%	20	55	36.4%
			Solifenacin ER 5mg q.d	[29]				7	54	13.0%	9	54	16.7%
A169	Zat'ura	2010	Cizolirtine citrate 400mg b.i.d	[57]	12	54	22.2%	8	54	14.8%	1	54	2.0%
			Placebo	[1]	16	54	29.6%	0	54	0.0%	0	54	0.0%
			Oxybutynin IR 5mg t.i.d	[7]	3	27	9.3%	2	27	7.4%	0	27	0.0%
A170	Zellner	2009	Trospium chloride IR 45mg t.i.d	[47]	188	828	22.7%	48	828	5.8%			
			Oxybutynin IR 2.5 - 5mg b.i.d	[24]	220	830	26.5%	68	830	8.2%			
A172	Zinner	2002	Tolterodine ER 4mg q.d	[4]	264	506	52.1%	27	506	5.3%	6	506	1.2%
			Placebo	[1]	246	507	48.5%	33	507	6.5%	9	507	1.8%
A173	Zinner	2005	Darifenacin ER 15mg q.d.	[40]				0	61	0.0%			
			Darifenacin ER 30mg q.d	[38]				1	61	1.6%			
			Oxybutynin IR 5mg t.i.d	[7]				3	61	4.9%			
			Placebo	[1]				0	61	0.0%			
A174	Zinner	2006	Darifenacin ER 15mg q.d	[40]	136	214	63.6%	17	214	7.9%	2	214	0.9%
			Placebo	[1]	110	225	48.9%	10	225	4.4%	5	225	2.2%

Appendix C

Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder

C.1 Example WinBUGS code for fixed effect NMA incorporating a missing data framework for continuous outcomes

```
Model{
#MISSING DATA FRAMEWORK
for(i in 1:ns){
        for (k in 1:na[i]){

                change_var[i,k] <- ((pow(b_sd_star[i,k],2) + pow(f_sd_star[i,k],2) -
2*rho[i,k]*b_sd_star[i,k]*f_sd_star[i,k])*equals(ind_c_miss[i,k],1)) +
(pow(sd[i,k],2)*(equals(ind_c_miss[i,k],0)))

                change_sd[i,k] <- sqrt(change_var[i,k])
                se[i,k] <- change_sd[i,k]/sqrt(numinclanalysis[i,k])


f_sd_star[i,k] <- (-0.011+(0.835*b_sd[i,k]))*(equals(ind_f_miss[i,k],1)) +
(f_sd[i,k])*(equals(ind_f_miss[i,k],0))


        b_sd_star[i,k] <- 0*equals(ind_b_miss[i,k],1) +
(b_sd[i,k])*(equals(ind_b_miss[i,k],0))


b_sd[i,k] ~ dunif(0,15)
f_sd[i,k] ~ dunif(0, 15)
sd[i,k] ~ dunif(0,25)

z[i,k]~ dnorm(z.star[i,k],z.prec)
z.star[i,k] ~ dnorm(0.67, 12.76)
rho[i,k]<- (exp(2*z[i,k])-1)/(exp(2*z[i,k])+1)

}
}

z.se <- 1/(sqrt(49-3))
z.prec<- pow(z.se,-2)


#FIXED EFFECT MODEL
for(i in 1:ns){                                                  # loop through studies
mu[i] ~ dnorm(0,.001)                          # vague priors for all trial baselines
        for (k in 1:na[i]) {
                                                                 # loop through arms/datapoints
var[i,k] <- pow(se[i,k])                                          # calculate variances
prec[i,k] <- 1/var[i,k]                                          # calculate precisions
y[i,k] ~ dnorm(theta[i,k],prec[i,k])                             # normal likelihood
theta[i,k] <- mu[i] + d[t[i,k]] - d[t[i,1]]                  # model for linear predictor

dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
}
```

```
resdev[i] <- sum(dev[i,1:na[i]])
}
totresdev <- sum(resdev[])                              #Total Residual Deviance

d[1]<-0                                                 # Reference treatment set to zero

                                                        # vague priors for treatment effects
for (k in 2:nt){
        d[k] ~ dnorm(0,.001) }


# All pairwise comparisons
for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[c,k] <- (d[k] - d[c] )}}

for (k in 1:nt) {
        rk[k] <- rank(d[],k)                            # Relative rankings
        best[k] <- equals(rk[k],1)}             #calculate probability that treat is best
}

}
```

C.2 Example WinBUGS code for random effects NMA incorporating a missing data framework for continuous outcomes

```
Model{

#MISSING DATA FRAMEWORK
for(i in 1:ns){
        for (k in 1:na[i]){

                change_var[i,k] <- ((pow(b_sd_star[i,k],2) + pow(f_sd_star[i,k],2) -
2*rho[i,k]*b_sd_star[i,k]*f_sd_star[i,k])*equals(ind_c_miss[i,k],1)) +
(pow(sd[i,k],2)*(equals(ind_c_miss[i,k],0)))

                change_sd[i,k] <- sqrt(change_var[i,k])
                se[i,k] <- change_sd[i,k]/sqrt(numinclanalysis[i,k])


f_sd_star[i,k] <- (-0.011+(0.835*b_sd[i,k]))*(equals(ind_f_miss[i,k],1)) +
(f_sd[i,k])*(equals(ind_f_miss[i,k],0))


        b_sd_star[i,k] <- 0*equals(ind_b_miss[i,k],1) +
(b_sd[i,k])*(equals(ind_b_miss[i,k],0))


b_sd[i,k] ~ dunif(0,15)
f_sd[i,k] ~ dunif(0, 15)
sd[i,k] ~ dunif(0,25)

z[i,k]~ dnorm(z.star[i,k],z.prec)
z.star[i,k] ~ dnorm(0.67, 12.76)
rho[i,k]<- (exp(2*z[i,k])-1)/(exp(2*z[i,k])+1)

}
}

z.se <- 1/(sqrt(49-3))
z.prec<- pow(z.se,-2)

#RANDOM EFFECTS MODEL
for(i in 1:ns){

                w[i,1] <- 0
                delta[i,1] <- 0
                mu[i] ~ dnorm(0,0.001)

        for (k in 1:na[i]) {

var[i,k] <- pow(se[i,k],2)
prec[i,k] <- 1/var[i,k]
y[i,k] ~ dnorm(theta[i,k],prec[i,k])
theta[i,k] <- mu[i] + delta[i,k]
```

```
dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
}

resdev[i] <- sum(dev[i,1:na[i]])

        for (k in 2:na[i]) {


                delta[i,k] ~ dnorm(md[i,k],taud[i,k])
                md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]
                taud[i,k] <- tau *2*(k-1)/k
                w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
                sw[i,k] <- sum(w[i,1:k-1])/(k-1)
                }
}

totresdev <- sum(resdev[])


# Treatment effects
d[1]<-0                         #Reference treatment set to zero

for(i in 2:nt){d[i] ~ dnorm(0, 0.001)}

between.study.sd ~ dunif(0,5)
tau <- pow(between.study.sd,-2)

#tau ~ dgamma(0.001,0.001)              #Sensitivity to variance parameters
#between.study.sd <- 1/sqrt(tau)

#between.study.sd ~ dnorm(0,1)I(0,)     #Sensitivity to variance parameters
#tau <- pow(between.study.sd,-2)


for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[c,k] <- (d[k] - d[c] )}}
for (k in 1:nt) {
        rk[k] <- rank(d[],k)

        best[k] <- equals(rk[k],1)}

}

}
```

C.3 Example WinBUGS code for age adjusted random effects NMA incorporating a missing data framework for continuous outcomes

```
Model{

#MISSING DATA FRAMEWORK
for(i in 1:ns){
        for (k in 1:na[i]){

                change_var[i,k] <- ((pow(b_sd_star[i,k],2) + pow(f_sd_star[i,k],2) -
2*rho[i,k]*b_sd_star[i,k]*f_sd_star[i,k])*equals(ind_c_miss[i,k],1)) +
(pow(sd[i,k],2)*(equals(ind_c_miss[i,k],0)))

                change_sd[i,k] <- sqrt(change_var[i,k])
                se[i,k] <- change_sd[i,k]/sqrt(numinclanalysis[i,k])


f_sd_star[i,k] <- (-0.011+(0.835*b_sd[i,k]))*(equals(ind_f_miss[i,k],1)) +
(f_sd[i,k])*(equals(ind_f_miss[i,k],0))


        b_sd_star[i,k] <- 0*equals(ind_b_miss[i,k],1) +
(b_sd[i,k])*(equals(ind_b_miss[i,k],0))


b_sd[i,k] ~ dunif(0,15)
f_sd[i,k] ~ dunif(0, 15)
sd[i,k] ~ dunif(0,25)

z[i,k]~ dnorm(z.star[i,k],z.prec)
z.star[i,k] ~ dnorm(0.67, 12.76)
rho[i,k]<- (exp(2*z[i,k])-1)/(exp(2*z[i,k])+1)
}
}

z.se <- 1/(sqrt(49-3))
z.prec<- pow(z.se,-2)

#AGE ADJUSTED RANDOM EFFECTS

for(i in 1:ns){
        w[i,1] <- 0
        delta[i,1] <- 0
        mu[i] ~ dnorm(0,0.001)
                for (k in 1:na[i]) {

var[i,k] <- pow(se[i,k],2)
prec[i,k] <- 1/var[i,k]
y[i,k] ~ dnorm(theta[i,k],prec[i,k])
```

```
theta[i,k] <- mu[i] + delta[i,k] +beta.age*(age[i]-57.5)
dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
}

resdev[i] <- sum(dev[i,1:na[i]])

        for (k in 2:na[i]) {
                delta[i,k] ~ dnorm(md[i,k],taud[i,k])
                md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]
                taud[i,k] <- tau *2*(k-1)/k
                w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
                sw[i,k] <- sum(w[i,1:k-1])/(k-1)
                }
}

totresdev <- sum(resdev[])

d[1]<-0                                 # Reference treatment set to zero

for(i in 2:nt){d[i] ~ dnorm(0, 0.001)}

between.study.sd ~ dunif(0,5)
tau <- pow(between.study.sd,-2)

beta.age ~ dnorm(0, 0.001)

for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[c,k] <- (d[k] - d[c] )}}
for (k in 1:nt) {
        rk[k] <- rank(d[],k)

        best[k] <- equals(rk[k],1)}

}

}
```

C.4 Example WinBUGS code for baseline risk NMA incorporating a missing data framework for continuous outcomes

```
Model{

for(i in 1:ns){
        for (k in 1:na[i]){

change_var[i,k] <- ((pow(b_sd_star[i,k],2) + pow(f_sd_star[i,k],2) -
2*rho[i,k]*b_sd_star[i,k]*f_sd_star[i,k])*equals(ind_c_miss[i,k],1)) +
(pow(sd[i,k],2)*(equals(ind_c_miss[i,k],0)))

change_sd[i,k] <- sqrt(change_var[i,k])
se[i,k] <- change_sd[i,k]/sqrt(numinclanalysis[i,k])

f_sd_star[i,k] <- (-0.011+(0.835*b_sd[i,k]))*(equals(ind_f_miss[i,k],1)) +
(f_sd[i,k])*(equals(ind_f_miss[i,k],0))

b_sd_star[i,k] <- 0*equals(ind_b_miss[i,k],1) +
(b_sd[i,k])*(equals(ind_b_miss[i,k],0))

b_sd[i,k] ~ dunif(0,15)
f_sd[i,k] ~ dunif(0, 15)
sd[i,k] ~ dunif(0,25)

z[i,k]~ dnorm(z.star[i,k],z.prec)
z.star[i,k] ~ dnorm(0.67, 12.76)
rho[i,k]<- (exp(2*z[i,k])-1)/(exp(2*z[i,k])+1)
}
}

z.se <- 1/(sqrt(49-3))
z.prec<- pow(z.se,-2)

for(i in 1:ns){
                w[i,1] <- 0
                delta[i,1] <- 0
                bl[i,1]<-0
                mu[i] ~ dnorm(0,0.001)

        for (k in 1:na[i]) {
var[i,k] <- pow(se[i,k],2)
prec[i,k] <- 1/var[i,k]
y[i,k] ~ dnorm(theta[i,k],prec[i,k])
theta[i,k] <- mu[i] + delta[i,k]
dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
}

resdev[i] <- sum(dev[i,1:na[i]])

        for (k in 2:na[i]) {
```

```
                delta[i,k] ~ dnorm(md[i,k],taud[i,k])
                md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k] + bl[i,k]
                taud[i,k] <- tau *2*(k-1)/k
                w[i,k] <- (delta[i,k] - ((d[t[i,k]] - d[t[i,1]]) +bl[i,k]))
                sw[i,k] <- sum(w[i,1:k-1])/(k-1)

                bl[i,k]<-(beta.baseline[t[i,k]]-beta.baseline[t[i,1]])*(mu[i])
                }
}

totresdev <- sum(resdev[])


# Treatment effects
d[1]<-0

for(i in 2:nt){d[i] ~ dnorm(0, 0.001)}

between.study.sd ~ dunif(0,5)

tau <- pow(between.study.sd,-2)
beta.baseline[1]<-0
for(k in 2:nt){

beta.baseline[k] ~ dnorm(0, 0.001)}

for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[c,k] <- (d[k] - d[c] )}}
for (k in 1:nt) {
        rk[k] <- rank(d[],k)

        best[k] <- equals(rk[k],1)}
}

}
```

C.5 Example WinBUGS code for fixed effect NMA of binary outcomes

```
Model{
# FIXED EFFECT MODEL
for(i in 1:ns){                                  # loop through studies
   mu[i] ~ dnorm(0,.01)                          # vague priors for trial baselines
   for (k in 1:na[i])  {                         # loop through arms/datapoints
      r[i,k] ~ dbin(p[i,k],n[i,k])               # binomial likelihood
      logit(p[i,k]) <- mu[i] + d[t[i,k]] - d[t[i,1]]
      rhat[i,k] <- p[i,k] * n[i,k]

      dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
         +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
    }

   resdev[i] <- sum(dev[i,1:na[i]])
    }
totresdev <- sum(resdev[])                       # Total Residual Deviance
d[1]<-0                                          # Reference treatment set to zero

                                                 # vague priors for treatment effects
for (k in 2:nt){
         d[k] ~ dnorm(0,.01) }


#Rank the treatment effects (with 1=worst) & record the worst treatment
for(k in 1:nt){
                  rkworst[k]<-nt+1-rank(d[],k)
                  worst[k]<-equals(rkworst[k],1)
                  rkbest[k]<-rank(d[],k)
                  best[k]<-equals(rkbest[k],1)}

#All pairwise log odds ratios and odds ratios
for(c in 1:(nt-1)){
         for(k in (c+1):nt){
                  #or[c,k]<-exp(lor[c,k])
                  lor[c,k]<-(d[k]-d[c])
         }}
}
```

C.6 Example WinBUGS code for random effects NMA of binary outcomes

```
Model{
# RANDOM EFFECT MODEL
for(i in 1:ns){                                          # loop through studies
  w[i,1] <- 0                  # adjustment for multi-arm trials is zero for control arm
  delta[i,t[i,1]] <- 0                        # treatment effect is zero for control arm
  mu[i] ~  dnorm(0.0,0.01)                          # vague priors for trial baselines
  for (k in 1:na[i]) {                              # loop through arms/datapoints
    r[i,k] ~ dbin(p[i,t[i,k]],n[i,k])                    # binomial likelihood
    logit(p[i,t[i,k]]) <- mu[i] + delta[i,t[i,k]]
    rhat[i,k] <- p[i,t[i,k]] * n[i,k]

    dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
      +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))}

  resdev[i] <- sum(dev[i,1:na[i]])

 for (k in 2:na[i]) {                                          # loop through arms
       delta[i,t[i,k]] ~ dnorm(md[i,t[i,k]],taud[i,t[i,k]])
       md[i,t[i,k]] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
       taud[i,t[i,k]] <- tau *2*(k-1)/k
       w[i,k] <- (delta[i,t[i,k]] - d[t[i,k]] + d[t[i,1]])
       sw[i,k] <- sum(w[i,1:k-1])/(k-1)
   }
 }
totresdev <- sum(resdev[])                               # Total Residual Deviance

d[1]<-0                                          # Reference treatment set to zero

for(i in 2:nt){
d[i] ~ dnorm(0,0.01) }

sd ~ dunif(0,2)                                  # vague prior for between-trial SD
tau <- pow(sd,-2)                                     # between-trial precision

#Rank the treatment effects (with 1=worst) & record the worst treatment
for(k in 1:nt){
                exp.d[k] <- exp(d[k])
                rkworst[k]<-nt+1-rank(d[],k)
                worst[k]<-equals(rkworst[k],1)}

for(c in 1:(nt-1)){
       for(k in (c+1):nt){
                or[c,k]<-exp(d[k]-d[c])
                lor[c,k]<-(d[k]-d[c])
       }}
}
```

C.7 Results for number of patients experiencing adverse events

TABLE C.1: Estimated odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from random effects network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Odds Ratio† (95%CrI)	Rank worst (95%CrI)	p(Worst)
	Propiverine ER 60mg q.d	[119]	1	43	8.06 (2.64,25.57)	3 (1,19)	0.20
	Propiverine IR 45mg t.i.d	[118]	1	49	7.8 (2.61,24)	3 (1,19)	0.17
	Darifenacin ER 30mg q.d	[38]	2	151	7.08 (3.92,13.24)	4 (1,11)	0.11
	Imidafenacin 0.25mg b.i.d	[37]	1	76	5.29 (2.6,10.88)	6 (1,21)	0.04
	Propiverine IR 30mg b.i.d	[117]	1	47	5.1 (1.67,15.88)	7 (2,39)	0.02
	Tolterodine IR 4mg b.i.d	[140]	1	16	4.65 (1.26,20.6)	8 (1,53)	0.09
	Solifenacin ER 20mg q.d	[33]	1	37	4.34 (1.86,10.32)	9 (2,34)	0.02
	Oxybutynin IR 5mg b.i.d	[18]	4	415	3.98 (2.58,6.22)	10 (4,21)	0.00
	Tarafenacin 0.4mg q.d	[82]	1	81	3.74 (1.65,8.6)	11 (2,40)	0.01
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	3.07 (0.35,34.97)	15 (1,77)	0.11
	Terodiline 25mg b.i.d	[28]	1	46	3.09 (1.13,8.8)	15 (2,58)	0.01
	Duloxetine 60mg b.i.d	[66]	1	153	3.06 (1.55,6.09)	15 (5,44)	0.00
	Oxbutynin patch 73.5mg	[15]	1	572	2.88 (1.77,4.71)	17 (7,37)	0.00
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	2	1018	2.78 (1.34,5.84)	18 (5,49)	0.00
	Propiverine IR 15mg t.i.d	[116]	1	149	2.64 (1.39,5.08)	20 (7,49)	0.00
	Oxybutynin IR 5mg t.i.d	[7]	3	257	2.63 (1.5,4.65)	20 (7,45)	0.00
	Fesoterodine ER 8mg q.d	[26]	4	1187	2.63 (1.99,3.45)	20 (11,33)	0.00
	Cizolirtine citrate 400mg b.i.d	[57]	2	81	2.49 (1.19,5.35)	22 (6,56)	0.00
	Solifenacin ER 5 - 15mg q.d	[34]	1	385	2.33 (1.29,4.24)	24 (8,53)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	25	2.28 (0.73,7.57)	25 (3,72)	0.01
	Solifenacin ER 5 - 10mg q.d	[31]	3	1160	2.28 (1.38,3.76)	25 (11,49)	0.00
	Darifenacin ER 15mg q.d.	[40]	2	321	2.31 (1.49,3.63)	25 (11,46)	0.00
	Trospium chloride IR 45mg t.i.d	[47]	1	828	2.27 (0.93,5.58)	26 (6,66)	0.00
	Propiverine ER 20mg q.d	[41]	6	1458	2.22 (1.66,3.01)	26 (15,41)	0.00
	OnaBoNTA 150u trigone sparing	[75]	1	50	2.14 (0.88,5.22)	28 (6,67)	0.00
	OnaBoNTA 300u trigone sparing	[76]	1	55	2.13 (0.91,5.07)	28 (7,66)	0.00
	Solifenacin ER 5 - 10mg q.d + BT	[77]	1	304	2.16 (1.01,4.59)	28 (8,63)	0.00
	Fesoterodine IR 12mg b.i.d	[122]	1	38	2.11 (0.59,8.39)	29 (3,74)	0.01
	Darifenacin ER 7.5mg q.d	[39]	2	377	2.06 (1.3,3.29)	30 (13,52)	0.00
	OnaBoNT-A 100u trigone sparing	[72]	3	186	2.01 (0.91,4.55)	31 (8,66)	0.00
	Oxybutynin vaginal ring 6mg q.d	[17]	1	147	2.02 (1.05,3.87)	31 (10,62)	0.00
	Propiverine IR 15mg b.i.d	[43]	3	541	1.96 (1.21,3.16)	32 (14,55)	0.00
	Trospium chloride IR 20mg b.i.d	[45]	2	596	1.97 (1.24,3.18)	32 (14,54)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	1.83 (0.99,3.39)	35 (12,64)	0.00
	Fesoterodine ER 4mg q.d	[25]	5	1368	1.81 (1.4,2.33)	36 (23,49)	0.00
	Propiverine ER 30mg q.d	[42]	1	391	1.78 (1.01,3.15)	36 (14,63)	0.00
	Oxybutynin vaginal ring 4mg q.d	[16]	1	143	1.76 (0.92,3.39)	37 (13,67)	0.00
	Solifenacin/trospium + placebo injection	[100]	1	127	1.64 (0.56,4.92)	41 (7,75)	0.00
	ZD0947IL 25mg/day	[58]	1	90	1.61 (0.75,3.49)	41 (12,71)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	4	602	1.64 (1.14,2.37)	41 (23,59)	0.00
	Fesoterodine IR 4mg b.i.d	[120]	1	43	1.61 (0.49,5.59)	42 (6,75)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1047	1.6 (1.16,2.2)	42 (26,58)	0.00
	Solifenacin ER 10mg q.d	[30]	2	375	1.6 (0.99,2.58)	42 (20,64)	0.00
	Trospium chloride ER 60mg q.d	[44]	2	578	1.58 (1.04,2.39)	43 (23,63)	0.00
	Fesoterodine IR 8mg b.id	[121]	1	47	1.51 (0.47,4.87)	45 (7,76)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	1.43 (0.7,2.89)	47 (17,72)	0.00
	Tolterodine IR 2mg b.i.d	[5]	13	1293	1.44 (1.06,1.97)	47 (32,61)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	3	186	1.45 (0.79,2.71)	47 (20,70)	0.00
	OnaBoNTA 50u trigone sparing	[74]	1	56	1.39 (0.57,3.36)	49 (14,75)	0.00
	Pregabalin 150mg b.i.d	[62]	1	105	1.36 (0.71,2.6)	49 (20,72)	0.00
	Pregabalin 150mg b.i.d + tolterodine ER 4mg q.d	[102]	1	102	1.37 (0.7,2.61)	49 (20,72)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	79	1.37 (0.63,3.05)	49 (15,74)	0.00
	Mirabegron ER 200mg q.d	[53]	1	167	1.3 (0.74,2.27)	52 (25,72)	0.00
	Netupitant 200mg q.d	[112]	1	61	1.26 (0.46,3.48)	53 (12,76)	0.00

TABLE C.1: Estimated odds ratios (95% credible intervals) for the proportion of patients with adverse events obtained from random effects network meta-analysis (cont.)

	ONO-8539 300mg b.i.d	[61]	1	88	1.25 (0.65,2.41)	53 (23,73)	0.00
	Solifenacin ER 5mg q.d	[29]	6	1133	1.23 (0.85,1.79)	54 (36,69)	0.00
	Tolterodine ER 4mg q.d	[4]	19	4729	1.23 (1.05,1.44)	54 (45,62)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	2	275	1.22 (0.67,2.24)	55 (26,73)	0.00
	Mirabegron ER 100mg q.d	[52]	4	1692	1.14 (0.88,1.5)	58 (44,69)	0.00
	Tolterodine IR 1mg b.i.d	[6]	4	297	1.12 (0.73,1.71)	59 (39,72)	0.00
	Mirabegron ER 25mg q.d	[50]	3	811	1.08 (0.79,1.49)	60 (44,71)	0.00
	Lipo-BoNTA 200U	[138]	1	29	1.08 (0,511)	61 (1,81)	0.17
	Mirabegron IR 150mg b.i.d	[49]	1	65	1.06 (0.47,2.34)	61 (24,76)	0.00
	Mirabegron ER 50mg q.d	[51]	9	3889	1.05 (0.87,1.28)	62 (51,70)	0.00
	Placebo	[1]	67	12078	NA	64 (57,70)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	105	0.99 (0.5,1.92)	64 (33,76)	0.00
	Trospium chloride IR 15mg t.i.d	[46]	1	22	0.95 (0.22,4)	65 (10,78)	0.00
	ONO-8539 30mg b.i.d	[59]	1	88	0.95 (0.49,1.82)	66 (35,76)	0.00
	Netupitant 100mg q.d	[111]	1	61	0.9 (0.31,2.58)	67 (21,78)	0.00
	Solabegron IR 50mg b.i.d	[54]	1	88	0.9 (0.42,1.93)	67 (33,77)	0.00
	Tolterodine IR 0.5mg b.i.d	[141]	1	21	0.88 (0.27,2.82)	68 (18,78)	0.00
	Solabegron IR 125mg b.i.d	[55]	1	85	0.86 (0.39,1.86)	68 (35,77)	0.00
	Netupitant 50mg q.d	[110]	1	62	0.78 (0.26,2.26)	70 (26,78)	0.00
	ONO-8539 100mg b.i.d	[60]	1	87	0.8 (0.41,1.55)	70 (44,77)	0.00
	Percutaneous tibial nerve stimulation	[83]	4	188	0.8 (0.3,2.08)	70 (30,77)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	0.73 (0.31,1.67)	71 (40,78)	0.00
	Solifenacin ER 2.5mg q.d	[32]	1	41	0.54 (0.18,1.43)	75 (47,79)	0.00
	Electrostimulation	[80]	2	48	0.14 (0,2.28)	78 (26,81)	0.00
	Sham therapy	[3]	2	141	0.1 (0,0.83)	79 (70,81)	0.00
	Control	[2]	1	75	0.03 (0,11.37)	80 (2,81)	0.02
	Bladder Training (BT)/Behaviour Therapy	[85]	2	103	0.03 (0,0.27)	80 (78,81)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment has the highest incidence of adverse events overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average.

p(Worst) denotes the probability that the intervention in question has the highest incidence of adverse events. The probability worst is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

C.8 Results for discontinuations due to adverse events

TABLE C.2: Estimated odds ratios (95% credible intervals) for discontinuations due to adverse events obtained from random effects network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Odds ratio† (95% CrI)	Rank worst (95% CrI)	p(Worst)
	Imidafenacin 0.25mg b.i.d	[37]	1	76	15.05 (4.69,61.35)	6 (1,20)	0.06
	Trospium 30mg/day + Solifenacin 10mg/day (continuous)	[126]	1	62	11.5 (0.87,1813)	8 (1,64)	0.20
	Oxybutynin 20mg intravesically q.d	[106]	1	23	11.31 (0.65,1002)	9 (1,73)	0.19
	Solabegron IR 50mg b.i.d	[54]	1	88	10.83 (1.9,134)	9 (1,38)	0.07
	Cizolirtine citrate 400mg b.i.d	[57]	1	54	10.97 (2.62,79.2)	9 (1,31)	0.05
	Oxybutynin trandermal 3.9mg/day	[10]	2	169	10.13 (2.71,53.75)	9 (2,30)	0.02
	Propiverine IR 15mg b.i.d	[43]	1	395	9.14 (1.75,210.9)	10 (1,42)	0.08
	Netupitant 50mg q.d	[110]	1	62	8.46 (0.94,364.7)	11 (1,65)	0.08
	Pregabalin 150mg b.i.d	[62]	1	105	7.08 (1.29,67.94)	13 (2,52)	0.02
	Duloxetine 60mg b.i.d	[66]	1	153	7.2 (3.37,18.09)	13 (4,26)	0.00
	Netupitant 200mg q.d	[112]	1	61	6.71 (0.61,280.8)	14 (1,73)	0.03
	Propiverine ER 30mg q.d	[42]	1	391	6.83 (1.21,151.2)	14 (2,57)	0.02
	Oxbutynin patch 73.5mg	[15]	1	572	6.02 (3.05,12.28)	15 (6,28)	0.00
	Solifenacin ER 20mg q.d	[33]	1	37	5.45 (1.42,17.85)	17 (5,49)	0.00
	Solabegron IR 125mg b.i.d	[55]	1	85	4.61 (0.72,108)	19 (2,69)	0.01
	Netupitant 100mg q.d	[111]	1	61	4.83 (0.46,271.9)	19 (2,76)	0.01
	Darifenacin ER 30mg q.d	[38]	3	212	4.85 (2.17,10.8)	19 (7,35)	0.00
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	102	4.08 (0.51,41.77)	21 (3,76)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	65	3.86 (0.92,21.04)	21 (5,66)	0.00
	Fesoterodine IR 12mg b.i.d	[122]	1	38	3.87 (0.58,31.74)	22 (2,73)	0.01
	Oxybutynin vaginal ring 6mg q.d	[17]	1	147	3.4 (1.12,13.19)	24 (7,59)	0.00
	OnaBoNTA 300u trigone sparing	[76]	1	55	3.12 (0.11,38.78)	26 (2,80)	0.01
	Oxybutynin IR 3mg t.i.d	[19]	1	244	3.2 (1.81,5.83)	26 (14,43)	0.00
	OnaBoNTA 150u trigone sparing	[75]	1	50	3.22 (0.15,55.66)	27 (2,80)	0.01
	Serlopitant 4mg q.d	[109]	1	114	2.99 (0.98,8.81)	27 (11,64)	0.00
	Oxybutynin IR 5mg t.i.d	[7]	5	378	3.03 (1.83,5.17)	27 (16,41)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	2.85 (0.64,14.02)	28 (6,73)	0.00
	Serlopitant 0.25mg q.d	[107]	1	110	2.68 (0.87,7.29)	29 (12,68)	0.00
	OnaBoNTA 50u trigone sparing	[74]	1	56	2.58 (0.06,58.42)	30 (2,80)	0.01
	Oxybutynin IR 5mg b.i.d	[18]	3	270	2.62 (1.47,4.67)	30 (17,48)	0.00
	Trospium 60mg/day + Solifenacin 20mg/day (cyclic)	[139]	1	58	2.6 (0.13,532.3)	31 (2,80)	0.02
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	105	2.45 (0.16,23.33)	33 (6,80)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	2.41 (0.36,11.93)	33 (9,77)	0.00
	Oxybutynin ER 15mg q.d	[9]	1	65	2.39 (0.83,6.97)	33 (13,69)	0.00
	ONO-8539 300mg b.i.d	[61]	1	88	2.24 (0.67,6.92)	35 (12,72)	0.00
	Oxybutynin gel 56mg/day	[135]	1	198	2.27 (1.02,5.26)	35 (15,63)	0.00
	Fesoterodine ER 8mg q.d	[26]	6	2645	2.02 (1.56,2.68)	38 (26,49)	0.00
	Darifenacin ER 15mg q.d.	[40]	3	382	1.99 (1.06,3.77)	38 (22,61)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	4	1485	1.89 (1.35,2.72)	39 (27,54)	0.00
	Oxybutynin gel 84mg/day	[134]	1	211	1.88 (0.76,4.83)	41 (16,71)	0.00
	Solifenacin ER 10mg q.d	[30]	6	1442	1.76 (1.28,2.43)	42 (30,55)	0.00
	Oxybutynin vaginal ring 4mg q.d	[16]	1	143	1.66 (0.45,8.42)	45 (12,76)	0.00
	Trospium chloride ER 60mg q.d	[44]	2	578	1.61 (0.91,3.02)	45 (26,68)	0.00
	Fesoterodine ER 4mg q.d	[25]	6	3016	1.58 (1.21,2.11)	46 (35,58)	0.00
	Trospium chloride IR 20mg b.i.d	[45]	2	596	1.59 (0.85,2.95)	46 (26,69)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	1.53 (0.76,3.28)	47 (25,71)	0.00
	Propiverine ER 20mg q.d	[41]	6	1832	1.51 (1.03,2.3)	47 (32,63)	0.00
	Mirabegron ER 200mg q.d	[53]	1	167	1.47 (0.59,3.52)	49 (23,75)	0.00
	OnaBoNT-A 100u trigone sparing	[72]	3	610	1.47 (0.43,4.98)	49 (18,76)	0.00
	ONO-8539 100mg b.i.d	[60]	1	87	1.41 (0.33,5.12)	50 (17,77)	0.00
	Terodiline 25mg b.i.d	[28]	2	80	1.4 (0.42,6.29)	51 (12,77)	0.00
	Darifenacin ER 7.5mg q.d	[39]	3	414	1.41 (0.68,3.18)	51 (26,73)	0.00
	Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	1	21	1.36 (0.01,100.8)	52 (2,81)	0.02
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	2	1018	1.35 (0.7,2.78)	52 (28,71)	0.00
	Solifenacin ER 5 - 10mg q.d	[31]	4	1738	1.35 (0.86,2.1)	52 (35,68)	0.00

TABLE C.2: Estimated odds ratios (95% credible intervals) for discontinuations due to adverse events obtained from random effects network meta-analysis (cont.)

	Mirabegron ER 25mg q.d	[50]	3	811	1.37 (0.86,2.22)	52 (34,69)	0.00
	Solifenacin ER 5mg q.d	[29]	11	2279	1.3 (0.92,1.8)	53 (40,67)	0.00
	Serlopitant 1mg q.d	[108]	1	110	1.24 (0.29,4.48)	55 (20,79)	0.00
	Mirabegron ER 50mg q.d	[51]	10	4331	1.11 (0.87,1.42)	59 (49,70)	0.00
	Naftopidil 25mg q.d	[114]	1	22	1.1 (0,124.9)	60 (1,81)	0.03
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	1.11 (0.1,10.29)	60 (9,80)	0.00
	Tolterodine ER 4mg q.d	[4]	27	8732	1.09 (0.9,1.29)	60 (51,69)	0.00
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	227	1.11 (0.55,2.2)	60 (34,76)	0.00
	Oxybutynin ER 10mg q.d	[8]	2	576	1.1 (0.64,1.93)	60 (39,74)	0.00
	Tolterodine IR 2mg b.i.d	[5]	15	2645	1.03 (0.78,1.36)	62 (52,72)	0.00
	Mirabegron ER 100mg q.d	[52]	5	2125	1.04 (0.77,1.4)	62 (49,72)	0.00
	ZD0947IL 25mg/day	[58]	1	90	1.04 (0.34,3.22)	62 (26,78)	0.00
	Placebo	[1]	77	16882	NA	64 (55,71)	0.00
	Trospium 30mg/day + Solifenacin 10mg/day (cyclic)	[123]	1	55	0.96 (0,205.2)	65 (2,81)	0.01
	Solifenacin ER 5 - 10mg q.d + BT	[77]	1	304	0.94 (0.41,2.19)	65 (35,77)	0.00
	Trospium chloride IR 45mg t.i.d	[47]	1	828	0.93 (0.42,2.13)	66 (37,77)	0.00
	Fesoterodine IR 8mg b.id	[121]	1	47	0.89 (0.11,10.11)	67 (10,80)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	1	52	0.91 (0,25.58)	68 (5,81)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	4	553	0.86 (0.47,1.6)	68 (46,77)	0.00
	Tolerodine ER 4mg q.d + Neurostimulation	[96]	1	20	0.8 (0,71.57)	69 (1,81)	0.03
	Solifenacin ER 5 - 15mg q.d	[34]	1	385	0.78 (0.35,1.7)	70 (42,78)	0.00
	Tolterodine IR 1mg b.i.d	[6]	3	281	0.79 (0.32,1.72)	70 (43,78)	0.00
	Solifenacin ER 2.5mg q.d	[32]	1	41	0.77 (0.04,4.42)	71 (20,81)	0.00
	ONO-8539 30mg b.i.d	[59]	1	88	0.56 (0.09,2.86)	74 (29,81)	0.00
	Fesoterodine IR 4mg b.i.d	[120]	1	43	0.51 (0.04,9.77)	75 (11,81)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	1	26	0.26 (0,3.72)	78 (24,81)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment has the highest incidence of adverse events overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average.

p(Worst) denotes the probability that the intervention in question has the highest incidence of adverse events. The probability worst is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

C.9 Results for discontinuations due to a lack of efficacy

TABLE C.3: Estimated odds ratios (95% credible intervals) for discontinuations due to a lack of efficacy obtained from random effects network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Odds ratio† (95% CrI)	Rank (95% CrI)	p(Worst)
	ONO-8539 30mg b.i.d	[59]	1	88	3.42 (0.25,74.38)	8 (1,42)	0.08
	Serlopitant 4mg q.d	[109]	1	114	3.48 (0.4,47.1)	8 (1,37)	0.06
	Cizolirtine citrate 400mg b.i.d	[57]	1	54	2.79 (0.05,1937)	9 (1,51)	0.13
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	1	188	3.01 (0.21,88.54)	9 (1,44)	0.09
	Netupitant 100mg q.d	[111]	1	61	2.32 (0.17,52.58)	10 (1,46)	0.06
	Serlopitant 1mg q.d	[108]	1	110	2.18 (0.19,32.73)	11 (1,45)	0.03
	Imidafenacin 0.1mg b.i.d	[36]	1	55	2.24 (0.47,10.49)	11 (2,33)	0.01
	Oxybutynin IR 5mg t.i.d	[7]	1	27	1.7 (0,1220)	13 (1,55)	0.10
	Solifenacin ER 5 - 10mg q.d	[31]	2	803	1.68 (0.43,7.84)	14 (3,36)	0.00
	Trospium 60mg/day + Solifenacin 20mg/day (cyclic)	[139]	1	58	1.4 (0.01,312)	15 (1,54)	0.07
	OnaBoNT-A 100u trigone sparing	[72]	2	333	1.43 (0.23,10.13)	15 (3,43)	0.01
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	227	1.4 (0.16,13.1)	16 (2,46)	0.01
	ZD0947IL 25mg/day	[58]	1	90	1.14 (0.11,12.19)	18 (2,48)	0.01
	Trospium 30mg/day + Solifenacin 10mg/day (cyclic)	[123]	1	55	1.11 (0,273.8)	19 (1,55)	0.07
	Propiverine ER 20mg q.d	[41]	1	176	1.11 (0.16,10.99)	19 (3,46)	0.01
	Placebo	[1]	43	10186	NA	20 (13,28)	0.00
	Trospium 30mg/day + Solifenacin 10mg/day (continuous)	[126]	1	62	0.97 (0,202.7)	21 (1,55)	0.06
	OnaBoNTA 50u trigone sparing	[74]	1	56	0.81 (0.08,7.58)	24 (4,49)	0.00
	Solifenacin ER 5mg q.d	[29]	8	1735	0.77 (0.3,1.89)	25 (12,41)	0.00
	Mirabegron ER 200mg q.d	[53]	1	167	0.72 (0.09,4.03)	26 (6,49)	0.00
	Fesoterodine ER 4mg q.d	[25]	4	2462	0.71 (0.36,1.37)	26 (14,40)	0.00
	Solifenacin ER 10mg q.d	[30]	4	950	0.71 (0.25,1.93)	26 (12,43)	0.00
	Solifenacin ER 5 - 10mg q.d + BT	[77]	1	304	0.68 (0.02,14.13)	27 (2,53)	0.01
	Duloxetine 60mg b.i.d	[66]	1	153	0.7 (0.19,2.51)	27 (9,46)	0.00
	Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	1	21	0.65 (0,155.4)	28 (1,55)	0.06
	Darifenacin ER 7.5mg q.d	[39]	3	414	0.66 (0.16,2.7)	28 (9,47)	0.00
	Naftopidil 25mg q.d	[114]	1	22	0.61 (0,132.7)	29 (1,55)	0.05
	Darifenacin ER 15mg q.d.	[40]	2	321	0.62 (0.14,2.47)	29 (9,48)	0.00
	Tolterodine IR 2mg b.i.d	[5]	5	936	0.61 (0.13,2.26)	29 (11,47)	0.00
	Fesoterodine ER 8mg q.d	[26]	4	2079	0.6 (0.28,1.23)	30 (15,43)	0.00
	ONO-8539 100mg b.i.d	[60]	1	87	0.53 (0,22.22)	32 (2,55)	0.01
	Tolterodine ER 4mg q.d	[4]	18	6491	0.52 (0.33,0.79)	32 (22,41)	0.00
	Trospium chloride IR 20mg b.i.d	[45]	1	267	0.51 (0.01,15.53)	33 (2,54)	0.01
	ONO-8539 300mg b.i.d	[61]	1	88	0.51 (0,19.88)	33 (2,55)	0.01
	Oxybutynin ER 10mg q.d	[8]	2	576	0.5 (0.08,2.94)	33 (9,50)	0.00
	Mirabegron ER 50mg q.d	[51]	10	4331	0.51 (0.31,0.87)	33 (21,42)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	105	0.48 (0.02,5.58)	34 (5,53)	0.00
	Serlopitant 0.25mg q.d	[107]	1	110	0.46 (0,11.39)	35 (4,54)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1182	0.44 (0.19,0.95)	36 (20,46)	0.00
	OnaBoNTA 150u trigone sparing	[75]	1	50	0.4 (0.01,4.87)	37 (6,54)	0.00

TABLE C.3: Estimated odds ratios (95% credible intervals) for discontinuations due to a lack of efficacy obtained from random effects network meta-analysis (cont.)

	Mirabegron ER 25mg q.d	[50]	3	811	0.4 (0.14,1.13)	37 (18,48)	0.00
	Mirabegron ER 100mg q.d	[52]	5	2125	0.4 (0.2,0.81)	37 (23,46)	0.00
	OnaBoNTA 300u trigone sparing	[76]	1	55	0.35 (0.01,4.32)	39 (7,54)	0.00
	Darifenacin ER 30mg q.d	[38]	1	115	0.36 (0.01,3.38)	39 (8,54)	0.00
	Solifenacin ER 2.5mg q.d	[32]	1	41	0.36 (0.01,3.19)	39 (8,53)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	65	0.33 (0,10.6)	40 (3,55)	0.01
	Netupitant 50mg q.d	[110]	1	62	0.31 (0,13.25)	40 (3,55)	0.01
	Netupitant 200mg q.d	[112]	1	61	0.31 (0,12.7)	40 (3,55)	0.01
	Oxybutynin IR 5mg b.i.d	[18]	2	154	0.31 (0.01,4.83)	40 (6,54)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	0.3 (0,10.42)	41 (3,55)	0.01
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	102	0.23 (0,4.1)	43 (6,55)	0.00
	Pregabalin 150mg b.i.d	[62]	1	105	0.16 (0,3.48)	46 (7,55)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	1	52	0.13 (0,2.74)	47 (10,55)	0.00
	Solifenacin ER 20mg q.d	[33]	1	37	0.13 (0,2.05)	47 (11,55)	0.00
	Oxybutynin intravesically 5mg t.i.d	[14]	1	23	0.03 (0,0.45)	52 (35,55)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment has the highest incidence of discontinuations due to a lack of efficacy overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Worst) denotes the probability that the intervention in question has the highest incidence of discontinuations. The probability worst is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

C.10 Convergence diagnostics

FIGURE C.1: Brooks-Gelman-Rubin plots for urinary incontinence episodes

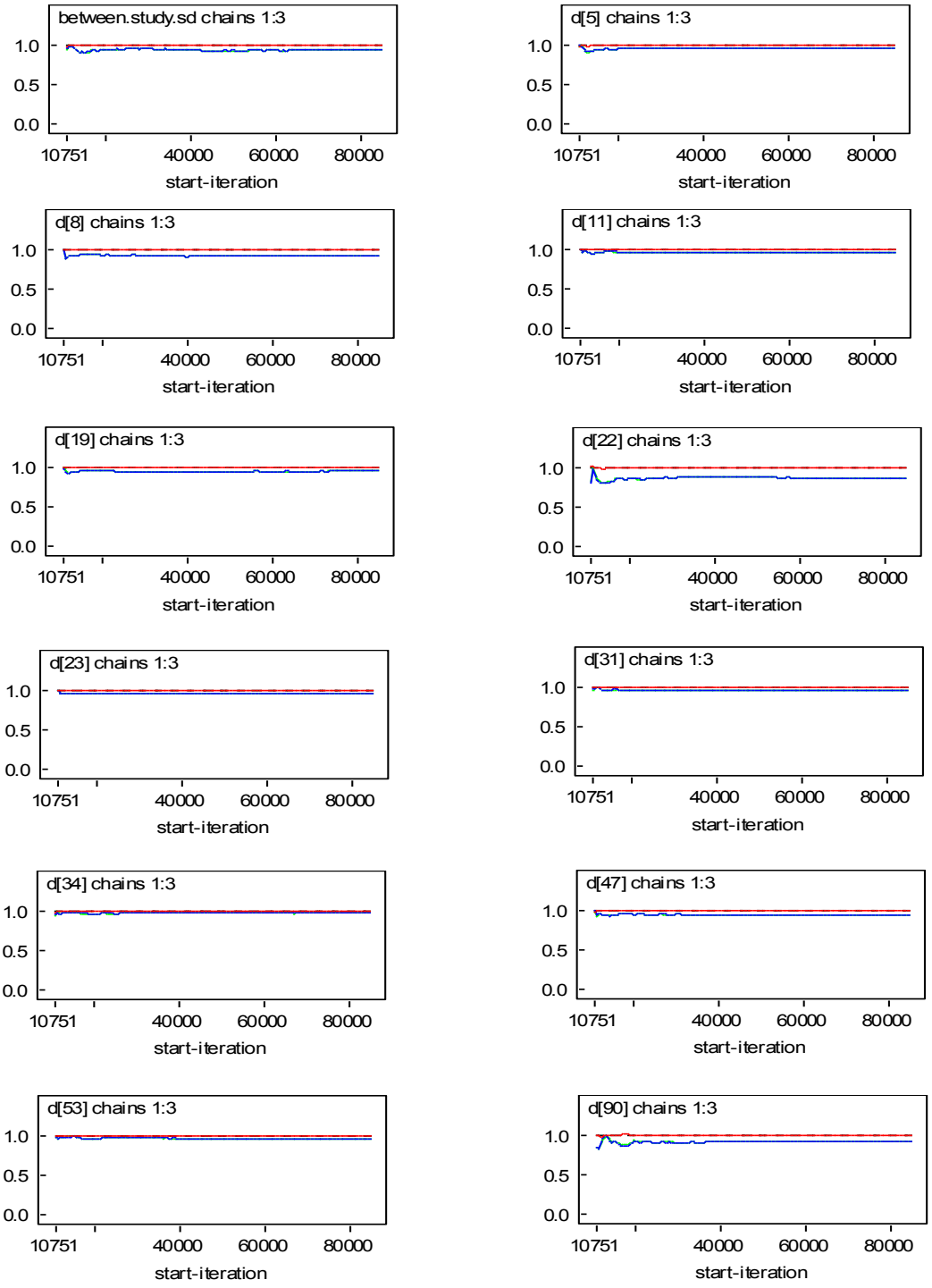


FIGURE C.2: Autocorrelation plots for urinary incontinence episodes

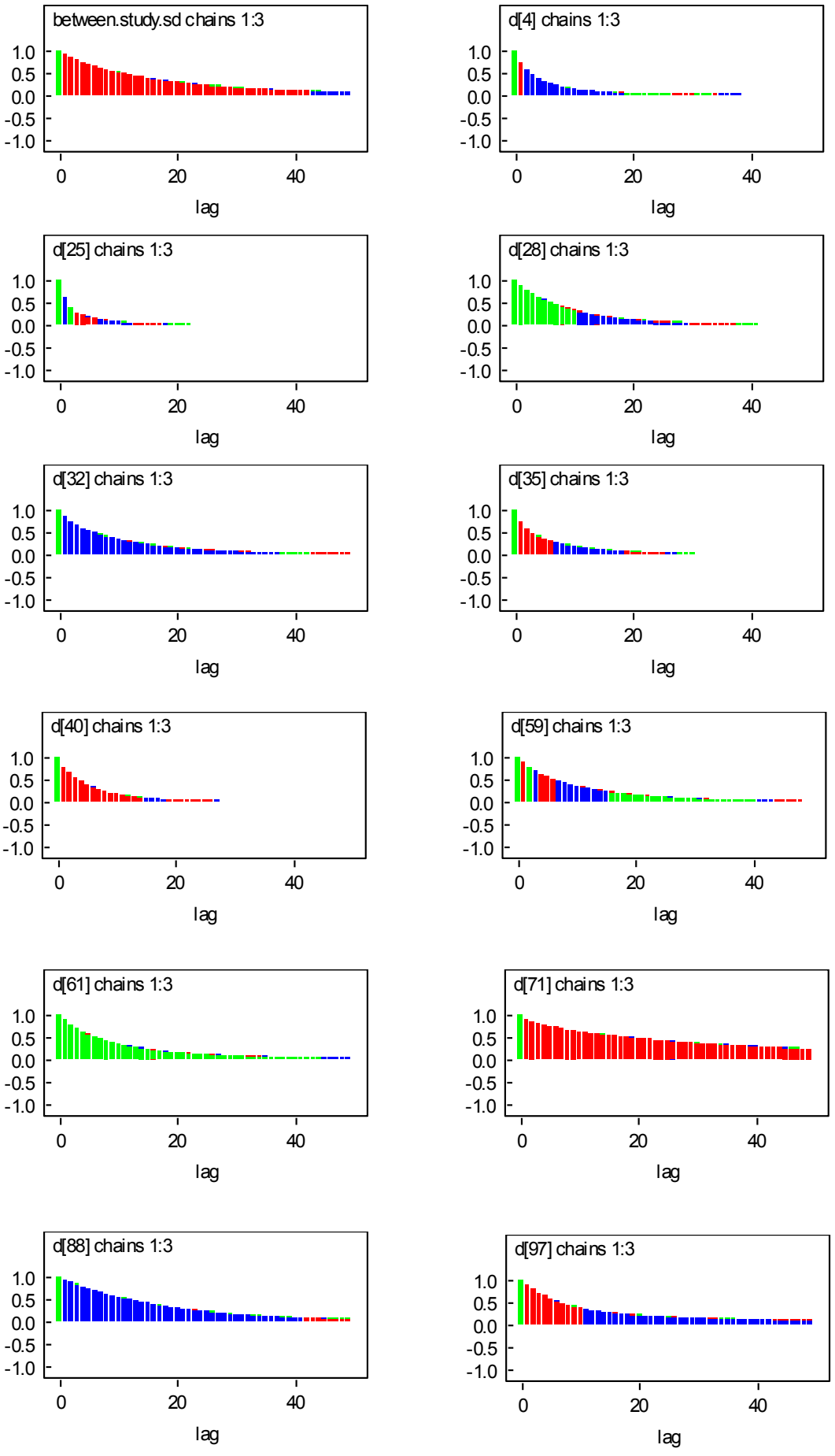


FIGURE C.3: History and trace plots for urinary incontinence episodes

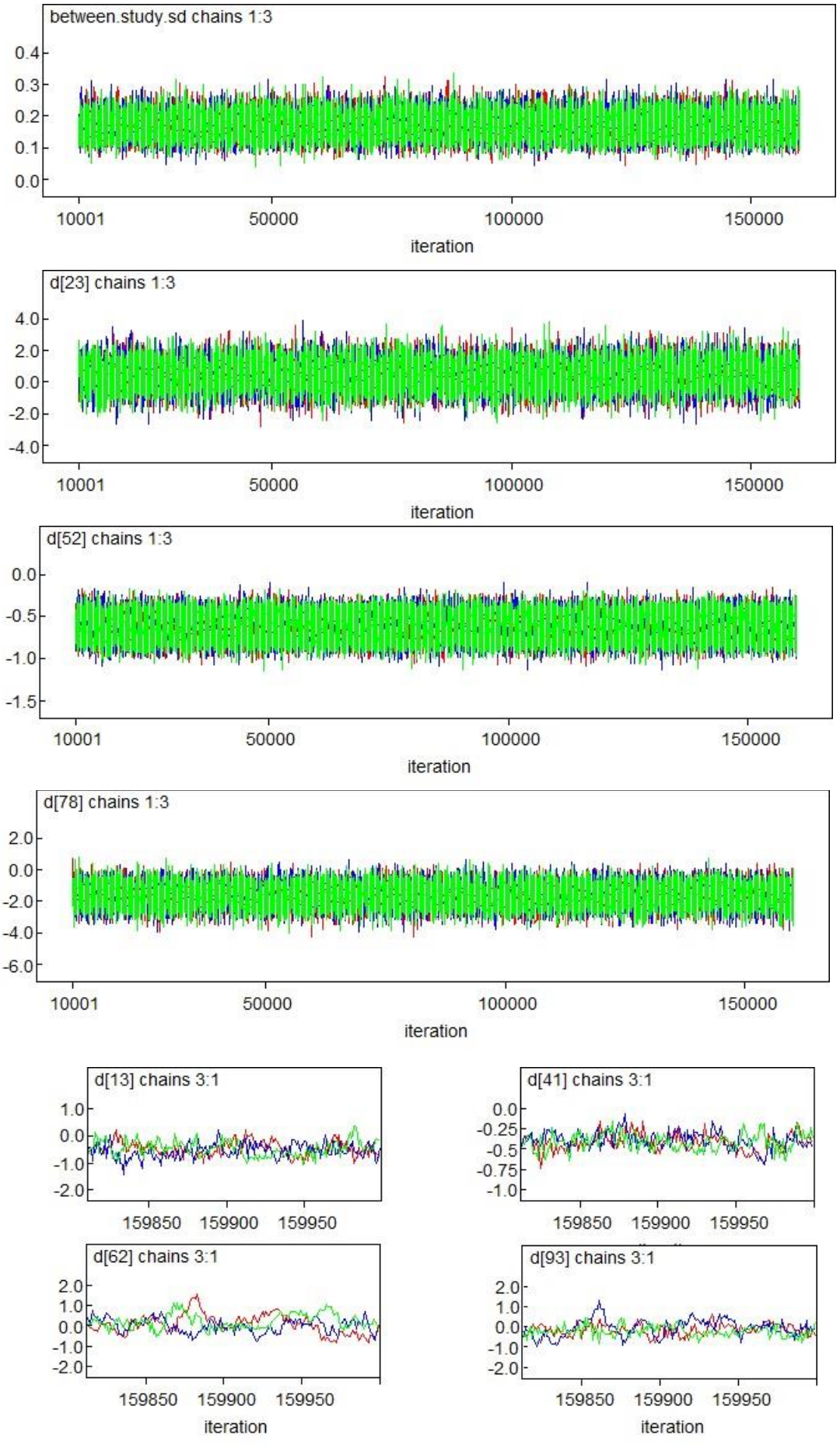
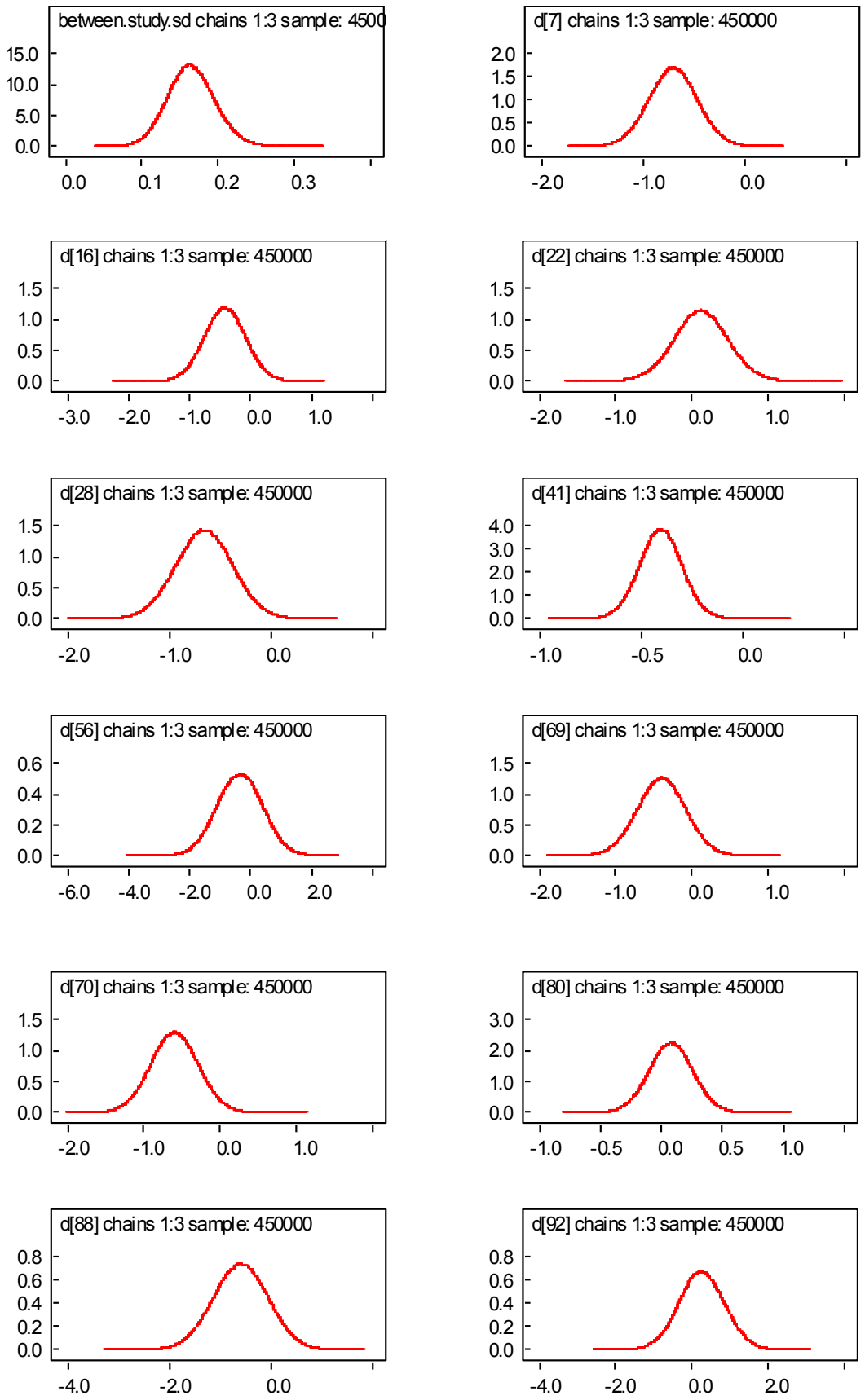


FIGURE C.4: Density plots for urinary incontinence episodes



C.11 Assessing inconsistencies

TABLE C.4: Inconsistency between direct and indirect information for urinary incontinence episodes

Node 1	Node 2	Indirect comparison	Direct comparison	Posterior probability
1	4	-0.262	-0.52	0.07
1	5	-0.2409	-0.4872	0.15
1	11	-1.162	0.1165	0.90
1	25	-0.2211	-0.4863	0.26
1	29	-0.4622	-0.7921	0.12
1	30	-1.332	-0.8163	0.94
1	31	-1.017	-0.5315	0.96
1	36	-0.5851	-0.3688	0.73
1	41	-0.361	-0.4104	0.45
1	43	-0.7479	-0.4249	0.77
1	50	-0.683	-0.07234	0.51
1	51	-0.6012	-0.4881	0.70
1	52	-0.6665	-0.5121	0.74
4	5	0.03369	0.168	0.76
4	10	0.2155	0.08031	0.37
4	25	0.02261	-0.03314	0.39
4	26	-0.3058	-0.1002	0.85
4	29	-0.2199	2.091	0.94
4	31	-0.006946	-0.4905	0.04
4	50	-0.0272	-0.4041	0.18
4	51	-0.009768	-0.127	0.29
4	52	-0.1145	-0.1479	0.45
5	7	-0.4066	-0.2069	0.64
5	24	0.7465	-0.4936	0.11
5	29	-0.3638	-0.2061	0.73
5	30	-0.3918	-0.489	0.35
5	43	0.02928	-0.2886	0.23
7	47	-0.07269	0.1418	0.63
10	11	-0.863	0.4357	0.90
11	24	-1.074	0.206	0.90
25	26	0.176	-0.2964	0.02
25	36	0.1291	-0.1469	0.25
29	30	-0.2858	-0.1519	0.65
29	41	0.3466	0.2952	0.42
30	41	0.5251	0.3656	0.29
36	41	0.09398	-0.08585	0.26
50	52	0.03779	0.1945	0.68

TABLE C.5: Inconsistency between direct and indirect information for voiding episodes

Node 1	Node 2	Indirect comparison	Direct comparison	Posterior probability
1	4	-0.7202	-0.6276	0.70
1	5	-0.9676	-0.6809	0.87
1	7	-0.5795	-0.7325	0.37
1	25	0.1793	-0.7518	0.07
1	29	-0.85	-0.6739	0.75
1	30	-1.041	-1.256	0.22
1	31	-0.8551	-0.7684	0.61
1	36	-1.374	-0.4423	0.93
1	41	-0.2099	-0.7151	0.25
1	50	-0.8145	0.01491	0.51
1	80	-0.4545	-1.909	0.01
2	84	1.773	-3.178	0.00
4	5	-0.07335	0.06098	0.72
4	10	-0.03612	0.2311	0.73
4	25	-0.03886	-0.1889	0.23
4	26	-0.5369	-0.2565	0.91
4	29	-0.07223	-0.1065	0.49
4	31	-0.1257	-0.21	0.39
4	51	0.02739	-0.2396	0.18
4	52	-0.1281	-0.05333	0.60
4	80	-1.1	0.3117	0.98
4	83	-0.5777	0.1136	0.60
4	84	1.963	-0.08327	0.07
4	85	0.2242	-0.1628	0.33
5	7	-2.952	0.1133	0.99
5	24	-0.7411	0.002918	0.76
5	29	0.04387	-0.1611	0.21
5	30	-0.483	-0.5196	0.45
5	83	0.07913	0.01491	0.50
5	85	-0.9623	0.6779	0.95
7	57	-1.457	1.582	0.99
24	80	-0.7609	0.02693	0.77
25	26	-0.1234	-0.3202	0.20
25	36	0.3208	-0.6548	0.06
29	30	-0.6967	-0.3901	0.84
29	41	0.06932	0.0246	0.43
30	41	0.5791	0.3317	0.18
36	41	-0.1261	-0.2606	0.32
50	52	0.1482	0.01491	0.50
80	83	0.8374	0.4498	0.45

TABLE C.6: Inconsistency between direct and indirect information for urgency episodes

Node 1	Node 2	Indirect comparison	Direct comparison	Posterior probability
1	4	-0.9524	-0.7435	0.62
1	5	-1.002	-0.6	0.67
1	25	0.04169	-1.125	0.12
1	29	-1.455	-0.8646	0.85
1	30	-0.8897	-1.567	0.16
1	31	-1.223	-0.7647	0.71
1	36	-1.619	-0.4641	0.87
1	41	-0.535	-0.936	0.34
1	50	-0.4109	-0.3444	0.53
4	25	-0.17	-0.3524	0.38
4	26	-0.6817	-0.5114	0.62
4	29	-0.2268	-0.7456	0.29
4	31	0.03496	-0.4335	0.29
4	51	0.1829	-0.11	0.32
4	52	-0.06047	0.123	0.60
25	26	-0.5699	-0.4332	0.58
25	36	0.6683	-0.507	0.12
29	30	-0.5881	-0.1529	0.70
29	41	0.08579	0.2307	0.58
30	41	0.4845	0.4729	0.49
36	41	-0.07747	-0.4173	0.35
52	50	0.1475	0.3407	0.58

TABLE C.7: Inconsistency between direct and indirect information for nocturia episodes

Node 1	Node 2	Indirect comparison	Direct comparison	Posterior probability
1	4	-0.1621	-0.1229	0.67
1	20	-1.012	-0.2248	0.99
1	29	-0.354	-0.1123	0.94
1	30	-0.1457	-0.1607	0.46
1	41	0.6432	-0.08942	0.05
1	51	-0.183	-0.1352	0.74
1	52	-0.08501	-0.1625	0.16
1	80	1.149	-0.4803	0.04
1	85	0.1602	-0.5495	0.01
4	25	-0.01638	0.01831	0.68
4	26	-0.06288	0.02783	0.86
4	29	-0.01588	0.95	0.82
4	51	-0.0634	-0.005543	0.75
4	52	-0.03479	0.044	0.82
4	83	-1.666	-0.09225	0.95
5	29	0.5963	-0.08628	0.03
5	30	-0.005768	0.002409	0.53
5	85	-0.5016	0.1953	0.98
25	26	0.05542	-0.01897	0.20
41	29	-0.4179	-0.02633	0.91
41	30	-0.05848	-0.0646	0.48
83	80	-0.2341	1.4	0.98

C.12 Sensitivity to variance priors

TABLE C.8: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters

Treatment	Code	Median difference† (95% CrI)	Median difference†† (95% CrI)	Median difference††† (95% CrI)
Sacral nerve stimulation	[81]	-8.72 (-11.33,-6.09)	-8.82 (-11.56,-6.16)	-8.68 (-11.28,-6.16)
OnaBoNT-A 200U trigone sparing	[73]	-2.3 (-3.16,-1.42)	-2.31 (-3.17,-1.41)	-2.3 (-3.16,-1.41)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-2.2 (-4.06,-0.36)	-2.26 (-4.15,-0.37)	-2.25 (-4.14,-0.37)
Solifenacin/trospium + placebo injection	[100]	-1.97 (-2.96,-1.01)	-1.98 (-2.97,-1.01)	-1.98 (-2.97,-0.99)
Electrostimulation + PFMT + BT	[97]	-1.93 (-2.94,-0.91)	-1.91 (-2.9,-0.92)	-1.91 (-2.91,-0.91)
OnaBoNT-A 100u trigone sparing	[72]	-1.88 (-2.31,-1.45)	-1.88 (-2.31,-1.45)	-1.88 (-2.32,-1.45)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.63 (-2.73,-0.54)	-1.65 (-2.71,-0.55)	-1.66 (-2.73,-0.57)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.39 (-4,1.06)	-1.44 (-3.89,0.96)	-1.48 (-3.9,1.06)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.29 (-1.7,-0.89)	-1.29 (-1.7,-0.9)	-1.29 (-1.71,-0.89)
Oxybutynin intravesically 5mg t.i.d	[14]	-1.19 (-2.49,0.1)	-1.19 (-2.43,0.08)	-1.19 (-2.48,0.1)
Mirabegron IR 100mg b.i.d	[48]	-1.12 (-1.98,-0.25)	-1.11 (-1.95,-0.24)	-1.1 (-1.96,-0.25)
Oxybutynin ER 10mg q.d	[8]	-0.98 (-1.55,-0.42)	-0.99 (-1.54,-0.42)	-0.99 (-1.55,-0.42)
Oxybutynin IR 3mg t.i.d	[19]	-0.9 (-1.28,-0.52)	-0.9 (-1.27,-0.53)	-0.9 (-1.28,-0.52)
Solifenacin ER 10mg q.d	[30]	-0.88 (-1.14,-0.63)	-0.88 (-1.14,-0.63)	-0.89 (-1.14,-0.63)
Imidafenacin 0.25mg b.i.d	[37]	-0.82 (-1.43,-0.21)	-0.83 (-1.44,-0.21)	-0.82 (-1.44,-0.21)
Propiverine ER 30mg q.d	[42]	-0.79 (-1.31,-0.28)	-0.79 (-1.3,-0.29)	-0.79 (-1.32,-0.29)
Tolterodine ER 4mg q.d + BT	[87]	-0.78 (-1.45,-0.12)	-0.77 (-1.43,-0.12)	-0.78 (-1.44,-0.11)
Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	-0.78 (-1.71,0.16)	-0.77 (-1.7,0.15)	-0.77 (-1.68,0.16)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	-0.75 (-1.45,-0.01)	-0.75 (-1.47,-0.02)	-0.75 (-1.48,0)
Darifenacin ER 30mg q.d	[38]	-0.74 (-1.36,-0.12)	-0.74 (-1.36,-0.12)	-0.74 (-1.37,-0.11)
Tolterodine ER 2mg b.i.d + Estrogen 0.625mg 2xwk	[99]	-0.75 (-1.24,-0.25)	-0.75 (-1.24,-0.25)	-0.75 (-1.25,-0.25)
Solifenacin ER 5 - 10mg q.d	[31]	-0.73 (-1,-0.45)	-0.73 (-1,-0.46)	-0.73 (-1,-0.45)
Solifenacin ER 5mg q.d	[29]	-0.71 (-0.96,-0.46)	-0.71 (-0.96,-0.46)	-0.71 (-0.96,-0.46)
Oxybutynin IR 5mg t.i.d	[7]	-0.71 (-1.18,-0.25)	-0.71 (-1.17,-0.24)	-0.71 (-1.17,-0.24)
Fesoterodine ER 8mg q.d	[26]	-0.69 (-0.89,-0.5)	-0.69 (-0.89,-0.5)	-0.69 (-0.89,-0.5)
Oxybutynin gel 56mg/day	[135]	-0.69 (-1.59,0.18)	-0.7 (-1.58,0.15)	-0.71 (-1.57,0.18)
Mirabegron ER 25mg q.d	[50]	-0.67 (-0.99,-0.35)	-0.67 (-0.98,-0.36)	-0.67 (-0.99,-0.35)
Terodiline 25mg b.i.d	[28]	-0.66 (-1.21,-0.1)	-0.66 (-1.21,-0.11)	-0.66 (-1.22,-0.11)
Trospium chloride ER 60mg q.d	[44]	-0.66 (-1.09,-0.24)	-0.66 (-1.08,-0.24)	-0.66 (-1.08,-0.24)
Cizolirtine citrate 400mg b.i.d	[57]	-0.63 (-1.17,-0.08)	-0.63 (-1.18,-0.1)	-0.63 (-1.17,-0.09)
Mirabegron ER 100mg q.d	[52]	-0.62 (-0.83,-0.41)	-0.62 (-0.82,-0.41)	-0.62 (-0.83,-0.41)
Solifenacin ER 5 - 15mg q.d	[34]	-0.61 (-1.11,-0.1)	-0.61 (-1.12,-0.11)	-0.61 (-1.12,-0.1)
Solabegron IR 125mg b.i.d	[55]	-0.6 (-0.94,-0.26)	-0.6 (-0.93,-0.27)	-0.6 (-0.94,-0.26)
Elocalcitol 75mg	[70]	-0.6 (-1.2,0.01)	-0.6 (-1.2,0)	-0.6 (-1.2,0.01)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.6 (-1.66,0.46)	-0.61 (-1.67,0.46)	-0.61 (-1.66,0.44)
Propiverine IR 15mg b.i.d	[43]	-0.57 (-1,-0.16)	-0.58 (-1,-0.16)	-0.58 (-1.01,-0.16)
Mirabegron ER 200mg q.d	[53]	-0.57 (-1.13,-0.01)	-0.57 (-1.13,-0.02)	-0.57 (-1.14,-0.01)
Mirabegron ER 50mg q.d	[51]	-0.57 (-0.76,-0.38)	-0.57 (-0.75,-0.38)	-0.57 (-0.76,-0.38)
Darifenacin ER 15mg q.d.	[40]	-0.51 (-0.97,-0.05)	-0.51 (-0.97,-0.05)	-0.52 (-0.97,-0.05)
Mirabegron IR 150mg b.i.d	[49]	-0.51 (-1.57,0.54)	-0.52 (-1.61,0.57)	-0.51 (-1.56,0.55)
Oxybutynin chloride topical gel 1g/day	[13]	-0.5 (-1.02,0.02)	-0.5 (-1.01,0.02)	-0.5 (-1.02,0.02)
Tolterodine ER 4mg q.d	[4]	-0.51 (-0.61,-0.4)	-0.5 (-0.61,-0.4)	-0.5 (-0.61,-0.4)
Fesoterodine ER 4mg q.d	[25]	-0.47 (-0.66,-0.29)	-0.47 (-0.65,-0.29)	-0.47 (-0.66,-0.28)
Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	-0.48 (-0.79,-0.17)	-0.48 (-0.78,-0.17)	-0.48 (-0.79,-0.17)
Tolterodine IR 2mg b.i.d	[5]	-0.45 (-0.6,-0.3)	-0.45 (-0.59,-0.3)	-0.45 (-0.6,-0.3)
Tolterodine IR 2mg b.i.d + PFMT	[95]	-0.45 (-1.07,0.18)	-0.45 (-1.07,0.17)	-0.45 (-1.08,0.18)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.44 (-1.11,0.23)	-0.43 (-1.09,0.23)	-0.43 (-1.09,0.23)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.43 (-1.08,0.22)	-0.42 (-1.06,0.23)	-0.42 (-1.07,0.23)
Propiverine ER 20mg q.d	[41]	-0.41 (-0.61,-0.2)	-0.4 (-0.61,-0.21)	-0.41 (-0.61,-0.2)
Imidafenacin 0.1mg b.i.d	[36]	-0.4 (-0.7,-0.11)	-0.4 (-0.69,-0.11)	-0.4 (-0.7,-0.11)
Cizolirtine citrate 200mg b.i.d	[56]	-0.4 (-1.83,1.03)	-0.37 (-1.82,1.06)	-0.36 (-1.84,1.1)
Elocalcitol 150mg	[69]	-0.4 (-1.03,0.22)	-0.4 (-1.03,0.23)	-0.4 (-1.02,0.23)
Oxybutynin trandermal 3.9mg/day	[10]	-0.33 (-0.67,0)	-0.33 (-0.66,-0.01)	-0.33 (-0.67,0)
Tolterodine IR 1mg b.i.d	[6]	-0.33 (-0.69,0.02)	-0.33 (-0.69,0.02)	-0.33 (-0.69,0.03)
Oxbutynin patch 73.5mg	[15]	-0.31 (-0.67,0.05)	-0.31 (-0.66,0.04)	-0.31 (-0.66,0.05)

TABLE C.8: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters (cont.)

Fesoterodine ER 4 - 8mg q.d	[27]	-0.28 (-0.52,-0.05)	-0.29 (-0.52,-0.05)	-0.29 (-0.53,-0.05)
Oxybutynin gel 84mg/day	[134]	-0.29 (-1.13,0.52)	-0.3 (-1.11,0.51)	-0.31 (-1.11,0.53)
Darifenacin ER 7.5mg q.d	[39]	-0.27 (-0.84,0.3)	-0.27 (-0.84,0.3)	-0.27 (-0.85,0.3)
Oxybutynin ER 15mg q.d	[9]	-0.28 (-1.41,0.86)	-0.29 (-1.38,0.82)	-0.29 (-1.38,0.82)
Duloxetine 40mg b.i.d	[65]	-0.26 (-0.75,0.25)	-0.26 (-0.75,0.23)	-0.26 (-0.76,0.24)
PFMT + BT	[89]	-0.26 (-0.87,0.37)	-0.24 (-0.86,0.37)	-0.23 (-0.85,0.37)
Imidafenacin 0.05mg b.i.d	[35]	-0.24 (-0.77,0.27)	-0.25 (-0.77,0.28)	-0.24 (-0.77,0.28)
Solabegron IR 50mg b.i.d	[54]	-0.2 (-0.54,0.14)	-0.2 (-0.53,0.13)	-0.2 (-0.54,0.14)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.14 (-0.81,0.52)	-0.15 (-0.8,0.51)	-0.14 (-0.8,0.52)
Tarafenacin 0.4mg q.d	[82]	-0.13 (-0.85,0.61)	-0.13 (-0.83,0.58)	-0.13 (-0.86,0.58)
ONO-8539 100mg b.i.d	[60]	-0.11 (-0.85,0.63)	-0.11 (-0.85,0.64)	-0.11 (-0.86,0.63)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.08 (-0.39,0.23)	-0.08 (-0.39,0.22)	-0.08 (-0.39,0.23)
ZD0947IL 25mg/day	[58]	-0.1 (-0.94,0.76)	-0.1 (-0.96,0.77)	-0.1 (-0.96,0.76)
Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	-0.08 (-0.72,0.55)	-0.07 (-0.7,0.55)	-0.06 (-0.69,0.56)
Lipo-BoNTA 200U	[138]	-0.06 (-0.96,0.83)	-0.06 (-0.94,0.83)	-0.05 (-0.95,0.83)
Oxybutynin transdermal 1.3mg/day	[11]	-0.03 (-0.66,0.61)	-0.03 (-0.66,0.6)	-0.04 (-0.67,0.61)
Placebo	[1]	NA	NA	NA
Estradiol 25mg	[68]	0 (-0.38,0.38)	0 (-0.37,0.37)	0 (-0.38,0.38)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	0 (-0.57,0.56)	0 (-0.56,0.55)	-0.01 (-0.57,0.55)
Pregabalin 150mg b.i.d	[62]	0 (-0.88,0.9)	-0.01 (-0.88,0.87)	0 (-0.88,0.88)
Bladder Training (BT)/Behaviour Therapy	[85]	0.02 (-0.59,0.62)	0.04 (-0.58,0.62)	0.05 (-0.56,0.63)
Oxybutynin ER 5 - 30mg/day	[22]	0.09 (-0.62,0.78)	0.11 (-0.61,0.79)	0.12 (-0.59,0.8)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	0.07 (-0.45,0.59)	0.07 (-0.44,0.58)	0.07 (-0.45,0.59)
Electrostimulation	[80]	0.08 (-0.28,0.44)	0.08 (-0.27,0.43)	0.08 (-0.28,0.43)
Tarafenacin 0.2mg q.d	[90]	0.11 (-0.63,0.86)	0.11 (-0.62,0.84)	0.11 (-0.63,0.85)
Percutaneous tibial nerve stimulation	[83]	0.18 (-1.16,1.54)	0.21 (-1.14,1.58)	0.2 (-1.17,1.56)
Oxybutynin ER 2.5mg q.d + BT	[92]	0.22 (-0.94,1.4)	0.23 (-0.97,1.39)	0.25 (-0.9,1.44)
Oxybutynin transdermal 2.6mg/day	[12]	0.29 (-0.37,0.95)	0.29 (-0.36,0.94)	0.29 (-0.36,0.94)
Oxybutynin IR 5mg b.i.d	[18]	0.35 (-0.3,0.99)	0.35 (-0.29,1)	0.35 (-0.3,1)
Control	[2]	0.33 (-0.67,1.31)	0.34 (-0.63,1.28)	0.37 (-0.59,1.35)
Emepronium bromide ER 200mg q.d	[63]	0.34 (-0.33,1.01)	0.35 (-0.33,1.03)	0.34 (-0.33,1.02)
Flavoxate chloride 200mg q.d	[64]	0.33 (-0.35,1.01)	0.34 (-0.34,1.02)	0.33 (-0.35,1.01)
Reflexology	[71]	0.33 (-0.76,1.41)	0.34 (-0.72,1.37)	0.37 (-0.69,1.44)
Vaginal oestrogen cream 1.25mg/day	[132]	0.38 (-0.15,0.91)	0.38 (-0.14,0.9)	0.38 (-0.15,0.9)
Oxybutynin IR 5 - 20mg	[23]	0.46 (-0.9,1.84)	0.49 (-0.91,1.86)	0.49 (-0.88,1.88)
Oxybutynin ER 2.5mg q.d	[20]	0.52 (-0.5,1.55)	0.53 (-0.48,1.52)	0.55 (-0.46,1.56)
Trospium chloride IR 15mg t.i.d	[46]	0.52 (-0.41,1.41)	0.53 (-0.38,1.45)	0.54 (-0.38,1.45)
ONO-8539 300mg b.i.d	[61]	0.51 (-0.23,1.25)	0.53 (-0.21,1.26)	0.53 (-0.22,1.27)
ONO-8539 30mg b.i.d	[59]	0.58 (-0.18,1.33)	0.58 (-0.16,1.33)	0.59 (-0.17,1.33)
Resiniferatoxin 50nM	[67]	0.58 (-1.08,2.28)	0.58 (-1.05,2.24)	0.59 (-1.04,2.21)
Sham therapy	[3]	0.83 (-0.62,2.27)	0.85 (-0.59,2.31)	0.83 (-0.62,2.3)
Oxybutynin ER 5-30mg q.d + BT	[86]	0.92 (-0.31,2.12)	0.94 (-0.29,2.17)	0.95 (-0.27,2.19)

† Between-study standard deviation based on an Uniform(0,5) prior distribution on the standard deviation scale

†† Between-study standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Between-study standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE C.9: Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters

Treatment	**Code**	**Median difference† (95% CrI)**	**Median difference†† (95% CrI)**	**Median difference††† (95% CrI)**
Sacral nerve stimulation	[81]	-7.89 (-12.03,-3.76)	-7.94 (-11.93,-3.88)	-7.95 (-11.93,-3.57)
Electrostimulation + PFMT + BT	[97]	-3.37 (-5.44,-1.23)	-3.4 (-5.48,-1.3)	-3.39 (-5.34,-1.4)
Oxybutynin IR 2.5mg t.i.d	[21]	-3.12 (-4.25,-2.01)	-3.11 (-4.2,-2)	-3.14 (-4.26,-2.02)
OnaBoNT-A 200U trigone sparing	[73]	-2.93 (-4.23,-1.61)	-2.93 (-4.24,-1.62)	-2.94 (-4.25,-1.63)
PFMT + BT	[89]	-2.16 (-4.1,-0.22)	-2.19 (-4.1,-0.32)	-2.19 (-3.96,-0.44)
Estradiol 3mg intravaginally	[128]	-1.87 (-3.8,-0.01)	-1.87 (-3.77,0.01)	-1.93 (-3.72,-0.04)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.93 (-2.7,-1.15)	-1.93 (-2.7,-1.15)	-1.93 (-2.74,-1.14)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.79 (-3.45,-0.05)	-1.77 (-3.43,-0.11)	-1.82 (-3.46,-0.07)
Oxybutynin intravesically 5mg t.i.d	[14]	-1.73 (-4.63,1.12)	-1.71 (-4.43,1.21)	-1.67 (-4.33,0.94)
Cizolirtine citrate 400mg b.i.d	[57]	-1.6 (-2.97,-0.3)	-1.61 (-2.93,-0.32)	-1.6 (-2.91,-0.19)
Propiverine IR 30mg b.i.d	[117]	-1.54 (-3.78,0.64)	-1.51 (-3.51,0.6)	-1.56 (-3.78,0.43)
Estriol 1mg intravesically	[131]	-1.51 (-2.74,-0.27)	-1.49 (-2.69,-0.23)	-1.5 (-2.76,-0.28)
OnaBoNT-A 100U trigone sparing	[72]	-1.5 (-1.9,-1.11)	-1.51 (-1.9,-1.12)	-1.51 (-1.9,-1.12)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.48 (-2.99,0)	-1.49 (-3,0.05)	-1.52 (-3.04,0.01)
Lipo-BoNTA 200U	[138]	-1.44 (-3.09,0.1)	-1.47 (-3.01,0.11)	-1.48 (-3.12,0.19)
Reflexology	[71]	-1.4 (-3.31,0.49)	-1.33 (-3.27,0.59)	-1.4 (-3.21,0.48)
Tolterodine IR 2mg b.i.d + BT	[93]	-1.41 (-2.45,-0.42)	-1.39 (-2.43,-0.46)	-1.39 (-2.51,-0.41)
Oxybutynin ER 10mg q.d	[8]	-1.33 (-1.99,-0.69)	-1.33 (-1.97,-0.68)	-1.33 (-1.98,-0.67)
Imidafenacin 0.25mg b.i.d	[37]	-1.25 (-2.06,-0.43)	-1.24 (-2.02,-0.44)	-1.26 (-2.06,-0.45)
Mirabegron IR 150mg b.i.d	[49]	-1.2 (-2.13,-0.29)	-1.19 (-2.13,-0.27)	-1.21 (-2.12,-0.27)
Solifenacin ER 10mg q.d	[30]	-1.16 (-1.41,-0.93)	-1.16 (-1.4,-0.93)	-1.16 (-1.41,-0.93)
Mirabegron IR 100mg b.i.d	[48]	-1.18 (-1.98,-0.39)	-1.17 (-1.93,-0.38)	-1.18 (-1.95,-0.41)
Electrostimulation	[80]	-1.15 (-1.86,-0.47)	-1.15 (-1.83,-0.49)	-1.16 (-1.84,-0.48)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-1.15 (-2.02,-0.31)	-1.15 (-1.98,-0.33)	-1.17 (-2.01,-0.32)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-1.11 (-1.81,-0.42)	-1.11 (-1.8,-0.43)	-1.11 (-1.81,-0.44)
Oxybutynin ER 2.5mg q.d + BT	[92]	-1.07 (-3.04,0.81)	-1.01 (-2.86,0.87)	-0.94 (-2.84,0.96)
Oxybutynin ER 2.5mg q.d	[20]	-1.05 (-2.67,0.58)	-1.01 (-2.64,0.59)	-0.97 (-2.59,0.63)
Oxybutynin vaginal ring 6mg q.d	[17]	-1 (-1.74,-0.26)	-1 (-1.72,-0.27)	-1 (-1.72,-0.27)
Fesoterodine ER 8mg q.d	[26]	-1.01 (-1.22,-0.8)	-1.01 (-1.21,-0.81)	-1.01 (-1.22,-0.8)
Solabegron IR 125mg b.i.d	[55]	-0.9 (-1.19,-0.61)	-0.9 (-1.17,-0.63)	-0.9 (-1.19,-0.61)
Mirabegron ER 50mg q.d	[51]	-0.8 (-1.01,-0.6)	-0.8 (-1,-0.61)	-0.8 (-1.01,-0.61)
Pregabalin 150mg b.i.d	[62]	-0.81 (-1.35,-0.27)	-0.81 (-1.34,-0.28)	-0.8 (-1.35,-0.27)
Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	-0.8 (-1.29,-0.32)	-0.8 (-1.28,-0.31)	-0.8 (-1.29,-0.32)
Oxybutynin IR 3mg t.i.d	[19]	-0.8 (-1.33,-0.26)	-0.79 (-1.32,-0.25)	-0.79 (-1.33,-0.26)
Solifenacin ER 5-10mg q.d	[31]	-0.79 (-1.08,-0.51)	-0.8 (-1.08,-0.52)	-0.8 (-1.08,-0.51)
Trospium chloride ER 60mg q.d	[44]	-0.77 (-1.17,-0.38)	-0.77 (-1.16,-0.38)	-0.77 (-1.16,-0.38)
Mirabegron ER 25mg q.d	[50]	-0.76 (-1.1,-0.43)	-0.76 (-1.09,-0.44)	-0.76 (-1.09,-0.43)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	-0.73 (-1.39,-0.1)	-0.73 (-1.4,-0.09)	-0.73 (-1.4,-0.09)
Fesoterodine ER 4mg q.d	[25]	-0.73 (-0.92,-0.54)	-0.73 (-0.91,-0.55)	-0.73 (-0.92,-0.55)
Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	-0.71 (-2.14,0.77)	-0.74 (-2.14,0.64)	-0.73 (-2.14,0.72)
Oxybutynin chloride topical gel 1g/day	[13]	-0.7 (-1.21,-0.19)	-0.7 (-1.19,-0.21)	-0.7 (-1.2,-0.2)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.7 (-1.4,0)	-0.69 (-1.39,-0.02)	-0.7 (-1.38,-0.01)
Fesoterodine ER 4 - 8mg q.d	[27]	-0.69 (-0.91,-0.48)	-0.7 (-0.9,-0.49)	-0.69 (-0.91,-0.48)
Solifenacin ER 5mg q.d	[29]	-0.69 (-0.92,-0.5)	-0.68 (-0.9,-0.5)	-0.69 (-0.92,-0.5)
Propiverine ER 20mg q.d	[41]	-0.69 (-0.92,-0.47)	-0.69 (-0.92,-0.47)	-0.69 (-0.92,-0.47)
Mirabegron ER 100mg q.d	[52]	-0.7 (-0.94,-0.46)	-0.7 (-0.93,-0.47)	-0.7 (-0.94,-0.46)
Oxybutynin gel 84mg/day	[134]	-0.7 (-1.35,-0.05)	-0.7 (-1.35,-0.05)	-0.7 (-1.34,-0.04)
Tolterodine IR 2mg b.i.d	[5]	-0.67 (-0.86,-0.5)	-0.67 (-0.85,-0.5)	-0.67 (-0.86,-0.5)
Tarafenacin 0.4mg q.d	[82]	-0.66 (-1.55,0.23)	-0.67 (-1.53,0.2)	-0.66 (-1.55,0.23)
Oxybutynin IR 5mg t.i.d	[7]	-0.65 (-1.09,-0.2)	-0.64 (-1.1,-0.21)	-0.64 (-1.11,-0.21)
Vaginal oestrogen cream 1.25mg/day	[132]	-0.65 (-1.5,0.17)	-0.65 (-1.46,0.15)	-0.67 (-1.49,0.15)
Tolterodine IR 1mg b.i.d	[6]	-0.63 (-1.11,-0.16)	-0.63 (-1.09,-0.16)	-0.63 (-1.1,-0.17)
Imidafenacin 0.05mg b.i.d	[35]	-0.62 (-1.26,0.02)	-0.62 (-1.25,0)	-0.63 (-1.26,0.01)
Tolterodine ER 4mg q.d	[4]	-0.62 (-0.74,-0.51)	-0.62 (-0.74,-0.51)	-0.62 (-0.74,-0.51)
Darifenacin ER 7.5mg q.d	[39]	-0.6 (-2.15,0.95)	-0.63 (-2.14,0.92)	-0.59 (-2.14,0.93)

TABLE C.9: Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters (cont.)

Trospium chloride IR 45mg t.i.d	[47]	-0.6 (-1.35,0.12)	-0.61 (-1.35,0.12)	-0.6 (-1.36,0.13)
Oxybutynin trandermal 3.9mg/day	[10]	-0.57 (-0.96,-0.18)	-0.58 (-0.96,-0.19)	-0.58 (-0.96,-0.2)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.51 (-1.05,0.03)	-0.51 (-1.05,0.02)	-0.5 (-1.05,0.04)
Imidafenacin 0.1mg b.i.d	[36]	-0.5 (-0.82,-0.18)	-0.49 (-0.81,-0.19)	-0.5 (-0.81,-0.19)
Solabegron IR 50mg b.i.d	[54]	-0.5 (-0.79,-0.21)	-0.5 (-0.77,-0.23)	-0.5 (-0.79,-0.21)
Propiverine IR 45mg t.i.d	[118]	-0.46 (-2.59,1.61)	-0.43 (-2.4,1.62)	-0.47 (-2.58,1.38)
Bladder Training (BT)/Behaviour Therapy	[85]	-0.43 (-1.42,0.49)	-0.41 (-1.41,0.53)	-0.42 (-1.43,0.52)
Darifenacin ER 15mg q.d	[40]	-0.4 (-1.25,0.45)	-0.4 (-1.24,0.43)	-0.39 (-1.22,0.46)
Serlopitant 4mg q.d	[109]	-0.41 (-0.9,0.08)	-0.41 (-0.89,0.08)	-0.41 (-0.91,0.08)
Serlopitant 0.25mg q.d	[107]	-0.41 (-0.9,0.08)	-0.41 (-0.9,0.08)	-0.41 (-0.91,0.08)
Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	-0.4 (-3.71,2.64)	-0.32 (-3,2.63)	-0.24 (-3.43,3.1)
Terodiline 25mg b.i.d	[28]	-0.34 (-0.81,0.11)	-0.35 (-0.81,0.11)	-0.34 (-0.8,0.12)
Propiverine IR 15mg b.i.d	[43]	-0.35 (-1.23,0.56)	-0.34 (-1.21,0.53)	-0.36 (-1.24,0.48)
Netupitant 200mg q.d	[112]	-0.32 (-1.47,0.85)	-0.31 (-1.48,0.81)	-0.33 (-1.52,0.85)
Oxybutynin 20mg intravesically q.d	[106]	-0.31 (-1.64,1.02)	-0.3 (-1.65,1.02)	-0.3 (-1.67,1.09)
Elocalcitol 150mg	[69]	-0.29 (-1.07,0.49)	-0.29 (-1.08,0.49)	-0.31 (-1.09,0.49)
Oxybutynin gel 56mg/day	[135]	-0.3 (-0.98,0.37)	-0.3 (-0.97,0.37)	-0.29 (-0.97,0.41)
Control	[2]	-0.27 (-1.74,1.22)	-0.22 (-1.77,1.37)	-0.26 (-1.73,1.24)
Oxybutynin ER 5-30mg q.d + BT	[86]	-0.22 (-2.32,1.87)	-0.2 (-2.26,1.85)	-0.24 (-2.4,1.85)
Oxybutynin ER 5 - 30mg/day	[22]	-0.23 (-1.47,1.04)	-0.21 (-1.48,1.02)	-0.23 (-1.51,1.01)
Cizolirtine citrate 200mg b.i.d	[56]	-0.21 (-1.7,1.26)	-0.19 (-1.69,1.31)	-0.2 (-1.68,1.36)
Netupitant 100mg q.d	[111]	-0.19 (-1.27,0.93)	-0.2 (-1.29,0.94)	-0.21 (-1.3,0.92)
Tarafenacin 0.2mg q.d	[90]	-0.15 (-1.07,0.76)	-0.16 (-1.07,0.76)	-0.15 (-1.07,0.76)
Serlopitant 1mg q.d	[108]	-0.11 (-0.61,0.38)	-0.11 (-0.6,0.37)	-0.11 (-0.6,0.38)
Elocalcitol 75mg	[70]	-0.1 (-0.85,0.64)	-0.1 (-0.84,0.65)	-0.09 (-0.84,0.65)
Oxybutynin transdermal 1.3mg/day	[11]	-0.08 (-0.73,0.58)	-0.08 (-0.73,0.55)	-0.08 (-0.73,0.55)
Oxybutynin transdermal 2.6mg/day	[12]	-0.08 (-0.7,0.55)	-0.09 (-0.7,0.54)	-0.09 (-0.71,0.53)
Netupitant 50mg q.d	[110]	-0.09 (-1.19,0.99)	-0.08 (-1.21,1.03)	-0.09 (-1.21,1.02)
Percutaneous tibial nerve stimulation	[83]	-0.08 (-1.41,1.13)	-0.11 (-1.31,1.12)	-0.13 (-1.44,1.12)
Electromagnetic stimulation	[125]	-0.07 (-2.09,1.94)	-0.08 (-2.04,1.9)	-0.12 (-2.03,1.99)
ONO-8539 100mg b.i.d	[60]	-0.05 (-0.8,0.71)	-0.04 (-0.79,0.7)	-0.05 (-0.8,0.7)
Oxybutynin ER 15mg q.d	[9]	-0.04 (-1.22,1.16)	-0.04 (-1.25,1.15)	-0.04 (-1.23,1.14)
Placebo	[1]	NA	NA	NA
Resiniferatoxin 50nM	[67]	0.06 (-1.22,1.41)	0.04 (-1.26,1.37)	0.09 (-1.23,1.44)
Oxybutynin IR 5mg b.i.d	[18]	0.12 (-0.88,1.13)	0.12 (-0.85,1.1)	0.12 (-0.88,1.1)
ONO-8539 300mg b.i.d	[61]	0.17 (-0.59,0.93)	0.17 (-0.59,0.93)	0.17 (-0.59,0.93)
Propantheline Bromide 15mg t.i.d	[113]	0.33 (-0.84,1.45)	0.33 (-0.82,1.46)	0.31 (-0.83,1.43)
Propiverine ER 60mg q.d	[119]	0.45 (-1.78,2.69)	0.46 (-1.57,2.6)	0.44 (-1.72,2.54)
ONO-8539 30mg b.i.d	[59]	0.47 (-0.27,1.19)	0.46 (-0.28,1.2)	0.47 (-0.29,1.2)
Estradiol 1mg intravaginally	[127]	0.49 (-1.18,2.17)	0.5 (-1.15,2.12)	0.49 (-1.2,2.15)
ZD0947IL 25mg/day	[58]	0.71 (-0.58,2.03)	0.69 (-0.62,2.03)	0.72 (-0.58,2.04)
Sham therapy	[3]	0.94 (-0.53,2.35)	0.93 (-0.44,2.32)	0.9 (-0.55,2.31)
Naftopidil 25mg q.d	[114]	3.24 (0.7,5.57)	3.14 (0.61,5.69)	3.18 (0.56,5.67)
Trospium chloride IR 15mg t.i.d	[46]	4.59 (1.8,7.32)	4.63 (2.15,7.1)	4.72 (2.11,7.79)
Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	5.19 (2.57,7.7)	5.16 (2.41,7.75)	5.13 (2.46,7.81)

† Between-study standard deviation based on an Uniform(0,5) prior distribution on the standard deviation scale

†† Between-study standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Between-study standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE C.10: Sensitivity analyses assessing the impact in change from baseline in urgency episodes for different choices of prior distribution on variance parameters

Treatment	**Code**	**Median difference† (95%CrI)**	**Median difference†† (95%CrI)**	**Median difference††† (95%CrI)**
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-6.94 (-8.65,-5.23)	-6.94 (-8.62,-5.29)	-6.93 (-8.64,-5.25)
Electrostimulation	[80]	-4.84 (-5.94,-3.74)	-4.84 (-5.93,-3.76)	-4.83 (-5.94,-3.74)
Vaginal oestrogen cream 1.25mg/day	[132]	-3.34 (-5.06,-1.62)	-3.34 (-5.04,-1.67)	-3.33 (-5.06,-1.64)
OnaBoNT-A 100u bladder body + trigone	[78]	-2.69 (-4.83,-0.54)	-2.69 (-4.8,-0.56)	-2.69 (-4.84,-0.52)
OnaBoNT-A 100u bladder base + trigone	[79]	-2.56 (-5.14,0.04)	-2.58 (-5.13,-0.02)	-2.57 (-5.16,0.03)
Cizolirtine citrate 400mg b.i.d	[57]	-2.62 (-4.66,-0.59)	-2.64 (-4.67,-0.62)	-2.63 (-4.66,-0.58)
Percutaneous tibial nerve stimulation	[83]	-2.5 (-4.37,-0.64)	-2.5 (-4.32,-0.7)	-2.48 (-4.34,-0.67)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-2.38 (-3.82,-0.94)	-2.37 (-3.77,-0.98)	-2.37 (-3.79,-0.96)
OnaBoNT-A 200U trigone sparing	[73]	-2.25 (-3.98,-0.51)	-2.25 (-3.96,-0.53)	-2.25 (-3.98,-0.53)
OnaBoNT-A 100u trigone sparing	[72]	-2.07 (-3.06,-1.08)	-2.07 (-3.04,-1.11)	-2.07 (-3.05,-1.08)
Oxybutynin IR 2.5mg t.i.d	[21]	-1.94 (-3.02,-0.82)	-1.96 (-3.02,-0.85)	-1.94 (-3.02,-0.82)
Darifenacin ER 7.5mg q.d	[39]	-1.68 (-3.34,-0.02)	-1.68 (-3.32,-0.06)	-1.67 (-3.34,-0.03)
Imidafenacin 0.25mg b.i.d	[37]	-1.48 (-3.08,0.12)	-1.48 (-3.05,0.08)	-1.48 (-3.07,0.1)
Fesoterodine ER 8mg q.d	[26]	-1.47 (-2.03,-0.9)	-1.47 (-2.01,-0.93)	-1.47 (-2.02,-0.91)
Lipo-BoNTA 200U	[138]	-1.33 (-3.59,0.94)	-1.33 (-3.56,0.94)	-1.33 (-3.59,0.95)
Solifenacin ER 10mg q.d	[30]	-1.34 (-2.02,-0.69)	-1.34 (-1.99,-0.7)	-1.34 (-2,-0.69)
Fesoterodine ER 4 - 8mg q.d	[27]	-1.23 (-2.01,-0.46)	-1.23 (-1.99,-0.47)	-1.23 (-2,-0.46)
Mirabegron IR 150mg b.i.d	[49]	-1.15 (-2.8,0.5)	-1.14 (-2.77,0.49)	-1.13 (-2.78,0.51)
Mirabegron IR 100mg b.i.d	[48]	-1.14 (-2.64,0.37)	-1.14 (-2.62,0.35)	-1.13 (-2.63,0.37)
Solifenacin ER 5mg q.d	[29]	-1.07 (-1.7,-0.49)	-1.07 (-1.68,-0.49)	-1.07 (-1.68,-0.49)
Oxybutynin trandermal 3.9mg/day	[10]	-1.02 (-2.38,0.34)	-1.02 (-2.35,0.3)	-1.01 (-2.39,0.34)
Fesoterodine ER 4mg q.d	[25]	-1.05 (-1.55,-0.53)	-1.05 (-1.54,-0.55)	-1.05 (-1.55,-0.54)
Solifenacin ER 5 - 10mg q.d	[31]	-0.9 (-1.68,-0.13)	-0.9 (-1.66,-0.16)	-0.9 (-1.68,-0.14)
Propiverine ER 20mg q.d	[41]	-0.9 (-1.58,-0.23)	-0.9 (-1.55,-0.24)	-0.9 (-1.56,-0.23)
Propiverine IR 15mg b.i.d	[43]	-0.88 (-2.79,0.99)	-0.87 (-2.73,0.96)	-0.88 (-2.77,1)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.77 (-2.42,0.85)	-0.76 (-2.37,0.82)	-0.75 (-2.39,0.86)
Pregabalin 150mg b.i.d	[62]	-0.78 (-2.02,0.45)	-0.78 (-1.97,0.41)	-0.78 (-1.99,0.43)
Tolterodine ER 4mg q.d	[4]	-0.77 (-1.16,-0.39)	-0.77 (-1.15,-0.4)	-0.77 (-1.16,-0.4)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.79 (-2.2,0.62)	-0.78 (-2.14,0.58)	-0.78 (-2.16,0.62)
Tarafenacin 0.4mg q.d	[82]	-0.74 (-2.53,1.02)	-0.74 (-2.47,1)	-0.74 (-2.5,1.03)
Elocalcitol 75mg	[70]	-0.7 (-2.17,0.77)	-0.7 (-2.14,0.73)	-0.7 (-2.16,0.77)
Imidafenacin 0.05mg b.i.d	[35]	-0.69 (-2.11,0.72)	-0.69 (-2.08,0.68)	-0.69 (-2.09,0.7)
Imidafenacin 0.1mg b.i.d	[36]	-0.72 (-1.58,0.13)	-0.71 (-1.55,0.11)	-0.71 (-1.57,0.12)
Mirabegron ER 100mg q.d	[52]	-0.69 (-1.4,0.01)	-0.69 (-1.38,-0.02)	-0.69 (-1.39,0)
Cizolirtine citrate 200mg b.i.d	[56]	-0.65 (-2.48,1.16)	-0.66 (-2.44,1.12)	-0.65 (-2.44,1.14)
Mirabegron ER 50mg q.d	[51]	-0.67 (-1.24,-0.1)	-0.67 (-1.23,-0.12)	-0.67 (-1.24,-0.11)
Tolterodine IR 2mg b.i.d	[5]	-0.66 (-1.64,0.29)	-0.66 (-1.61,0.26)	-0.65 (-1.62,0.29)
Netupitant 100mg q.d	[111]	-0.54 (-1.93,0.86)	-0.54 (-1.89,0.8)	-0.54 (-1.92,0.82)
Elocalcitol 150mg	[69]	-0.4 (-1.86,1.07)	-0.4 (-1.83,1.02)	-0.4 (-1.85,1.06)
Mirabegron ER 25mg q.d	[50]	-0.41 (-1.37,0.54)	-0.42 (-1.34,0.51)	-0.41 (-1.36,0.54)
Oxybutynin ER 2.5mg q.d + BT	[92]	-0.38 (-3.06,2.27)	-0.36 (-2.99,2.21)	-0.36 (-3,2.27)
Tarafenacin 0.2mg q.d	[90]	-0.39 (-2.14,1.34)	-0.39 (-2.1,1.32)	-0.39 (-2.13,1.35)
Tolterodine IR 2mg b.i.d + PFMT	[95]	-0.36 (-2.08,1.33)	-0.36 (-2.03,1.27)	-0.35 (-2.07,1.34)
Darifenacin ER 15mg q.d.	[40]	-0.37 (-1.82,1.09)	-0.36 (-1.78,1.05)	-0.37 (-1.81,1.08)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.38 (-1.69,0.93)	-0.38 (-1.66,0.88)	-0.38 (-1.67,0.9)
Oxybutynin ER 2.5mg q.d	[20]	-0.19 (-2.82,2.46)	-0.17 (-2.75,2.41)	-0.17 (-2.77,2.46)
Bladder Training (BT)/Behaviour Therapy	[85]	-0.17 (-1.84,1.5)	-0.15 (-1.79,1.45)	-0.15 (-1.82,1.5)
ONO-8539 300mg b.i.d	[61]	-0.19 (-1.65,1.25)	-0.18 (-1.61,1.25)	-0.18 (-1.62,1.26)
Netupitant 50mg q.d	[110]	-0.17 (-1.55,1.22)	-0.17 (-1.51,1.16)	-0.17 (-1.54,1.2)
ZD0947IL 25mg/day	[58]	-0.09 (-2.04,1.87)	-0.09 (-2.01,1.86)	-0.09 (-2.04,1.88)
Netupitant 200mg q.d	[112]	-0.08 (-1.47,1.31)	-0.08 (-1.42,1.26)	-0.08 (-1.46,1.29)
Placebo	[1]	NA	NA	NA
ONO-8539 100mg b.i.d	[60]	0.06 (-1.39,1.51)	0.07 (-1.37,1.49)	0.07 (-1.38,1.51)
ONO-8539 30mg b.i.d	[59]	0.87 (-0.58,2.3)	0.88 (-0.54,2.29)	0.88 (-0.55,2.31)

† Between-study standard deviation based on an Uniform(0,5) prior distribution on the standard deviation scale

†† Between-study standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Between-study standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE C.11: Sensitivity analyses assessing the impact in change from baseline in nocturia episodes for different choices of prior distribution on variance parameters

Treatment	**Code**	**Median difference† (95%CrI)**	**Median difference†† (95%CrI)**	**Median difference††† (95%CrI)**
Oxybutynin 20mg intravesically q.d	[106]	-2.61 (-3.84,-1.39)	-2.59 (-3.94,-1.35)	-2.59 (-3.68,-1.44)
Estriol 1mg intravesically	[131]	-2.29 (-3.41,-1.32)	-2.33 (-3.5,-1.18)	-2.32 (-3.28,-1.27)
Estradiol 3mg intravaginally	[128]	-1.8 (-2.98,-0.74)	-1.89 (-2.89,-0.81)	-1.84 (-2.95,-0.83)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.64 (-1.25,0)	-0.64 (-1.3,-0.01)	-0.64 (-1.28,-0.01)
Percutaneous tibial nerve stimulation	[83]	-0.44 (-0.99,0.17)	-0.44 (-1.05,0.14)	-0.45 (-1.03,0.07)
Bladder Training (BT)/Behaviour Therapy	[85]	-0.41 (-0.69,-0.15)	-0.42 (-0.71,-0.14)	-0.42 (-0.71,-0.15)
Electrostimulation	[80]	-0.41 (-0.75,-0.06)	-0.39 (-0.73,-0.05)	-0.41 (-0.78,-0.04)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.41 (-0.85,0.04)	-0.39 (-0.8,0.02)	-0.4 (-0.86,0.03)
Mirabegron IR 100mg b.i.d	[48]	-0.36 (-0.7,-0.06)	-0.35 (-0.66,-0.04)	-0.36 (-0.68,-0.06)
Tolterodine ER 4mg q.d + BT	[87]	-0.35 (-0.64,-0.09)	-0.37 (-0.63,-0.09)	-0.36 (-0.64,-0.07)
Imidafenacin 0.1mg b.i.d	[36]	-0.32 (-0.66,0.01)	-0.34 (-0.7,0)	-0.33 (-0.69,0.02)
Estradiol 1mg intravaginally	[127]	-0.32 (-1.77,1.04)	-0.35 (-1.49,0.88)	-0.33 (-1.6,1.03)
OnaBoNT-A 100u trigone sparing	[72]	-0.26 (-0.41,-0.1)	-0.26 (-0.41,-0.11)	-0.26 (-0.41,-0.11)
Oxybutynin ER 2.5mg q.d	[20]	-0.26 (-0.51,0.01)	-0.26 (-0.53,0.03)	-0.27 (-0.54,0.01)
Fesoterodine ER 4 - 8mg q.d	[27]	-0.24 (-0.35,-0.13)	-0.24 (-0.34,-0.13)	-0.24 (-0.35,-0.13)
Mirabegron ER 25mg q.d	[50]	-0.25 (-0.43,-0.06)	-0.24 (-0.42,-0.06)	-0.24 (-0.43,-0.05)
Oxybutynin ER 2.5mg q.d + BT	[92]	-0.22 (-0.72,0.38)	-0.23 (-0.82,0.33)	-0.24 (-0.8,0.36)
Mirabegron ER 50mg q.d	[51]	-0.15 (-0.23,-0.07)	-0.15 (-0.23,-0.06)	-0.15 (-0.23,-0.06)
Solifenacin ER 10mg q.d	[30]	-0.15 (-0.28,-0.02)	-0.15 (-0.28,-0.02)	-0.15 (-0.3,-0.01)
Mirabegron IR 150mg b.i.d	[49]	-0.14 (-0.48,0.17)	-0.14 (-0.5,0.17)	-0.14 (-0.46,0.18)
Solifenacin ER 5mg q.d	[29]	-0.14 (-0.28,0)	-0.14 (-0.27,0)	-0.14 (-0.29,0)
Tolterodine ER 4mg q.d	[4]	-0.13 (-0.2,-0.07)	-0.13 (-0.2,-0.07)	-0.13 (-0.2,-0.06)
Mirabegron ER 100mg q.d	[52]	-0.13 (-0.22,-0.02)	-0.12 (-0.22,-0.03)	-0.12 (-0.22,-0.02)
Fesoterodine ER 4mg q.d	[25]	-0.13 (-0.2,-0.06)	-0.13 (-0.2,-0.06)	-0.13 (-0.21,-0.05)
Fesoterodine ER 8mg q.d	[26]	-0.13 (-0.22,-0.05)	-0.13 (-0.22,-0.05)	-0.13 (-0.22,-0.05)
Tarafenacin 0.4mg q.d	[82]	-0.13 (-0.45,0.22)	-0.11 (-0.43,0.21)	-0.12 (-0.45,0.22)
Tolterodine IR 2mg b.i.d	[5]	-0.13 (-0.42,0.1)	-0.13 (-0.44,0.13)	-0.14 (-0.43,0.1)
Oxybutynin trandermal 3.9mg/day	[10]	-0.12 (-0.4,0.16)	-0.11 (-0.4,0.18)	-0.11 (-0.41,0.18)
Oxbutynin patch 73.5mg	[15]	-0.1 (-0.22,0.01)	-0.1 (-0.22,0.02)	-0.1 (-0.23,0.03)
Oxybutynin chloride topical gel 1g/day	[13]	-0.1 (-0.3,0.1)	-0.1 (-0.3,0.09)	-0.1 (-0.3,0.11)
Propiverine ER 20mg q.d	[41]	-0.07 (-0.15,0)	-0.07 (-0.15,0.01)	-0.07 (-0.16,0.01)
Resiniferatoxin 50nM	[67]	-0.06 (-0.62,0.58)	-0.06 (-0.59,0.52)	-0.04 (-0.63,0.55)
Oxybutynin ER 5 - 30mg/day	[22]	-0.02 (-0.48,0.47)	-0.06 (-0.48,0.4)	-0.04 (-0.5,0.43)
Sham therapy	[3]	-0.04 (-0.74,0.62)	-0.05 (-0.77,0.65)	-0.06 (-0.73,0.58)
Placebo	[1]	NA	NA	NA
Oxybutynin IR 2.5mg t.i.d	[21]	0.13 (-0.15,0.4)	0.12 (-0.15,0.4)	0.12 (-0.16,0.4)
Terodiline 25mg b.i.d	[28]	0.14 (-0.04,0.31)	0.14 (-0.04,0.31)	0.14 (-0.04,0.31)
Tarafenacin 0.2mg q.d	[90]	0.14 (-0.16,0.46)	0.14 (-0.16,0.44)	0.14 (-0.18,0.46)
Darifenacin ER 7.5mg q.d	[39]	0.29 (-0.34,0.91)	0.27 (-0.37,0.9)	0.26 (-0.37,0.89)
Vaginal oestrogen cream 1.25mg/day	[132]	0.3 (-0.12,0.73)	0.32 (-0.08,0.75)	0.29 (-0.15,0.73)

† Between-study standard deviation based on an Uniform(0,5) prior distribution on the standard deviation scale

†† Between-study standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Between-study standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

C.13 Sensitivity to potentially biased studies

TABLE C.12: Sensitivity analysis assessing the impact in change from baseline in urgency episodes having excluded studies A157 and A158

Treatment	**Code**	**Median difference (95% CrI)**	**Rank (95% CrI)**	**p(Best)**
Cizolirtine citrate 400mg b.i.d	[57]	-2.64 (-4.36,-0.9)	3 (1,22)	0.25
OnaBoNT-A 100u bladder base + trigone	[79]	-2.55 (-4.8,-0.34)	3 (1,37)	0.25
OnaBoNT-A 100u bladder body + trigone	[78]	-2.67 (-4.4,-0.97)	3 (1,20)	0.22
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-2.31 (-3.21,-1.4)	4 (1,12)	0.06
OnaBoNT-A 200U trigone sparing	[73]	-2.26 (-3.61,-0.89)	5 (1,22)	0.09
OnaBoNT-A 100u trigone sparing	[72]	-2.07 (-2.7,-1.42)	6 (2,12)	0.00
Darifenacin ER 7.5mg q.d	[39]	-1.55 (-2.74,-0.39)	10 (2,36)	0.01
Fesoterodine ER 8mg q.d	[26]	-1.5 (-1.83,-1.15)	10 (6,17)	0.00
Imidafenacin 0.25mg b.i.d	[37]	-1.44 (-2.69,-0.18)	11 (3,40)	0.01
Lipo-BoNTA 200U	[138]	-1.35 (-3.33,0.66)	13 (1,49)	0.03
Solifenacin ER 10mg q.d	[30]	-1.3 (-1.71,-0.9)	13 (7,22)	0.00
Fesoterodine ER 4 - 8mg q.d	[27]	-1.22 (-1.69,-0.75)	15 (8,27)	0.00
Mirabegron IR 150mg b.i.d	[49]	-1.11 (-2.45,0.23)	17 (4,47)	0.00
Mirabegron IR 100mg b.i.d	[48]	-1.11 (-2.27,0.07)	17 (4,45)	0.00
Fesoterodine ER 4mg q.d	[25]	-1.07 (-1.37,-0.76)	18 (11,28)	0.00
Oxybutynin trandermal 3.9mg/day	[10]	-1.02 (-1.87,-0.17)	19 (7,41)	0.00
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.92 (-4.13,2.44)	21 (1,51)	0.07
Solifenacin ER 5mg q.d	[29]	-0.95 (-1.34,-0.59)	21 (13,32)	0.00
Propiverine ER 20mg q.d	[41]	-0.87 (-1.28,-0.47)	23 (13,35)	0.00
Solifenacin ER 5 - 10mg q.d	[31]	-0.84 (-1.29,-0.4)	24 (13,37)	0.00
Propiverine IR 15mg b.i.d	[43]	-0.79 (-2.17,0.56)	25 (5,49)	0.00
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.76 (-1.79,0.29)	26 (8,47)	0.00
Pregabalin 150mg b.i.d	[62]	-0.75 (-1.54,0.03)	26 (10,44)	0.00
Tarafenacin 0.4mg q.d	[82]	-0.74 (-2.15,0.66)	27 (5,50)	0.00
Tolterodine IR 2mg b.i.d + BT	[93]	-0.68 (-1.7,0.32)	28 (9,47)	0.00
Elocalcitol 75mg	[70]	-0.71 (-1.72,0.31)	28 (8,48)	0.00
Tolterodine ER 4mg q.d	[4]	-0.7 (-0.94,-0.47)	28 (20,36)	0.00
Cizolirtine citrate 200mg b.i.d	[56]	-0.65 (-2.12,0.81)	29 (5,50)	0.00
Mirabegron ER 100mg q.d	[52]	-0.67 (-1.09,-0.26)	29 (17,40)	0.00
Mirabegron ER 50mg q.d	[51]	-0.66 (-0.99,-0.32)	29 (19,39)	0.00
Imidafenacin 0.05mg b.i.d	[35]	-0.64 (-1.67,0.37)	30 (9,48)	0.00
Imidafenacin 0.1mg b.i.d	[36]	-0.63 (-1.21,-0.07)	30 (15,43)	0.00
Tolterodine IR 2mg b.i.d	[5]	-0.58 (-1.2,0.04)	31 (17,43)	0.00
Netupitant 100mg q.d	[111]	-0.53 (-1.41,0.34)	33 (11,48)	0.00
Tarafenacin 0.2mg q.d	[90]	-0.4 (-1.77,0.97)	36 (8,51)	0.00
Elocalcitol 150mg	[69]	-0.4 (-1.4,0.6)	36 (12,50)	0.00
Mirabegron ER 25mg q.d	[50]	-0.4 (-0.98,0.17)	36 (20,47)	0.00
Darifenacin ER 15mg q.d.	[40]	-0.37 (-1.35,0.62)	37 (12,50)	0.00
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.35 (-1.26,0.54)	37 (14,50)	0.00
Oxybutynin ER 2.5mg q.d + BT	[92]	-0.27 (-2.34,1.79)	39 (4,52)	0.00
Tolterodine IR 2mg b.i.d + PFMT	[95]	-0.27 (-1.38,0.82)	39 (13,51)	0.00
ZD0947IL 25mg/day	[58]	-0.12 (-1.76,1.53)	41 (8,52)	0.00
ONO-8539 300mg b.i.d	[61]	-0.14 (-1.25,0.97)	41 (14,51)	0.00
Netupitant 50mg q.d	[110]	-0.17 (-1.03,0.7)	41 (19,50)	0.00
Oxybutynin ER 2.5mg q.d	[20]	-0.08 (-2.14,1.95)	42 (6,53)	0.00
Bladder Training (BT)/Behaviour Therapy	[85]	-0.09 (-1.16,0.96)	42 (18,51)	0.00
Netupitant 200mg q.d	[112]	-0.07 (-0.95,0.8)	42 (21,51)	0.00
Percutaneous tibial nerve stimulation	[83]	-0.03 (-2.04,2)	43 (7,52)	0.00
Placebo	[1]	NA	44 (38,49)	0.00
ONO-8539 100mg b.i.d	[60]	0.11 (-0.99,1.21)	45 (20,52)	0.00
Electrostimulation	[80]	1.18 (-1.94,4.47)	51 (8,52)	0.00
ONO-8539 30mg b.i.d	[59]	0.92 (-0.17,2.03)	51 (41,53)	0.00
Vaginal oestrogen cream 1.25mg/day	[132]	2.68 (-0.55,6.06)	53 (35,53)	0.00

C.14 Sensitivity to the inclusion of sacral nerve stimulation

TABLE C.13: Sensitivity analysis assessing the impact in change from baseline in incontinence episodes having excluded sacral nerve stimulation from the treatment network

Treatment	**Code**	**Mean difference (95% CrI)**	**Mean difference† (95% CrI)**
Sacral nerve stimulation	[81]	-8.72 (-11.33,-6.09)	NA
OnaBoNT-A 200U trigone sparing	[73]	-2.3 (-3.16,-1.42)	-2.29 (-3.15,-1.38)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-2.2 (-4.06,-0.36)	-2.29 (-4.16,-0.37)
Solifenacin/trospium + placebo injection	[100]	-1.97 (-2.96,-1.01)	-1.98 (-2.96,-0.98)
Electrostimulation + PFMT + BT	[97]	-1.93 (-2.94,-0.91)	-1.9 (-2.92,-0.88)
OnaBoNT-A 100u trigone sparing	[72]	-1.88 (-2.31,-1.45)	-1.88 (-2.32,-1.45)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.63 (-2.73,-0.54)	-1.66 (-2.75,-0.58)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.39 (-4,1.06)	-1.46 (-3.95,1.08)
Tolterodine ER 4mg q.d + Neurostimulation	[96]	-1.29 (-1.7,-0.89)	-1.3 (-1.7,-0.89)
Oxybutynin intravesically 5mg t.i.d	[14]	-1.19 (-2.49,0.1)	-1.21 (-2.48,0.08)
Mirabegron IR 100mg b.i.d	[48]	-1.12 (-1.98,-0.25)	-1.11 (-1.97,-0.24)
Oxybutynin ER 10mg q.d	[8]	-0.98 (-1.55,-0.42)	-0.99 (-1.55,-0.42)
Oxybutynin IR 3mg t.i.d	[19]	-0.9 (-1.28,-0.52)	-0.9 (-1.28,-0.52)
Solifenacin ER 10mg q.d	[30]	-0.88 (-1.14,-0.63)	-0.89 (-1.14,-0.63)
Imidafenacin 0.25mg b.i.d	[37]	-0.82 (-1.43,-0.21)	-0.83 (-1.43,-0.21)
Propiverine ER 30mg q.d	[42]	-0.79 (-1.31,-0.28)	-0.79 (-1.31,-0.28)
Tolterodine ER 4mg q.d + BT	[87]	-0.78 (-1.45,-0.12)	-0.78 (-1.44,-0.11)
Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	-0.78 (-1.71,0.16)	-0.77 (-1.68,0.18)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	-0.75 (-1.45,-0.01)	-0.75 (-1.48,-0.02)
Darifenacin ER 30mg q.d	[38]	-0.74 (-1.36,-0.12)	-0.74 (-1.36,-0.12)
Tolterodine ER 2mg b.i.d + Estrogen 0.625mg 2xwk	[99]	-0.75 (-1.24,-0.25)	-0.74 (-1.24,-0.24)
Solifenacin ER 5 - 10mg q.d	[31]	-0.73 (-1,-0.45)	-0.73 (-1,-0.45)
Solifenacin ER 5mg q.d	[29]	-0.71 (-0.96,-0.46)	-0.71 (-0.96,-0.46)
Oxybutynin IR 5mg t.i.d	[7]	-0.71 (-1.18,-0.25)	-0.71 (-1.18,-0.25)
Fesoterodine ER 8mg q.d	[26]	-0.69 (-0.89,-0.5)	-0.69 (-0.89,-0.5)
Oxybutynin gel 56mg/day	[135]	-0.69 (-1.59,0.18)	-0.7 (-1.59,0.17)
Mirabegron ER 25mg q.d	[50]	-0.67 (-0.99,-0.35)	-0.67 (-0.99,-0.36)
Terodiline 25mg b.i.d	[28]	-0.66 (-1.21,-0.1)	-0.67 (-1.22,-0.12)
Trospium chloride ER 60mg q.d	[44]	-0.66 (-1.09,-0.24)	-0.66 (-1.08,-0.24)
Cizilirtine citrate 400mg b.i.d	[57]	-0.63 (-1.17,-0.08)	-0.63 (-1.18,-0.08)
Mirabegron ER 100mg q.d	[52]	-0.62 (-0.83,-0.41)	-0.62 (-0.83,-0.41)
Solifenacin ER 5 - 15mg q.d	[34]	-0.61 (-1.11,-0.1)	-0.61 (-1.12,-0.1)
Solabegron IR 125mg b.i.d	[55]	-0.6 (-0.94,-0.26)	-0.6 (-0.94,-0.26)
Elocalcitol 75mg	[70]	-0.6 (-1.2,0.01)	-0.6 (-1.2,0.01)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.6 (-1.66,0.46)	-0.59 (-1.69,0.47)
Propiverine IR 15mg b.i.d	[43]	-0.57 (-1,-0.16)	-0.58 (-1.01,-0.16)
Mirabegron ER 200mg q.d	[53]	-0.57 (-1.13,-0.01)	-0.57 (-1.13,0)
Mirabegron ER 50mg q.d	[51]	-0.57 (-0.76,-0.38)	-0.57 (-0.76,-0.38)
Darifenacin ER 15mg q.d.	[40]	-0.51 (-0.97,-0.05)	-0.52 (-0.97,-0.06)
Mirabegron IR 150mg b.i.d	[49]	-0.51 (-1.57,0.54)	-0.52 (-1.56,0.56)
Oxybutynin chloride topical gel 1g/day	[13]	-0.5 (-1.02,0.02)	-0.5 (-1.02,0.02)
Tolterodine ER 4mg q.d	[4]	-0.51 (-0.61,-0.4)	-0.5 (-0.61,-0.4)
Fesoterodine ER 4mg q.d	[25]	-0.47 (-0.66,-0.29)	-0.47 (-0.66,-0.28)
Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	-0.48 (-0.79,-0.17)	-0.48 (-0.79,-0.16)
Tolterodine IR 2mg b.i.d	[5]	-0.45 (-0.6,-0.3)	-0.45 (-0.6,-0.3)
Tolterodine IR 2mg b.i.d + PFMT	[95]	-0.45 (-1.07,0.18)	-0.45 (-1.08,0.19)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.44 (-1.11,0.23)	-0.42 (-1.09,0.24)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.43 (-1.08,0.22)	-0.41 (-1.06,0.23)
Propiverine ER 20mg q.d	[41]	-0.41 (-0.61,-0.2)	-0.41 (-0.62,-0.2)
Imidafenacin 0.1mg b.i.d	[36]	-0.4 (-0.7,-0.11)	-0.4 (-0.7,-0.11)
Cizilirtine citrate 200mg b.i.d	[56]	-0.4 (-1.83,1.03)	-0.39 (-1.89,1.09)
Elocalcitol 150mg	[69]	-0.4 (-1.03,0.22)	-0.4 (-1.04,0.23)
Oxybutynin trandermal 3.9mg/day	[10]	-0.33 (-0.67,0)	-0.33 (-0.67,0)
Tolterodine IR 1mg b.i.d	[6]	-0.33 (-0.69,0.02)	-0.33 (-0.69,0.03)
Oxbutynin patch 73.5mg	[15]	-0.31 (-0.67,0.05)	-0.31 (-0.67,0.05)

TABLE C.13: Sensitivity analysis assessing the impact in change from baseline in incontinence episodes having excluded sacral nerve stimulation from the treatment network (cont.)

Fesoterodine ER 4 - 8mg q.d	[27]	-0.28 (-0.52,-0.05)	-0.29 (-0.53,-0.05)
Oxybutynin gel 84mg/day	[134]	-0.29 (-1.13,0.52)	-0.3 (-1.13,0.52)
Darifenacin ER 7.5mg q.d	[39]	-0.27 (-0.84,0.3)	-0.28 (-0.84,0.31)
Oxybutynin ER 15mg q.d	[9]	-0.28 (-1.41,0.86)	-0.3 (-1.38,0.83)
Duloxetine 40mg b.i.d	[65]	-0.26 (-0.75,0.25)	-0.26 (-0.76,0.25)
PFMT + BT	[89]	-0.26 (-0.87,0.37)	-0.23 (-0.85,0.39)
Imidafenacin 0.05mg b.i.d	[35]	-0.24 (-0.77,0.27)	-0.25 (-0.77,0.28)
Solabegron IR 50mg b.i.d	[54]	-0.2 (-0.54,0.14)	-0.2 (-0.54,0.14)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.14 (-0.81,0.52)	-0.15 (-0.81,0.52)
Tarafenacin 0.4mg q.d	[82]	-0.13 (-0.85,0.61)	-0.13 (-0.84,0.59)
ONO-8539 100mg b.i.d	[60]	-0.11 (-0.85,0.63)	-0.1 (-0.85,0.64)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.08 (-0.39,0.23)	-0.08 (-0.39,0.23)
ZD0947IL 25mg/day	[58]	-0.1 (-0.94,0.76)	-0.1 (-0.96,0.75)
Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	-0.08 (-0.72,0.55)	-0.06 (-0.69,0.58)
Lipo-BoNTA 200U	[138]	-0.06 (-0.96,0.83)	-0.06 (-0.95,0.83)
Oxybutynin transdermal 1.3mg/day	[11]	-0.03 (-0.66,0.61)	-0.03 (-0.68,0.62)
Placebo	[1]	NA	NA
Estradiol 25mg	[68]	0 (-0.38,0.38)	0 (-0.38,0.38)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	0 (-0.57,0.56)	0 (-0.56,0.56)
Pregabalin 150mg b.i.d	[62]	0 (-0.88,0.9)	0.01 (-0.89,0.9)
Bladder Training (BT)/Behaviour Therapy	[85]	0.02 (-0.59,0.62)	0.04 (-0.56,0.64)
Oxybutynin ER 5 - 30mg/day	[22]	0.09 (-0.62,0.78)	0.11 (-0.59,0.81)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	0.07 (-0.45,0.59)	0.06 (-0.46,0.59)
Electrostimulation	[80]	0.08 (-0.28,0.44)	0.07 (-0.28,0.43)
Tarafenacin 0.2mg q.d	[90]	0.11 (-0.63,0.86)	0.12 (-0.62,0.85)
Percutaneous tibial nerve stimulation	[83]	0.18 (-1.16,1.54)	0.18 (-1.2,1.55)
Oxybutynin ER 2.5mg q.d + BT	[92]	0.22 (-0.94,1.4)	0.24 (-0.93,1.42)
Oxybutynin transdermal 2.6mg/day	[12]	0.29 (-0.37,0.95)	0.3 (-0.36,0.95)
Oxybutynin IR 5mg b.i.d	[18]	0.35 (-0.3,0.99)	0.36 (-0.3,1.01)
Control	[2]	0.33 (-0.67,1.31)	0.36 (-0.61,1.31)
Emepronium bromide ER 200mg q.d	[63]	0.34 (-0.33,1.01)	0.34 (-0.34,1.01)
Flavoxate chloride 200mg q.d	[64]	0.33 (-0.35,1.01)	0.33 (-0.35,1.02)
Reflexology	[71]	0.33 (-0.76,1.41)	0.36 (-0.71,1.41)
Vaginal oestrogen cream 1.25mg/day	[132]	0.38 (-0.15,0.91)	0.37 (-0.15,0.9)
Oxybutynin IR 5 - 20mg	[23]	0.46 (-0.9,1.84)	0.49 (-0.92,1.89)
Oxybutynin ER 2.5mg q.d	[20]	0.52 (-0.5,1.55)	0.54 (-0.48,1.52)
Trospium chloride IR 15mg t.i.d	[46]	0.52 (-0.41,1.41)	0.55 (-0.37,1.47)
ONO-8539 300mg b.i.d	[61]	0.51 (-0.23,1.25)	0.52 (-0.22,1.27)
ONO-8539 30mg b.i.d	[59]	0.58 (-0.18,1.33)	0.59 (-0.17,1.34)
Resiniferatoxin 50nM	[67]	0.58 (-1.08,2.28)	0.6 (-1.08,2.24)
Sham therapy	[3]	0.83 (-0.62,2.27)	0.83 (-0.66,2.29)
Oxybutynin ER 5-30mg q.d + BT	[86]	0.92 (-0.31,2.12)	0.92 (-0.31,2.17)

† Results obtained from unadjusted random effects network meta-analysis excluding sacral nerve stimulation from the network of treatment comparisons

Appendix D

Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder

D.1 Example WinBUGS code for hierarchical NMA of continuous outcomes

```
Model{

for(i in 1:ns){
        for (k in 1:na[i]){

change_var[i,k] <- ((pow(b_sd_star[i,k],2) + pow(f_sd_star[i,k],2) -
2*rho[i,k]*b_sd_star[i,k]*f_sd_star[i,k])*equals(ind_c_miss[i,k],1)) +
(pow(sd[i,k],2)*(equals(ind_c_miss[i,k],0)))

change_sd[i,k] <- sqrt(change_var[i,k])
se[i,k] <- change_sd[i,k]/sqrt(numinclanalysis[i,k])

f_sd_star[i,k] <- (-0.011+(0.835*b_sd[i,k]))*(equals(ind_f_miss[i,k],1)) +
(f_sd[i,k])*(equals(ind_f_miss[i,k],0))

b_sd_star[i,k] <- 0*equals(ind_b_miss[i,k],1) +
(b_sd[i,k])*(equals(ind_b_miss[i,k],0))

b_sd[i,k] ~ dunif(0,15)
f_sd[i,k] ~ dunif(0, 15)
sd[i,k] ~ dunif(0,25)

z[i,k]~ dnorm(z.star[i,k],z.prec)
z.star[i,k] ~ dnorm(0.67, 12.76)
rho[i,k]<- (exp(2*z[i,k])-1)/(exp(2*z[i,k])+1)

} }

z.se <- 1/(sqrt(49-3))
z.prec<- pow(z.se,-2)

for(i in 1:ns){

        w[i,1] <- 0
        delta[i,1] <- 0
        mu[i] ~ dnorm(0,0.001)

for (k in 1:na[i]) {
        var[i,k] <- pow(se[i,k],2)
        prec[i,k] <- 1/var[i,k]
        y[i,k] ~ dnorm(theta[i,k],prec[i,k])
        theta[i,k] <- mu[i] + delta[i,k]

dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
}
```

```
resdev[i] <- sum(dev[i,1:na[i]])

        for (k in 2:na[i]) {
                delta[i,k] ~ dnorm(md[i,k],taud[i,k])
                md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]
                taud[i,k] <- tau *2*(k-1)/k
                w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
                sw[i,k] <- sum(w[i,1:k-1])/(k-1)
                }
}

totresdev <- sum(resdev[])

d[1]<-0                                         # Set reference intervention to be zero
                                                # Hierarchical NMA
for(i in 2:3){d[i] ~ dnorm(D.d[1], D.d.prec[1])} #Control/Sham
for(i in 4:6){d[i] ~ dnorm(D.d[2], D.d.prec[2])} #Tolterodine
for(i in 7:24){d[i]  ~ dnorm(D.d[3], D.d.prec[3])} #Oxybutynin
for(i in 25:27){d[i]  ~ dnorm(D.d[4], D.d.prec[4])} #Fesoterodine
d[28] ~ dnorm(D.d[5], D.d.prec[5]) #Terodiline
for(i in 29:31){d[i] ~ dnorm(D.d[6], D.d.prec[6])} #Solifenacin
for(i in 32:33){d[i] ~ dnorm(D.d[41], D.d.prec[41])} #Tarafenacin
d[34] ~ dnorm(D.d[6], D.d.prec[6]) #Solifenacin
for(i in 35:37){d[i] ~ dnorm(D.d[8], D.d.prec[8])} #Imidafenacin
for(i in 38:40){d[i] ~ dnorm(D.d[9], D.d.prec[9])} #Darifenacin
for(i in 41:43){d[i] ~ dnorm(D.d[10], D.d.prec[10])} #Propiverine
d[44] ~ dnorm(D.d[11], D.d.prec[11]) #Trospium
d[45] ~ dnorm(D.d[42], D.d.prec[42]) #Lipo-BoNTA
d[46] ~ dnorm(D.d[11], D.d.prec[11]) #Trospium
d[47]  ~ dnorm(D.d[3], D.d.prec[3]) #Oxybutynin
for(i in 48:53){d[i] ~ dnorm(D.d[12], D.d.prec[12])} #Mirabegron
for(i in 54:55){d[i] ~ dnorm(D.d[13], D.d.prec[13])} #Solabegron
for(i in 56:57){d[i] ~ dnorm(D.d[14], D.d.prec[14])} #Cizilirtine
d[58] ~ dnorm(D.d[15], D.d.prec[15]) #ZD0947IL
for(i in 59:61){d[i] ~ dnorm(D.d[16], D.d.prec[16])} #ONO-8539
d[62] ~ dnorm(D.d[17], D.d.prec[17]) #Pregabalin
d[63] ~ dnorm(D.d[18], D.d.prec[18]) #Empronium Bromide
d[64] ~ dnorm(D.d[19], D.d.prec[19]) #Flavoxate chloride
d[65] ~ dnorm(D.d[20], D.d.prec[20]) #Duloxetine
d[66]  ~ dnorm(D.d[3], D.d.prec[3]) #Oxybutynin
d[67] ~ dnorm(D.d[22], D.d.prec[22]) #Resiniferatoxin
d[68] ~ dnorm(D.d[23], D.d.prec[23]) #Estradiol
for(i in 69:70){d[i] ~ dnorm(D.d[21], D.d.prec[21])} #Elocalcitol
d[71] ~ dnorm(D.d[25], D.d.prec[25]) #Reflexology
for(i in 72:73){d[i] ~ dnorm(D.d[26], D.d.prec[26])} #Botox
d[74] ~ dnorm(D.d[43], D.d.prec[43]) #Electrostim + Estrogen
```

```
d[75] ~ dnorm(D.d[44], D.d.prec[44]) #Estrogen
d[76] ~ dnorm(D.d[28], D.d.prec[28]) #Tolterodine + Pilocarpine
d[77] ~ dnorm(D.d[29], D.d.prec[29]) #Anticholinergic + Saline
for(i in 78:79){d[i] ~ dnorm(D.d[26], D.d.prec[26])} #Botox
d[80] ~ dnorm(D.d[7], D.d.prec[7]) #Electrostim
d[81] ~ dnorm(D.d[30], D.d.prec[30]) #Sacral Nerve Stimulation
d[82] ~ dnorm(D.d[31], D.d.prec[31]) #Tolterodine + Estrogen
d[83] ~ dnorm(D.d[7], D.d.prec[7]) #Percutaneous tibial nerve stimulation
d[84] ~ dnorm(D.d[32], D.d.prec[32]) #Physiotherapy
d[85] ~ dnorm(D.d[33], D.d.prec[33]) #Bladder training
d[86] ~ dnorm(D.d[27], D.d.prec[27]) #Oxybutynin + Behaviour
d[87] ~ dnorm(D.d[36], D.d.prec[36]) #Tolterodine + Behaviour
d[88] ~ dnorm(D.d[38], D.d.prec[38]) #Pregabalin + Tolterodine
d[89] ~ dnorm(D.d[40], D.d.prec[40]) #BT + PFE + Behaviour
d[90] ~ dnorm(D.d[24], D.d.prec[24]) #Oxybutynin + Salivary pastilles
d[91] ~ dnorm(D.d[39], D.d.prec[39]) #Trospium + Physiotherapy
d[92] ~ dnorm(D.d[27], D.d.prec[27]) #Oxybutynin + BT
d[93] ~ dnorm(D.d[36], D.d.prec[36]) #Tolterodine + BT
d[94] ~ dnorm(D.d[38], D.d.prec[38]) #Pregabalin + Tolterodine
d[95] ~ dnorm(D.d[37], D.d.prec[37]) #Tolterodine + PFE
d[96] ~ dnorm(D.d[34], D.d.prec[34]) #Tolterodine +Neurostim
d[97] ~ dnorm(D.d[35], D.d.prec[35]) #Electrostim + PFE+ BT

for(i in 1:44){
D.d[i] ~ dnorm(0, 0.001)
D.d.sd[i] ~ dunif(0,5)
D.d.prec[i] <- pow(D.d.sd[i],-2)
}

between.study.sd ~ dunif(0,5)
tau <- pow(between.study.sd,-2)

#between.study.sd ~ dnorm(0,1)I(0,)
#tau <- pow(between.study.sd,-2)

#tau ~ dgamma(0.001,0.001)
#between.study.sd <- 1/sqrt(tau)

for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[c,k] <- (d[k] - d[c] )}}
for (k in 1:nt) {
        rk[k] <- rank(d[],k)
        best[k] <- equals(rk[k],1)}
        }
}
```

D.2 Example WinBUGS code for hierarchical NMA of continuous outcomes incorporating dose constraints

```
Model{

for(i in 1:ns){
        for (k in 1:na[i]){
change_var[i,k] <- ((pow(b_sd_star[i,k],2) + pow(f_sd_star[i,k],2) -
2*rho[i,k]*b_sd_star[i,k]*f_sd_star[i,k])*equals(ind_c_miss[i,k],1)) +
(pow(sd[i,k],2)*(equals(ind_c_miss[i,k],0)))

change_sd[i,k] <- sqrt(change_var[i,k])
se[i,k] <- change_sd[i,k]/sqrt(numinclanalysis[i,k])

f_sd_star[i,k] <- (-0.011+(0.835*b_sd[i,k]))*(equals(ind_f_miss[i,k],1)) +
(f_sd[i,k])*(equals(ind_f_miss[i,k],0))

b_sd_star[i,k] <- 0*equals(ind_b_miss[i,k],1) +
(b_sd[i,k])*(equals(ind_b_miss[i,k],0))

b_sd[i,k] ~ dunif(0,15)
f_sd[i,k] ~ dunif(0, 15)
sd[i,k] ~ dunif(0,25)

z[i,k]~ dnorm(z.star[i,k],z.prec)
z.star[i,k] ~ dnorm(0.67, 12.76)
rho[i,k]<- (exp(2*z[i,k])-1)/(exp(2*z[i,k])+1)

}}

z.se <- 1/(sqrt(49-3))
z.prec<- pow(z.se,-2)

for(i in 1:ns){
        w[i,1] <- 0
        delta[i,1] <- 0
        mu[i] ~ dnorm(0,0.001)
for (k in 1:na[i]) {
        var[i,k] <- pow(se[i,k],2)
        prec[i,k] <- 1/var[i,k]
        y[i,k] ~ dnorm(theta[i,k],prec[i,k])
        theta[i,k] <- mu[i] + delta[i,k]
        dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
}

resdev[i] <- sum(dev[i,1:na[i]])
        for (k in 2:na[i]) {
                delta[i,k] ~ dnorm(md[i,k],taud[i,k])
                md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]
```

```
                taud[i,k] <- tau *2*(k-1)/k
                w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
                sw[i,k] <- sum(w[i,1:k-1])/(k-1)
                }
}

totresdev <- sum(resdev[])

d[1]<-0                                          # Set reference treatment to be zero
                                                                  # Hierarchical NMA
for(i in 2:3){d[i] ~ dnorm(D.d[1], D.d.prec[1])} #Control/Sham
for(i in 4:6){d[i] ~ dnorm(D.d[2], D.d.prec[2])} #Tolterodine
for(i in 7:24){d[i]  ~ dnorm(D.d[3], D.d.prec[3])} #Oxybutynin
for(i in 25:27){d[i]  ~ dnorm(D.d[4], D.d.prec[4])} #Fesoterodine
d[28] ~ dnorm(D.d[5], D.d.prec[5]) #Terodiline
for(i in 29:31){d[i] ~ dnorm(D.d[6], D.d.prec[6])} #Solifenacin
for(i in 32:33){d[i] ~ dnorm(D.d[41], D.d.prec[41])} #Tarafenacin
d[34] ~ dnorm(D.d[6], D.d.prec[6]) #Solifenacin
for(i in 35:37){d[i] ~ dnorm(D.d[8], D.d.prec[8])} #Imidafenacin
for(i in 38:40){d[i] ~ dnorm(D.d[9], D.d.prec[9])} #Darifenacin
for(i in 41:43){d[i] ~ dnorm(D.d[10], D.d.prec[10])} #Propiverine
d[44] ~ dnorm(D.d[11], D.d.prec[11]) #Trospium
d[45] ~ dnorm(D.d[42], D.d.prec[42]) #Lipo-BoNTA
d[46] ~ dnorm(D.d[11], D.d.prec[11]) #Trospium
d[47]  ~ dnorm(D.d[3], D.d.prec[3]) #Oxybutynin
for(i in 48:53){d[i] ~ dnorm(D.d[12], D.d.prec[12])} #Mirabegron
for(i in 54:55){d[i] ~ dnorm(D.d[13], D.d.prec[13])} #Solabegron
for(i in 56:57){d[i] ~ dnorm(D.d[14], D.d.prec[14])} #Cizilirtine
d[58] ~ dnorm(D.d[15], D.d.prec[15]) #ZD0947IL
for(i in 59:61){d[i] ~ dnorm(D.d[16], D.d.prec[16])} #ONO-8539
d[62] ~ dnorm(D.d[17], D.d.prec[17]) #Pregabalin
d[63] ~ dnorm(D.d[18], D.d.prec[18]) #Empronium Bromide
d[64] ~ dnorm(D.d[19], D.d.prec[19]) #Flavoxate chloride
d[65] ~ dnorm(D.d[20], D.d.prec[20]) #Duloxetine
d[66]  ~ dnorm(D.d[3], D.d.prec[3]) #Oxybutynin
d[67] ~ dnorm(D.d[22], D.d.prec[22]) #Resiniferatoxin
d[68] ~ dnorm(D.d[23], D.d.prec[23]) #Estradiol
for(i in 69:70){d[i] ~ dnorm(D.d[21], D.d.prec[21])} #Elocalcitol
d[71] ~ dnorm(D.d[25], D.d.prec[25]) #Reflexology
for(i in 72:73){d[i] ~ dnorm(D.d[26], D.d.prec[26])} #Botox
d[74] ~ dnorm(D.d[43], D.d.prec[43]) #Electrostim + Estrogen
d[75] ~ dnorm(D.d[44], D.d.prec[44]) #Estrogen
d[76] ~ dnorm(D.d[28], D.d.prec[28]) #Tolterodine + Pilocarpine
d[77] ~ dnorm(D.d[29], D.d.prec[29]) #Anticholinergic + Saline
for(i in 78:79){d[i] ~ dnorm(D.d[26], D.d.prec[26])} #Botox
d[80] ~ dnorm(D.d[7], D.d.prec[7]) #Electrostim
```

```
d[81] ~ dnorm(D.d[30], D.d.prec[30]) #Sacral Nerve Stimulation
d[82] ~ dnorm(D.d[31], D.d.prec[31]) #Tolterodine + Estrogen
d[83] ~ dnorm(D.d[7], D.d.prec[7]) #Percutaneous tibial nerve stimulation
d[84] ~ dnorm(D.d[32], D.d.prec[32]) #Physiotherapy
d[85] ~ dnorm(D.d[33], D.d.prec[33]) #Bladder training
d[86] ~ dnorm(D.d[27], D.d.prec[27]) #Oxybutynin + Behaviour
d[87] ~ dnorm(D.d[36], D.d.prec[36]) #Tolterodine + Behaviour
d[88] ~ dnorm(D.d[38], D.d.prec[38]) #Pregabalin + Tolterodine
d[89] ~ dnorm(D.d[40], D.d.prec[40]) #BT + PFE + Behaviour
d[90] ~ dnorm(D.d[24], D.d.prec[24]) #Oxybutynin + Salivary pastilles
d[91] ~ dnorm(D.d[39], D.d.prec[39]) #Trospium + Physiotherapy
d[92] ~ dnorm(D.d[27], D.d.prec[27]) #Oxybutynin + BT
d[93] ~ dnorm(D.d[36], D.d.prec[36]) #Tolterodine + BT
d[94] ~ dnorm(D.d[38], D.d.prec[38]) #Pregabalin + Tolterodine
d[95] ~ dnorm(D.d[37], D.d.prec[37]) #Tolterodine + PFE
d[96] ~ dnorm(D.d[34], D.d.prec[34]) #Tolterodine +Neurostim
d[97] ~ dnorm(D.d[35], D.d.prec[35]) #Electrostim + PFE+ BT

for(i in 1:44){
D.d[i] ~ dnorm(0, 0.001)
D.d.sd[i] ~ dunif(0,5)
D.d.prec[i] <- pow(D.d.sd[i],-2)
}

b1<-1                                   #Dose constraints lower dose first
b1~dbern(cons1)
cons1<- step(d[6]-d[5])-equals(d[6],d[5])

b2<-1
b2~dbern(cons2)
cons2<- step(d[11]-d[12])-equals(d[11],d[12])

b3<-1
b3~dbern(cons3)
cons3<- step(d[12]-d[10])-equals(d[12],d[10])

b4<-1
b4~dbern(cons4)
cons4<- step(d[16]-d[17])-equals(d[16],d[17])

b5<-1
b5~dbern(cons5)
cons5<- step(d[21]-d[19])-equals(d[21],d[19])

b7<-1
b7~dbern(cons7)
```

```
cons7<- step(d[19]-d[18])-equals(d[19],d[18])

b8<-1
b8~dbern(cons8)
cons8<- step(d[18]-d[7])-equals(d[18],d[7])

b9<-1
b9~dbern(cons9)
cons9<- step(d[20]-d[8])-equals(d[20],d[8])

b10<-1
b10~dbern(cons10)
cons10<- step(d[8]-d[9])-equals(d[8],d[9])

b14<-1
b14~dbern(cons14)
cons14<- step(d[25]-d[27])-equals(d[25],d[27])

b15<-1
b15~dbern(cons15)
cons15<- step(d[27]-d[26])-equals(d[27],d[26])

b16<-1
b16~dbern(cons16)
cons16<- step(d[29]-d[31])-equals(d[29],d[31])

b17<-1
b17~dbern(cons17)
cons17<- step(d[31]-d[30])-equals(d[31],d[30])

b18<-1
b18~dbern(cons18)
cons18<- step(d[35]-d[36])-equals(d[35],d[36])

b19<-1
b19~dbern(cons19)
cons19<- step(d[36]-d[37])-equals(d[36],d[37])

b20<-1
b20~dbern(cons20)
cons20<- step(d[39]-d[40])-equals(d[39],d[40])

b21<-1
b21~dbern(cons21)
cons21<- step(d[40]-d[38])-equals(d[40],d[38])
```

```
b37<-1
b37~dbern(cons37)
cons37<- step(d[41]-d[42])-equals(d[41],d[42])

b12<-1
b12~dbern(cons12)
cons12<- step(d[33]-d[32])-equals(d[33],d[32])

b24<-1
b24~dbern(cons24)
cons24<- step(d[48]-d[49])-equals(d[48],d[49])

b25<-1
b25~dbern(cons25)
cons25<- step(d[50]-d[51])-equals(d[50],d[51])

b26<-1
b26~dbern(cons26)
cons26<- step(d[51]-d[52])-equals(d[51],d[52])

b13<-1
b13~dbern(cons13)
cons13<- step(d[52]-d[53])-equals(d[52],d[53])

b27<-1
b27~dbern(cons27)
cons27<- step(d[54]-d[55])-equals(d[54],d[55])

b28<-1
b28~dbern(cons28)
cons28<- step(d[56]-d[57])-equals(d[56],d[57])

b29<-1
b29~dbern(cons29)
cons29<- step(d[59]-d[60])-equals(d[59],d[60])

b30<-1
b30~dbern(cons30)
cons30<- step(d[60]-d[61])-equals(d[60],d[61])

b31<-1
b31~dbern(cons31)
cons31<- step(d[70]-d[69])-equals(d[70],d[69])

b33<-1
b33~dbern(cons33)
```

```
cons33<- step(d[94]-d[88])-equals(d[94],d[88])

b34<-1
b34~dbern(cons34)
cons34<- step(d[72]-d[73])-equals(d[72],d[73])

b35<-1
b35~dbern(cons35)
cons35<- step(d[13]-d[47])-equals(d[13],d[47])

b36<-1
b36~dbern(cons36)
cons36<- step(d[47]-d[60])-equals(d[47],d[60])


between.study.sd ~ dunif(0,5)
tau <- pow(between.study.sd,-2)

#between.study.sd ~ dnorm(0,1)I(0,)
#tau <- pow(between.study.sd,-2)

#tau ~ dgamma(0.001,0.001)
#between.study.sd <- 1/sqrt(tau)

for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[c,k] <- (d[k] - d[c] )}}
for (k in 1:nt) {
        rk[k] <- rank(d[],k)
        best[k] <- equals(rk[k],1)}
        }
}
```

D.3 Example WinBUGS code for hierarchical NMA of binary outcomes

```
Model{
        for(i in 1:ns){
        w[i,1] <- 0
        delta[i,t[i,1]] <- 0
        mu[i] ~  dnorm(0.0,0.01)
for (k in 1:na[i]) {
        r[i,k] ~ dbin(p[i,t[i,k]],n[i,k])
        logit(p[i,t[i,k]]) <- mu[i] + delta[i,t[i,k]]
        rhat[i,k] <- p[i,t[i,k]] * n[i,k]
        dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k])) + (n[i,k]-r[i,k]) * (log(n[i,k]-
r[i,k]) - log(n[i,k]-rhat[i,k])))}

resdev[i] <- sum(dev[i,1:na[i]])

 for (k in 2:na[i]) {
        delta[i,t[i,k]] ~ dnorm(md[i,t[i,k]],taud[i,t[i,k]])
        md[i,t[i,k]] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
        taud[i,t[i,k]] <- tau *2*(k-1)/k
        w[i,k] <- (delta[i,t[i,k]] - d[t[i,k]] + d[t[i,1]])
        sw[i,k] <- sum(w[i,1:k-1])/(k-1)}
 }
totresdev <- sum(resdev[])

d[1]<-0                                  # reference treatment effect is set to zero
                                                        #Hierarchical NMA
for(k in 2:3){d[k]~dnorm(D.d[1], prec.d[1])} #Control
for(k in 4:6){d[k]~dnorm(D.d[2], prec.d[2])} #Tolterodine
d[7]~dnorm(D.d[3], prec.d[3]) #Oxybutynin
for(k in 8:9){d[k]~dnorm(D.d[2], prec.d[2])} #Tolterodine
for(k in 10:12){d[k]~dnorm(D.d[4], prec.d[4])} #Fesoterodine
for(k in 13:18){d[k]~dnorm(D.d[3], prec.d[3])} #Oxybutynin
for(k in 19:22){d[k]~dnorm(D.d[5], prec.d[5])} #Propiverine
d[23]~dnorm(D.d[6], prec.d[6]) #Netupitant
d[24]~dnorm(D.d[3], prec.d[3]) #Oxybutynin
for(k in 25:27){d[k]~dnorm(D.d[4], prec.d[4])} #Fesoterodine
d[28]~dnorm(D.d[7], prec.d[7]) #Terodiline
for(k in 29:34){d[k]~dnorm(D.d[8], prec.d[8])} #Solifenacin
for(k in 35:37){d[k]~dnorm(D.d[9], prec.d[9])} #Imidafenacin
for(k in 38:40){d[k]~dnorm(D.d[10], prec.d[10])} #Darifenacin
for(k in 41:43){d[k]~dnorm(D.d[5], prec.d[5])} #Propiverine
for(k in 44:47){d[k]~dnorm(D.d[11], prec.d[11])} #Trospium chloride
for(k in 48:53){d[k]~dnorm(D.d[12], prec.d[12])} #Mirabegron
for(k in 54:55){d[k]~dnorm(D.d[13], prec.d[13])} #Solabegron
for(k in 56:57){d[k]~dnorm(D.d[14], prec.d[14])} #Cizolirtine citrate
d[58]~dnorm(D.d[15], prec.d[15]) #ZD0947IL
for(k in 59:61){d[k]~dnorm(D.d[16], prec.d[16])} #ONO-8539
```

```
d[62]~dnorm(D.d[17], prec.d[17]) #Pregabalin
for(k in 63:64){d[k]~dnorm(D.d[6], prec.d[6])} #Netupitant
d[65]~dnorm(D.d[18], prec.d[18]) #Pregabalin + Tolterodine
d[66]~dnorm(D.d[19], prec.d[19]) #Duloxetine
d[67]~dnorm(D.d[18], prec.d[18]) #Pregabalin + Tolterodine
d[68]~dnorm(D.d[20], prec.d[20]) #Solif/Trospium + placebo injection
d[69]~dnorm(D.d[21], prec.d[21]) #Tolterodine + BT
d[70]~dnorm(D.d[22], prec.d[22]) #BT
d[71]~dnorm(D.d[23], prec.d[23]) #Electrostim
for(k in 72:76){d[k]~dnorm(D.d[24], prec.d[24])} #OnaBoNTA
d[77]~dnorm(D.d[25], prec.d[25]) #Solifenacin + BT
d[78]~dnorm(D.d[23], prec.d[23]) #Electrostim
d[79]~dnorm(D.d[26], prec.d[26]) #LipoBoNTA
for(k in 80:81){d[k]~dnorm(D.d[27], prec.d[27])} #Tarafenacin

for (i in 1:27){
        D.d[i]~dnorm(0.0,0.01)
        prec.d[i]<-1/(sd.d[i]*sd.d[i])
        sd.d[i]~dunif(0,2)  }   # vague priors on class effects

sd ~ dunif(0,2)
tau <- pow(sd,-2)

#sd ~ dnorm(0,1)I(0,)
#tau <- pow(sd,-2)

#tau ~ dgamma(0.001,0.001)
#sd <- 1/sqrt(tau)

#Rank the treatment effects (with 1=worst) & record the worst treatment
for(k in 1:nt){
                exp.d[k] <- exp(d[k])
                rkworst[k]<-nt+1-rank(d[],k)
                worst[k]<-equals(rkworst[k],1)
                rkbest[k]<-rank(d[],k)
                best[k]<-equals(rkbest[k],1)}

#All pairwise log odds ratios and odds ratios
for(c in 1:(nt-1)){
        for(k in (c+1):nt){
                or[c,k]<-exp(d[k]-d[c])
                lor[c,k]<-(d[k]-d[c])
        }}
}
```

D.4 Example WinBUGS code for hierarchical NMA of binary outcomes incorporating dose constraints

```
Model{
        for(i in 1:ns){
        w[i,1] <- 0
        delta[i,t[i,1]] <- 0
        mu[i] ~  dnorm(0.0,0.01)
for (k in 1:na[i]) {
        r[i,k] ~ dbin(p[i,t[i,k]],n[i,k])
        logit(p[i,t[i,k]]) <- mu[i] + delta[i,t[i,k]]
        rhat[i,k] <- p[i,t[i,k]] * n[i,k]
        dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k])) + (n[i,k]-r[i,k]) * (log(n[i,k]-
r[i,k]) - log(n[i,k]-rhat[i,k])))}

resdev[i] <- sum(dev[i,1:na[i]])

 for (k in 2:na[i]) {
        delta[i,t[i,k]] ~ dnorm(md[i,t[i,k]],taud[i,t[i,k]])
        md[i,t[i,k]] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
        taud[i,t[i,k]] <- tau *2*(k-1)/k
        w[i,k] <- (delta[i,t[i,k]] - d[t[i,k]] + d[t[i,1]])
        sw[i,k] <- sum(w[i,1:k-1])/(k-1)}
 }
totresdev <- sum(resdev[])

d[1]<-0                                    # reference treatment effect is set to zero
                                                    #Hierarchical NMA
for(k in 2:3){d[k]~dnorm(D.d[1], prec.d[1])} #Control
for(k in 4:6){d[k]~dnorm(D.d[2], prec.d[2])} #Tolterodine
d[7]~dnorm(D.d[3], prec.d[3]) #Oxybutynin
for(k in 8:9){d[k]~dnorm(D.d[2], prec.d[2])} #Tolterodine
for(k in 10:12){d[k]~dnorm(D.d[4], prec.d[4])} #Fesoterodine
for(k in 13:18){d[k]~dnorm(D.d[3], prec.d[3])} #Oxybutynin
for(k in 19:22){d[k]~dnorm(D.d[5], prec.d[5])} #Propiverine
d[23]~dnorm(D.d[6], prec.d[6]) #Netupitant
d[24]~dnorm(D.d[3], prec.d[3]) #Oxybutynin
for(k in 25:27){d[k]~dnorm(D.d[4], prec.d[4])} #Fesoterodine
d[28]~dnorm(D.d[7], prec.d[7]) #Terodiline
for(k in 29:34){d[k]~dnorm(D.d[8], prec.d[8])} #Solifenacin
for(k in 35:37){d[k]~dnorm(D.d[9], prec.d[9])} #Imidafenacin
for(k in 38:40){d[k]~dnorm(D.d[10], prec.d[10])} #Darifenacin
for(k in 41:43){d[k]~dnorm(D.d[5], prec.d[5])} #Propiverine
for(k in 44:47){d[k]~dnorm(D.d[11], prec.d[11])} #Trospium chloride
for(k in 48:53){d[k]~dnorm(D.d[12], prec.d[12])} #Mirabegron
for(k in 54:55){d[k]~dnorm(D.d[13], prec.d[13])} #Solabegron
for(k in 56:57){d[k]~dnorm(D.d[14], prec.d[14])} #Cizolirtine citrate
d[58]~dnorm(D.d[15], prec.d[15]) #ZD0947IL
for(k in 59:61){d[k]~dnorm(D.d[16], prec.d[16])} #ONO-8539
```

```
d[62]~dnorm(D.d[17], prec.d[17]) #Pregabalin
for(k in 63:64){d[k]~dnorm(D.d[6], prec.d[6])} #Netupitant
d[65]~dnorm(D.d[18], prec.d[18]) #Pregabalin + Tolterodine
d[66]~dnorm(D.d[19], prec.d[19]) #Duloxetine
d[67]~dnorm(D.d[18], prec.d[18]) #Pregabalin + Tolterodine
d[68]~dnorm(D.d[20], prec.d[20]) #Solif/Trospium + placebo injection
d[69]~dnorm(D.d[21], prec.d[21]) #Tolterodine + BT
d[70]~dnorm(D.d[22], prec.d[22]) #BT
d[71]~dnorm(D.d[23], prec.d[23]) #Electrostim
for(k in 72:76){d[k]~dnorm(D.d[24], prec.d[24])} #OnaBoNTA
d[77]~dnorm(D.d[25], prec.d[25]) #Solifenacin + BT
d[78]~dnorm(D.d[23], prec.d[23]) #Electrostim
d[79]~dnorm(D.d[26], prec.d[26]) #LipoBoNTA
for(k in 80:81){d[k]~dnorm(D.d[27], prec.d[27])} #Tarafenacin

for (i in 1:27){
        D.d[i]~dnorm(0.0,0.01)
        prec.d[i]<-1/(sd.d[i]*sd.d[i])
        sd.d[i]~dunif(0,2)  }   # vague priors on class effects

b1<-1                                   #Dose constraints higher dose first
b1~dbern(cons1)
cons1<- step(d[5]-d[6])-equals(d[5],d[6])

b2<-1
b2~dbern(cons2)
cons2<- step(d[7]-d[18])-equals(d[7],d[18])

b3<-1
b3~dbern(cons3)
cons3<- step(d[8]-d[5])-equals(d[8],d[5])

b4<-1
b4~dbern(cons4)
cons4<- step(d[6]-d[9])-equals(d[6],d[9])

b5<-1
b5~dbern(cons5)
cons5<- step(d[17]-d[16])-equals(d[17],d[16])

b6<-1
b6~dbern(cons6)
cons6<- step(d[10]-d[11])-equals(d[10],d[11])

b7<-1
b7~dbern(cons7)
```

```
cons7<- step(d[11]-d[12])-equals(d[11],d[12])

b8<-1
b8~dbern(cons8)
cons8<- step(d[26]-d[27])-equals(d[26],d[27])

b9<-1
b9~dbern(cons9)
cons9<- step(d[27]-d[25])-equals(d[27],d[25])

b10<-1
b10~dbern(cons10)
cons10<- step(d[33]-d[30])-equals(d[33],d[30])

b11<-1
b11~dbern(cons11)
cons11<- step(d[30]-d[31])-equals(d[30],d[31])

b12<-1
b12~dbern(cons12)
cons12<- step(d[31]-d[29])-equals(d[31],d[29])

b13<-1
b13~dbern(cons13)
cons13<- step(d[29]-d[32])-equals(d[29],d[32])

b14<-1
b14~dbern(cons14)
cons14<- step(d[81]-d[80])-equals(d[81],d[80])

b15<-1
b15~dbern(cons15)
cons15<- step(d[37]-d[36])-equals(d[37],d[36])

b16<-1
b16~dbern(cons16)
cons16<- step(d[36]-d[35])-equals(d[36],d[35])

b17<-1
b17~dbern(cons17)
cons17<- step(d[38]-d[40])-equals(d[38],d[40])

b18<-1
b18~dbern(cons18)
cons18<- step(d[40]-d[39])-equals(d[40],d[39])
```

```
b19<-1
b19~dbern(cons19)
cons19<- step(d[20]-d[21])-equals(d[20],d[21])

b20<-1
b20~dbern(cons20)
cons20<- step(d[19]-d[42])-equals(d[19],d[42])

b21<-1
b21~dbern(cons21)
cons21<- step(d[42]-d[41])-equals(d[42],d[41])

b22<-1
b22~dbern(cons22)
cons22<- step(d[41]-d[43])-equals(d[41],d[43])

b23<-1
b23~dbern(cons23)
cons23<- step(d[46]-d[45])-equals(d[46],d[45])

b24<-1
b24~dbern(cons24)
cons24<- step(d[23]-d[63])-equals(d[23],d[63])

b25<-1
b25~dbern(cons25)
cons25<- step(d[63]-d[64])-equals(d[63],d[64])

b26<-1
b26~dbern(cons26)
cons26<- step(d[49]-d[48])-equals(d[49],d[48])

b27<-1
b27~dbern(cons27)
cons27<- step(d[55]-d[54])-equals(d[55],d[54])

b28<-1
b28~dbern(cons28)
cons28<- step(d[53]-d[52])-equals(d[53],d[52])

b29<-1
b29~dbern(cons29)
cons29<- step(d[52]-d[51])-equals(d[52],d[51])

b30<-1
b30~dbern(cons30)
```

```
cons30<- step(d[51]-d[50])-equals(d[51],d[50])

b31<-1
b31~dbern(cons31)
cons31<- step(d[57]-d[56])-equals(d[57],d[56])

b32<-1
b32~dbern(cons32)
cons32<- step(d[61]-d[60])-equals(d[61],d[60])

b33<-1
b33~dbern(cons33)
cons33<- step(d[60]-d[59])-equals(d[60],d[59])

b38<-1
b38~dbern(cons38)
cons38<- step(d[67]-d[65])-equals(d[67],d[65])

b34<-1
b34~dbern(cons34)
cons34<- step(d[76]-d[73])-equals(d[76],d[73])

b35<-1
b35~dbern(cons35)
cons35<- step(d[73]-d[75])-equals(d[73],d[75])

b36<-1
b36~dbern(cons36)
cons36<- step(d[75]-d[72])-equals(d[75],d[72])

b37<-1
b37~dbern(cons37)
cons37<- step(d[72]-d[74])-equals(d[72],d[74])


#prec.d ~ dgamma(0.001,0.001)
#sd.d <- 1/sqrt(prec.d)

sd ~ dunif(0,2)
tau <- pow(sd,-2)

#sd ~ dnorm(0,1)I(0,)
#tau <- pow(sd,-2)

#tau ~ dgamma(0.001,0.001)
#sd <- 1/sqrt(tau)
```

```
#Rank the treatment effects (with 1=worst) & record the worst treatment
for(k in 1:nt){
                exp.d[k] <- exp(d[k])
                rkworst[k]<-nt+1-rank(d[],k)
                worst[k]<-equals(rkworst[k],1)
                rkbest[k]<-rank(d[],k)
                best[k]<-equals(rkbest[k],1)}

#All pairwise log odds ratios and odds ratios
for(c in 1:(nt-1)){
        for(k in (c+1):nt){
                or[c,k]<-exp(d[k]-d[c])
                lor[c,k]<-(d[k]-d[c])
        }}
}
```

D.5 Results for efficacy outcomes

TABLE D.1: Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age-adjusted hierarchical network meta-analysis assuming a common regression coefficient

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95% CrI)	Rank (95% CrI)	p(Best)
	Sacral nerve stimulation	[81]	1	25	-5.35 (-9.43,-1.11)	1 (1,19)	0.78
	Electrostimulation + PFMT + BT	[97]	1	25	-3.35 (-5.36,-1.24)	2 (1,15)	0.17
	OnaBoNT-A 200U trigone sparing	[73]	1	86	-2.28 (-3.74,-1.29)	5 (2,15)	0.02
	PFMT + BT	[89]	2	67	-2.16 (-3.96,-0.27)	5 (2,74)	0.00
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-1.93 (-2.72,-1.15)	7 (3,19)	0.00
	Estradiol 3mg intravaginally	[128]	1	15	-1.83 (-3.72,-0.15)	8 (2,79)	0.01
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-1.81 (-3.17,-0.59)	8 (2,58)	0.00
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-1.69 (-2.87,-0.47)	9 (3,67)	0.00
	Lipo-BoNTA 200U	[138]	1	29	-1.52 (-3.18,0.09)	11 (2,86)	0.00
	OnaBoNT-A 100U trigone sparing	[72]	4	603	-1.56 (-1.97,-1.17)	11 (6,19)	0.00
	Cizolirtine citrate 400mg b.i.d	[57]	2	68	-1.46 (-2.84,-0.1)	12 (3,81)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	-1.48 (-2.42,-0.58)	12 (4,58)	0.00
	Estriol 1mg intravesically	[131]	1	21	-1.42 (-2.63,-0.23)	13 (3,77)	0.00
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-1.19 (-2.14,-0.62)	17 (5,54)	0.00
	Solifenacin ER 10mg q.d	[30]	5	1376	-1.13 (-1.38,-0.89)	19 (12,31)	0.00
	Electrostimulation	[80]	6	236	-1.08 (-1.65,-0.52)	20 (10,64)	0.00
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-1.08 (-1.83,-0.32)	20 (7,74)	0.00
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	96	-1.04 (-1.74,-0.37)	22 (8,72)	0.00
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-1 (-1.85,-0.3)	23 (7,75)	0.00
	Fesoterodine ER 8mg q.d	[26]	5	2319	-0.97 (-1.2,-0.75)	25 (15,45)	0.00
	Oxybutynin ER 10mg q.d	[8]	1	185	-0.94 (-1.54,-0.52)	27 (11,63)	0.00
	Propiverine IR 30mg b.i.d	[117]	1	45	-0.9 (-2.64,0.15)	29 (3,86)	0.00
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.9 (-2.27,0.47)	29 (5,91)	0.00
	Solabegron IR 125mg b.i.d	[55]	1	85	-0.89 (-1.2,-0.57)	30 (16,62)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	-0.84 (-1.35,-0.53)	34 (14,64)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	63	-0.83 (-1.36,-0.51)	34 (14,65)	0.00
	Oxybutynin vaginal ring 6mg q.d	[17]	1	96	-0.82 (-1.35,-0.38)	36 (14,70)	0.00
	Solifenacin ER 5-10mg q.d	[31]	3	1431	-0.81 (-1.09,-0.53)	36 (19,65)	0.00
	Mirabegron ER 50mg q.d	[51]	6	2014	-0.82 (-1.01,-0.63)	36 (22,57)	0.00
	Pregabalin 150mg b.i.d	[62]	1	94	-0.81 (-1.36,-0.27)	36 (13,77)	0.00
	Mirabegron ER 25mg q.d	[50]	2	619	-0.8 (-1.06,-0.54)	37 (20,64)	0.00
	Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	1	130	-0.8 (-1.3,-0.31)	37 (14,75)	0.00
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-0.78 (-1.53,-0.18)	39 (12,78)	0.00
	Trospium chloride ER 60mg q.d	[44]	2	565	-0.76 (-1.16,-0.37)	41 (17,73)	0.00
	Mirabegron ER 100mg q.d	[52]	3	1136	-0.76 (-0.97,-0.54)	41 (24,64)	0.00
	Oxybutynin intravesically 5mg t.i.d	[14]	1	9	-0.76 (-1.59,-0.09)	42 (11,81)	0.00
	Oxybutynin IR 3mg t.i.d	[19]	1	244	-0.75 (-1.17,-0.35)	42 (18,72)	0.00
	Oxybutynin gel 84mg/day	[134]	1	211	-0.76 (-1.2,-0.31)	42 (17,73)	0.00
	Fesoterodine ER 4mg q.d	[25]	7	3019	-0.74 (-0.92,-0.55)	43 (25,64)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	4	1485	-0.72 (-0.92,-0.5)	45 (25,68)	0.00
	Oxybutynin ER 5-30mg q.d + BT	[86]	1	32	-0.72 (-2.27,0.92)	46 (5,95)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.71 (-1.09,-0.32)	46 (21,74)	0.00
	Oxybutynin IR 2.5 - 5mg b.i.d	[24]	3	842	-0.72 (-1.16,-0.27)	46 (19,75)	0.00
	Oxybutynin IR 5mg t.i.d	[7]	6	355	-0.72 (-1.06,-0.36)	46 (22,71)	0.00
	Oxybutynin trandermal 3.9mg/day	[10]	3	292	-0.7 (-1,-0.36)	47 (24,71)	0.00
	Solifenacin ER 5mg q.d	[29]	10	1270	-0.7 (-0.93,-0.51)	47 (26,67)	0.00
	Oxybutynin vaginal ring 4mg q.d	[16]	1	115	-0.69 (-1.12,-0.21)	49 (20,77)	0.00
	Propiverine ER 20mg q.d	[41]	5	1144	-0.68 (-0.92,-0.46)	49 (26,69)	0.00
	Tolterodine IR 2mg b.i.d	[5]	18	3173	-0.67 (-0.84,-0.52)	51 (34,66)	0.00
	Oxybutynin 20mg intravesically q.d	[106]	1	21	-0.67 (-1.21,0.04)	51 (17,85)	0.00
	Oxybutynin ER 5 - 30mg/day	[22]	2	92	-0.64 (-1.15,0.05)	53 (20,84)	0.00
	Tolterodine IR 1mg b.i.d	[6]	3	279	-0.64 (-0.97,-0.31)	53 (24,75)	0.00
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.62 (-1.51,0.26)	55 (11,89)	0.00
	Tolterodine ER 4mg q.d	[4]	27	7085	-0.63 (-0.74,-0.51)	55 (41,68)	0.00
	Oxybutynin ER 15mg q.d	[9]	1	53	-0.61 (-1.11,0.15)	56 (20,87)	0.00

TABLE D.1: Estimated posterior median difference (and 95% credible interval) in change from baseline for voiding episodes relative to placebo obtained from age-adjusted hierarchical network meta-analysis assuming a common beta coefficient (cont.)

	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.61 (-1.2,-0.06)	57 (17,83)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	6	234	-0.6 (-1.24,0.02)	58 (17,84)	0.00
	Propiverine IR 45mg t.i.d	[118]	1	48	-0.58 (-1.79,0.75)	59 (9,93)	0.00
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	-0.58 (-1.31,0.15)	59 (16,88)	0.00
	Darifenacin ER 7.5mg q.d	[39]	1	29	-0.57 (-2,0.72)	60 (6,94)	0.00
	Trospium chloride IR 45mg t.i.d	[47]	1	615	-0.58 (-1.15,-0.01)	60 (19,84)	0.00
	Oxybutynin gel 56mg/day	[135]	1	198	-0.56 (-0.96,0)	60 (26,84)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	97	-0.55 (-1.09,0)	61 (20,84)	0.00
	Oxybutynin transdermal 1.3mg/day	[11]	1	128	-0.53 (-0.94,0.07)	62 (27,85)	0.00
	Oxybutynin IR 5mg b.i.d	[18]	1	116	-0.54 (-0.99,0.25)	62 (24,89)	0.00
	Oxybutynin transdermal 2.6mg/day	[12]	1	131	-0.52 (-0.93,0.08)	63 (28,86)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	3	507	-0.51 (-0.84,-0.2)	64 (33,79)	0.00
	Solabegron IR 50mg b.i.d	[54]	1	88	-0.51 (-0.83,-0.2)	64 (34,80)	0.00
	Propiverine IR 15mg b.i.d	[43]	2	145	-0.49 (-1.17,0.34)	65 (18,90)	0.00
	Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	2	32	-0.45 (-3.17,2.91)	68 (3,98)	0.00
	Darifenacin ER 15mg q.d	[40]	1	212	-0.42 (-1.25,0.4)	69 (16,91)	0.00
	Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	4	83	-0.41 (-1.52,0.87)	69 (12,94)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.39 (-1.83,1.11)	71 (8,96)	0.00
	Terodiline 25mg b.i.d	[28]	3	126	-0.35 (-0.82,0.12)	72 (34,88)	0.00
	Serlopitant 4mg q.d	[109]	1	114	-0.37 (-0.83,0.07)	72 (34,86)	0.00
	Serlopitant 0.25mg q.d	[107]	1	110	-0.36 (-0.83,0.08)	72 (34,86)	0.00
	Propiverine ER 60mg q.d	[119]	1	39	-0.28 (-1.36,1.42)	75 (15,97)	0.00
	Elocalcitol 150mg	[69]	1	87	-0.27 (-1.02,0.47)	75 (22,92)	0.00
	Netupitant 200mg q.d	[112]	1	55	-0.25 (-1.35,0.78)	76 (14,94)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	77	-0.22 (-1.13,0.7)	77 (19,94)	0.00
	Netupitant 100mg q.d	[111]	1	59	-0.2 (-1.25,0.86)	78 (16,94)	0.00
	Serlopitant 1mg q.d	[108]	1	110	-0.21 (-0.65,0.29)	78 (53,90)	0.00
	Percutaneous tibial nerve stimulation	[83]	5	198	-0.19 (-1.41,0.99)	78 (13,94)	0.00
	Netupitant 50mg q.d	[110]	1	60	-0.16 (-1.17,0.88)	79 (18,95)	0.00
	Electromagnetic stimulation	[125]	1	33	-0.12 (-2.01,1.87)	80 (6,97)	0.00
	Elocalcitol 75mg	[70]	1	84	-0.13 (-0.85,0.62)	80 (31,94)	0.00
	Reflexology	[71]	1	54	-0.05 (-1.78,1.74)	82 (9,96)	0.00
	Placebo	[1]	77	14550	NA	84 (77,90)	0.00
	ONO-8539 100mg b.i.d	[60]	1	83	0.07 (-0.66,0.74)	85 (52,94)	0.00
	Resiniferatoxin 50nM	[67]	1	34	0.05 (-1.21,1.38)	85 (17,97)	0.00
	ONO-8539 300mg b.i.d	[61]	1	82	0.18 (-0.52,0.87)	88 (63,95)	0.00
	Estradiol 1mg intravaginally	[127]	1	15	0.27 (-1.47,1.84)	89 (13,97)	0.00
	Propantheline Bromide 15mg t.i.d	[113]	1	23	0.25 (-0.83,1.33)	89 (35,97)	0.00
	ONO-8539 30mg b.i.d	[59]	1	87	0.34 (-0.35,1.08)	90 (73,96)	0.00
	ZD0947IL 25mg/day	[58]	1	92	0.68 (-0.62,1.99)	93 (55,97)	0.00
	Sham therapy	[3]	3	157	0.88 (-0.45,2.15)	95 (69,98)	0.00
	Control	[2]	4	138	1.07 (-0.12,2.5)	96 (82,98)	0.00
	Naftopidil 25mg q.d	[114]	1	22	3.14 (0.65,5.51)	98 (93,100)	0.00
	Trospium chloride IR 15mg t.i.d	[46]	2	30	4.38 (1.6,7.2)	99 (97,100)	0.00
	Solifenacin ER 5mg q.d + Naftopidil 25mg q.d	[115]	1	21	5.11 (2.63,7.7)	100 (98,100)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

TABLE D.2: Estimated posterior median difference (and 95% credible interval) in change from baseline for urgency episodes relative to placebo obtained from a hierarchical network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95%CrI)	Rank (95%CrI)	p(Best)
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-6.77 (-8.3,-5.22)	1 (1,1)	1.00
	Electrostimulation	[80]	4	161	-4.67 (-5.7,-3.65)	2 (2,3)	0.00
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	-3.17 (-4.74,-1.61)	4 (3,13)	0.00
	Percutaneous tibial nerve stimulation	[83]	2	52	-2.68 (-4.55,-0.86)	5 (3,29)	0.00
	Cizolirtine citrate 400mg b.i.d	[57]	1	16	-2.35 (-4.33,-0.44)	7 (3,40)	0.00
	Tolterodine ER 4mg q.d + Neurostimulation	[96]	1	20	-2.36 (-3.65,-1.09)	7 (3,23)	0.00
	OnaBoNT-A 100u bladder base + trigone	[79]	1	33	-2.22 (-3.93,-0.8)	8 (3,31)	0.00
	OnaBoNT-A 100u bladder body + trigone	[78]	1	35	-2.27 (-3.89,-1.03)	8 (3,25)	0.00
	OnaBoNT-A 200U trigone sparing	[73]	1	86	-2.19 (-3.44,-1.05)	8 (4,24)	0.00
	OnaBoNT-A 100u trigone sparing	[72]	3	592	-2.08 (-2.9,-1.26)	9 (5,19)	0.00
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-1.75 (-2.76,-0.8)	12 (5,32)	0.00
	Fesoterodine ER 8mg q.d	[26]	5	2319	-1.39 (-1.88,-0.94)	16 (10,28)	0.00
	Darifenacin ER 7.5mg q.d	[39]	1	29	-1.44 (-2.96,-0.07)	16 (4,46)	0.00
	Lipo-BoNTA 200U	[138]	1	29	-1.34 (-3.54,0.88)	17 (3,53)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1182	-1.24 (-1.81,-0.66)	19 (11,36)	0.00
	Solifenacin ER 10mg q.d	[30]	5	1341	-1.23 (-1.8,-0.73)	19 (11,33)	0.00
	Oxybutynin trandermal 3.9mg/day	[10]	1	48	-1.14 (-2.19,0.01)	22 (8,47)	0.00
	Fesoterodine ER 4mg q.d	[25]	7	3037	-1.11 (-1.53,-0.67)	22 (14,36)	0.00
	Imidafenacin 0.25mg b.i.d	[37]	1	76	-1.1 (-2.55,0.04)	23 (6,48)	0.00
	Solifenacin ER 5mg q.d	[29]	9	1246	-1.06 (-1.54,-0.58)	24 (14,37)	0.00
	Solifenacin ER 5 - 10mg q.d	[31]	3	1431	-1.01 (-1.55,-0.39)	25 (14,42)	0.00
	Oxybutynin ER 2.5mg q.d	[20]	1	16	-0.89 (-2.48,1.4)	28 (7,54)	0.00
	Propiverine ER 20mg q.d	[41]	4	1000	-0.89 (-1.47,-0.3)	28 (15,44)	0.00
	Propiverine IR 15mg b.i.d	[43]	1	100	-0.9 (-2.35,0.52)	28 (7,52)	0.00
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	-0.86 (-2.22,0.48)	29 (8,51)	0.00
	Cizolirtine citrate 200mg b.i.d	[56]	1	20	-0.86 (-2.58,0.83)	29 (6,53)	0.00
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.8 (-2.95,1.46)	31 (5,54)	0.00
	Pregabalin 150mg b.i.d	[62]	1	94	-0.78 (-1.88,0.33)	31 (11,51)	0.00
	Tolterodine ER 4mg q.d	[4]	16	4782	-0.76 (-1.11,-0.43)	32 (22,42)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	65	-0.75 (-1.61,-0.09)	32 (14,47)	0.00
	Imidafenacin 0.05mg b.i.d	[35]	1	91	-0.75 (-1.84,0.37)	33 (11,51)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	3	488	-0.74 (-1.51,0.02)	33 (15,48)	0.00
	Mirabegron ER 100mg q.d	[52]	3	1136	-0.72 (-1.25,-0.2)	33 (19,45)	0.00
	Mirabegron ER 50mg q.d	[51]	6	2014	-0.71 (-1.19,-0.25)	33 (21,44)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	63	-0.74 (-1.62,-0.05)	33 (13,47)	0.00
	Tolterodine IR 2mg b.i.d	[5]	5	746	-0.71 (-1.4,0.03)	34 (18,47)	0.00
	Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1	96	-0.71 (-1.95,0.5)	34 (10,52)	0.00
	Elocalcitol 75mg	[70]	1	84	-0.65 (-1.95,0.65)	35 (10,53)	0.00
	Mirabegron ER 25mg q.d	[50]	2	618	-0.65 (-1.23,0.02)	35 (20,48)	0.00
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.68 (-2.28,0.89)	35 (8,53)	0.00
	Darifenacin ER 15mg q.d.	[40]	1	212	-0.54 (-1.78,0.77)	38 (12,53)	0.00
	Elocalcitol 150mg	[69]	1	87	-0.45 (-1.71,0.85)	40 (12,53)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	77	-0.46 (-2.05,1.12)	40 (10,54)	0.00
	Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1	97	-0.44 (-1.58,0.71)	40 (14,53)	0.00
	Tolterodine IR 2mg b.i.d + PFMT	[95]	1	223	-0.4 (-1.85,1.05)	41 (11,54)	0.00
	Netupitant 100mg q.d	[111]	1	59	-0.41 (-1.57,0.72)	41 (14,53)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	2	44	-0.36 (-1.68,0.96)	42 (14,53)	0.00
	Netupitant 50mg q.d	[110]	1	60	-0.23 (-1.35,0.91)	44 (17,54)	0.00
	Netupitant 200mg q.d	[112]	1	55	-0.18 (-1.31,0.98)	45 (18,54)	0.00
	ZD0947IL 25mg/day	[58]	1	92	-0.09 (-1.98,1.76)	46 (10,54)	0.00
	Placebo	[1]	42	9362	NA	48 (42,52)	0.00
	ONO-8539 300mg b.i.d	[61]	1	82	0.02 (-1.28,1.22)	48 (18,54)	0.00
	ONO-8539 100mg b.i.d	[60]	1	83	0.16 (-1.09,1.35)	49 (23,54)	0.00
	ONO-8539 30mg b.i.d	[59]	1	87	0.61 (-0.62,1.94)	53 (37,54)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

TABLE D.3: Estimated posterior median difference (and 95% credible interval) in change from baseline for nocturia episodes relative to placebo obtained from a hierarchical network meta-analysis

Treatment pathway	Treatment	Code	Number of studies	Number of participants	Median difference† (95%CrI)	Rank	p(Best)
	Estriol 1mg intravesically	[131]	1	21	-2.19 (-3.15,-1)	1 (1,3)	0.64
	Estradiol 3mg intravaginally	[128]	1	15	-1.78 (-2.82,-0.67)	2 (1,4)	0.28
	Tolterodine IR 2mg b.i.d + BT	[93]	1	31	-0.54 (-1.11,-0.05)	5 (3,30)	0.00
	Percutaneous tibial nerve stimulation	[83]	3	162	-0.43 (-0.87,0.03)	7 (3,34)	0.00
	Electrostimulation	[80]	4	161	-0.47 (-0.81,-0.11)	7 (3,26)	0.00
	Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	1	102	-0.46 (-0.93,-0.02)	7 (3,32)	0.00
	Estradiol 1mg intravaginally	[127]	1	15	-0.4 (-1.74,0.96)	8 (2,40)	0.00
	Bladder Training (BT)/Behaviour Therapy	[85]	4	155	-0.38 (-0.63,-0.14)	8 (4,21)	0.00
	Imidafenacin 0.1mg b.i.d	[36]	1	77	-0.35 (-0.72,0)	9 (3,33)	0.00
	Tolterodine ER 4mg q.d + BT	[87]	1	134	-0.38 (-0.63,-0.12)	9 (4,25)	0.00
	Oxybutynin 20mg intravesically q.d	[106]	1	21	-0.26 (-2.89,0.01)	12 (1,34)	0.08
	OnaBoNT-A 100u trigone sparing	[72]	3	566	-0.26 (-0.41,-0.11)	12 (7,27)	0.00
	Mirabegron IR 100mg b.i.d	[48]	1	58	-0.2 (-0.45,-0.06)	15 (7,30)	0.00
	Fesoterodine ER 4 - 8mg q.d	[27]	3	1047	-0.21 (-0.32,-0.11)	15 (9,27)	0.00
	Mirabegron ER 25mg q.d	[50]	1	179	-0.19 (-0.36,-0.07)	16 (9,29)	0.00
	Oxybutynin ER 2.5mg q.d + BT	[92]	1	12	-0.16 (-0.72,0.43)	19 (4,40)	0.00
	Mirabegron ER 50mg q.d	[51]	5	1803	-0.16 (-0.23,-0.08)	20 (13,29)	0.00
	Mirabegron IR 150mg b.i.d	[49]	1	54	-0.16 (-0.34,0.01)	20 (9,34)	0.00
	Oxybutynin ER 2.5mg q.d	[20]	2	62	-0.16 (-0.46,0)	20 (7,33)	0.00
	Solifenacin ER 10mg q.d	[30]	2	445	-0.15 (-0.27,-0.02)	21 (11,33)	0.00
	Mirabegron ER 100mg q.d	[52]	3	1415	-0.14 (-0.23,-0.05)	22 (14,32)	0.00
	Fesoterodine ER 8mg q.d	[26]	5	1692	-0.14 (-0.22,-0.06)	22 (13,31)	0.00
	Fesoterodine ER 4mg q.d	[25]	7	2369	-0.14 (-0.21,-0.06)	23 (14,31)	0.00
	Solifenacin ER 5mg q.d	[29]	5	530	-0.14 (-0.27,-0.01)	23 (12,34)	0.00
	Tolterodine ER 4mg q.d	[4]	10	3363	-0.14 (-0.2,-0.07)	23 (15,30)	0.00
	Tolterodine IR 2mg b.i.d	[5]	2	143	-0.13 (-0.37,0.1)	24 (9,37)	0.00
	Tarafenacin 0.4mg q.d	[82]	1	76	-0.11 (-0.42,0.25)	26 (7,38)	0.00
	Oxybutynin trandermal 3.9mg/day	[10]	1	48	-0.11 (-0.36,0.13)	27 (9,37)	0.00
	Oxybutynin chloride topical gel 1g/day	[13]	1	389	-0.11 (-0.3,0.06)	27 (10,35)	0.00
	Oxbutynin patch 73.5mg	[15]	1	481	-0.1 (-0.21,0)	27 (15,34)	0.00
	Oxybutynin ER 5 - 30mg/day	[22]	1	60	-0.08 (-0.38,0.35)	29 (9,39)	0.00
	Propiverine ER 20mg q.d	[41]	5	1290	-0.07 (-0.15,0)	30 (20,35)	0.00
	Resiniferatoxin 50nM	[67]	1	34	-0.04 (-0.65,0.56)	32 (4,40)	0.00
	Sham therapy	[3]	1	110	-0.03 (-0.58,0.52)	32 (6,40)	0.00
	Oxybutynin IR 2.5mg t.i.d	[21]	2	47	-0.01 (-0.2,0.34)	33 (16,39)	0.00
	Placebo	[1]	35	6665	NA	34 (30,37)	0.00
	Terodiline 25mg b.i.d	[28]	3	126	0.13 (-0.04,0.31)	37 (31,40)	0.00
	Tarafenacin 0.2mg q.d	[90]	1	77	0.11 (-0.2,0.43)	37 (17,40)	0.00
	Vaginal oestrogen cream 1.25mg/day	[132]	1	98	0.24 (-0.26,0.67)	38 (12,40)	0.00
	Darifenacin ER 7.5mg q.d	[39]	1	29	0.27 (-0.32,0.93)	39 (11,40)	0.00

† median relative to a placebo intervention.

Rank denotes the posterior median estimate (95% credible interval). Ranks are calculated as the average treatment rank over all iterations and ranked according to the probability that each treatment is the best overall. Due to similarities in treatment effects and uncertainty around the point estimates, treatments are ranked in a different order at each iteration. Therefore, treatments that share a rank have a similar treatment effect on average. p(Best) denotes the probability that the intervention in question is the best. The probability best is calculated based on the number of iterations for which the intervention is ranked in first place.

First line therapy
Second line therapy
Second line therapy (if other second line therapies are contraindicated, clinically ineffective, or present unacceptable side effects)
Third line therapy
Not currently recommended

D.6 Assessing inconsistencies

TABLE D.4: Inconsistency between direct and indirect information obtained from hierarchical NMA for urinary incontinence episodes

Node 1	**Node 2**	**Indirect comparison**	**Direct comparison**	**Posterior probability**
1	4	-0.42	-0.51	0.22
1	5	-0.37	-0.47	0.28
1	11	-0.34	-0.56	0.30
1	25	-0.35	-0.48	0.37
1	29	-0.54	-0.68	0.35
1	30	-0.50	-0.88	0.21
1	31	-0.54	-0.67	0.45
1	36	-0.49	-0.43	0.57
1	41	-0.69	-0.40	0.66
1	43	-0.55	-0.54	0.51
1	50	-0.68	-0.07	0.51
1	51	-0.60	-0.52	0.69
1	52	-0.64	-0.57	0.59
4	5	0.04	0.16	0.75
4	10	0.11	-0.04	0.36
4	25	0.01	-0.04	0.41
4	26	-0.28	-0.11	0.80
4	29	-0.21	-0.30	0.45
4	31	-0.06	-0.49	0.05
4	50	-0.11	-0.40	0.19
4	51	-0.09	-0.14	0.40
4	52	-0.13	-0.15	0.45
5	7	-0.08	-0.22	0.34
5	24	-0.07	-0.51	0.18
5	29	-0.35	-0.19	0.76
5	30	-0.31	-0.48	0.23
5	43	0.00	-0.28	0.22
7	57	-0.12	0.05	0.61
10	11	0.09	0.25	0.64
11	24	-0.29	-1.36	0.12
25	26	0.11	-0.32	0.01
25	36	0.12	-0.13	0.27
29	30	-0.08	-0.15	0.36
29	41	0.29	0.30	0.52
30	41	0.41	0.34	0.39
36	41	0.06	-0.11	0.26
50	52	0.01	0.04	0.57

TABLE D.5: Inconsistency between direct and indirect information obtained from hierarchical NMA for number of patients experiencing adverse events

Node 1	**Node 2**	**Indirect comparison**	**Direct comparison**	**Posterior probability**
1	4	0.02	0.21	0.85
1	5	0.30	0.31	0.54
1	18	1.07	1.25	0.67
1	25	0.51	0.66	0.66
1	29	0.16	0.53	0.77
1	36	0.62	0.42	0.28
1	41	0.99	0.71	0.20
1	43	0.39	0.88	0.88
1	45	0.61	0.53	0.43
1	51	0.12	0.02	0.25
1	52	0.08	0.18	0.70
3	71	0.54	3.50	0.90
3	78	3.14	-0.25	0.06
4	25	0.43	0.10	0.13
4	26	0.74	0.45	0.16
4	29	-0.16	0.70	0.97
4	50	-0.20	0.20	0.87
4	51	-0.07	-0.19	0.25
4	52	-0.07	-0.10	0.47
4	71	-1.70	0.16	0.97
5	7	0.08	1.11	0.97
5	18	0.68	1.25	0.91
5	30	0.18	0.21	0.53
5	43	0.50	-0.03	0.11
5	71	0.05	-1.76	0.04
7	57	-0.60	1.33	1.00
18	29	-0.77	-1.58	0.10
18	45	-0.63	-0.44	0.66
25	26	0.08	0.41	0.91
25	36	-0.20	0.12	0.77
29	30	0.26	0.22	0.47
29	41	0.45	1.55	0.95
29	51	-0.27	-0.04	0.75
36	41	0.21	0.39	0.71
38	39	-1.32	-1.06	0.66
38	40	-1.24	-0.78	0.85
39	40	-0.13	0.37	0.87
50	52	-0.02	0.21	0.86
71	78	-1.33	-0.08	0.66
72	73	-0.40	0.08	0.80

D.7 Convergence diagnostics

FIGURE D.1: Brooks-Gelman-Rubin plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes

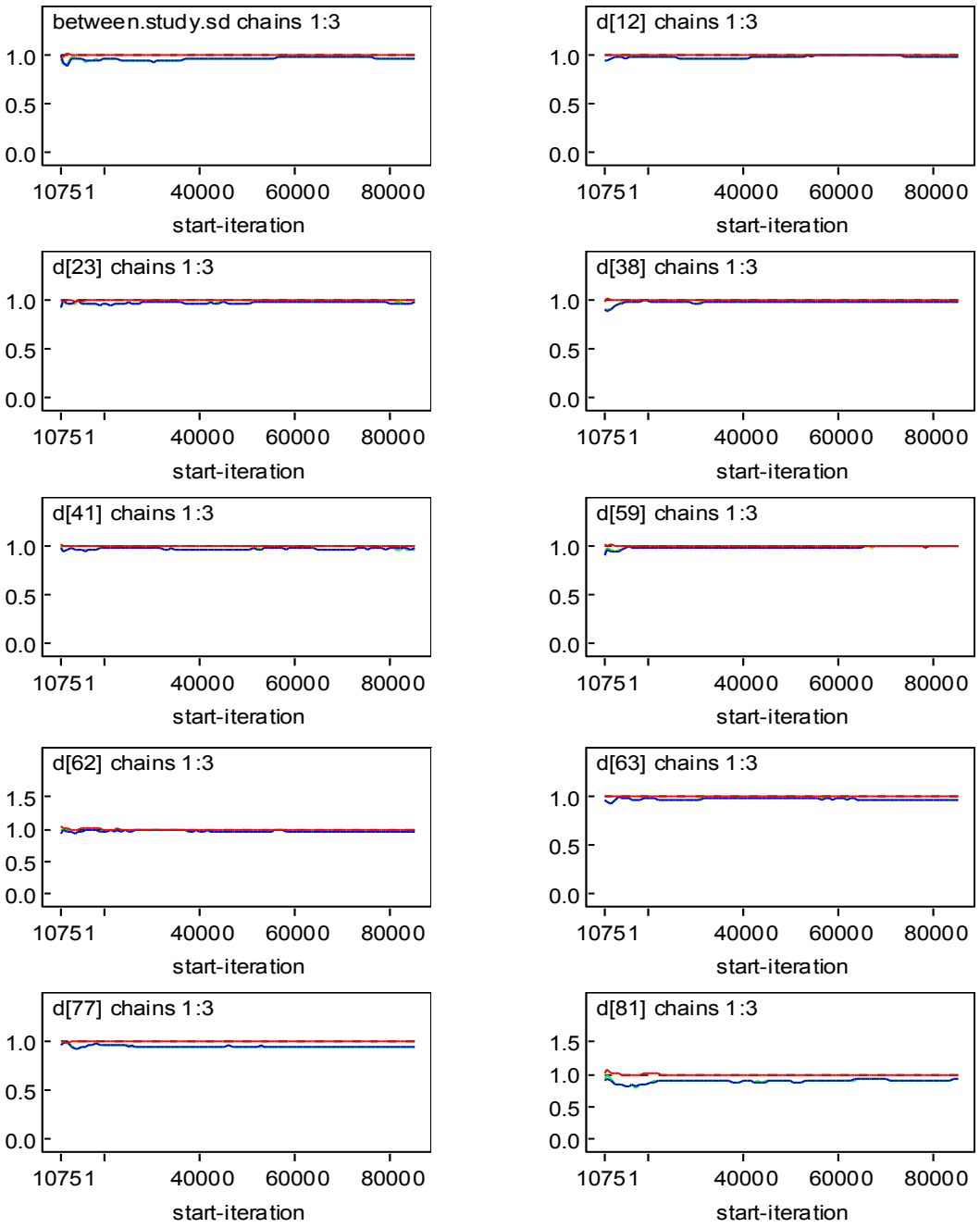


FIGURE D.2: Brooks-Gelman-Rubin plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events

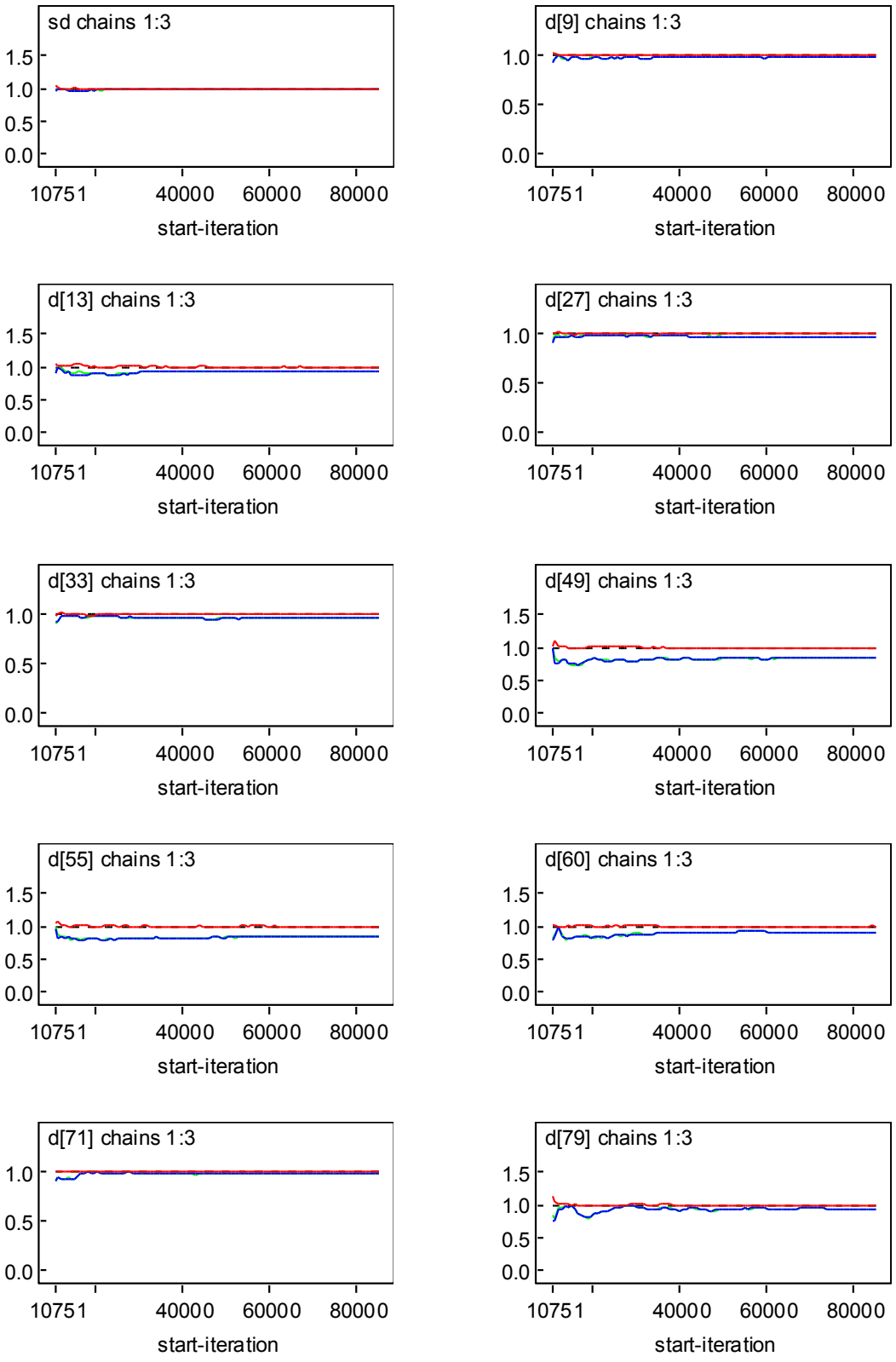


FIGURE D.3: Autocorrelation plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes

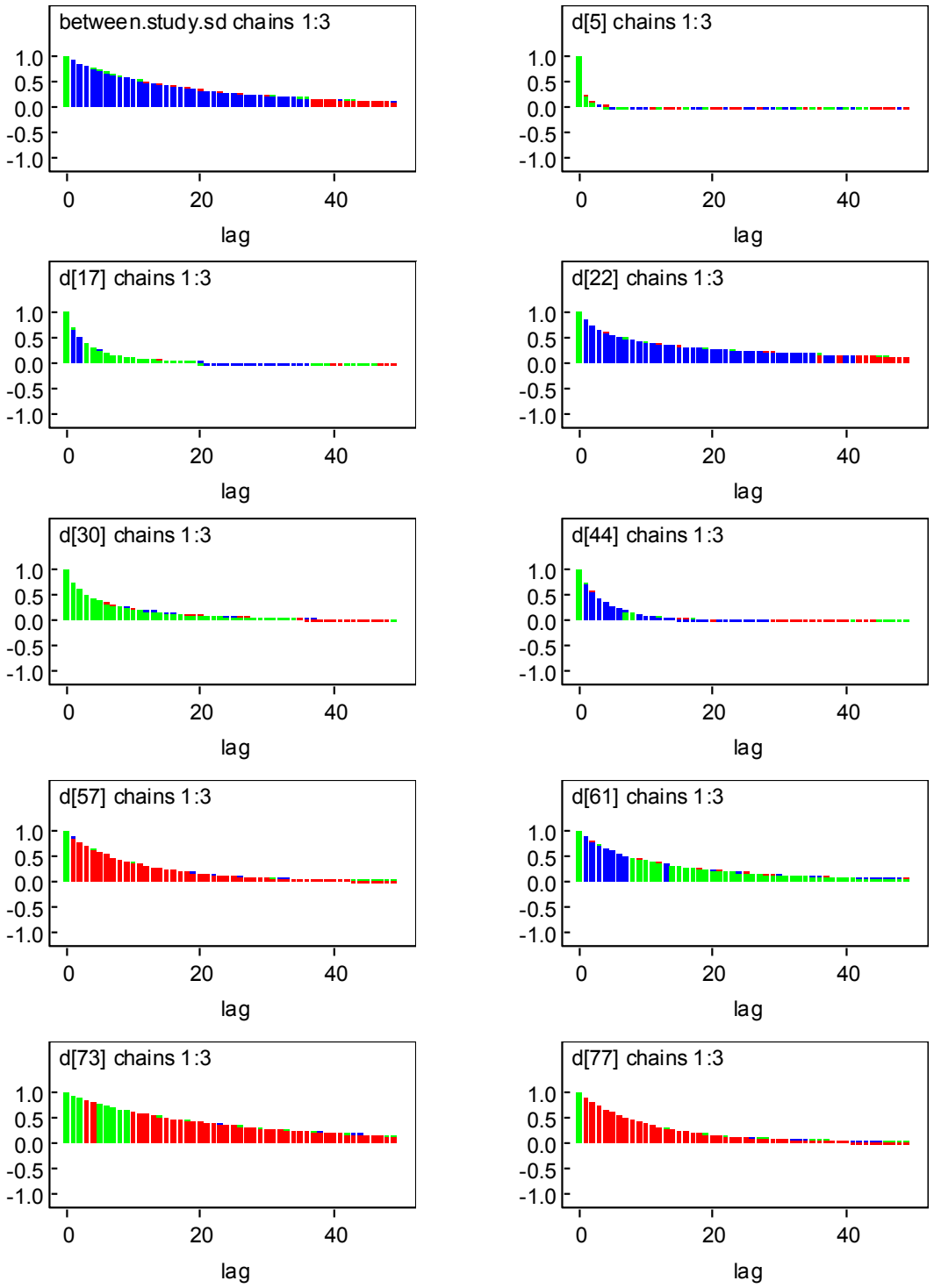


FIGURE D.4: Autocorrelation plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events

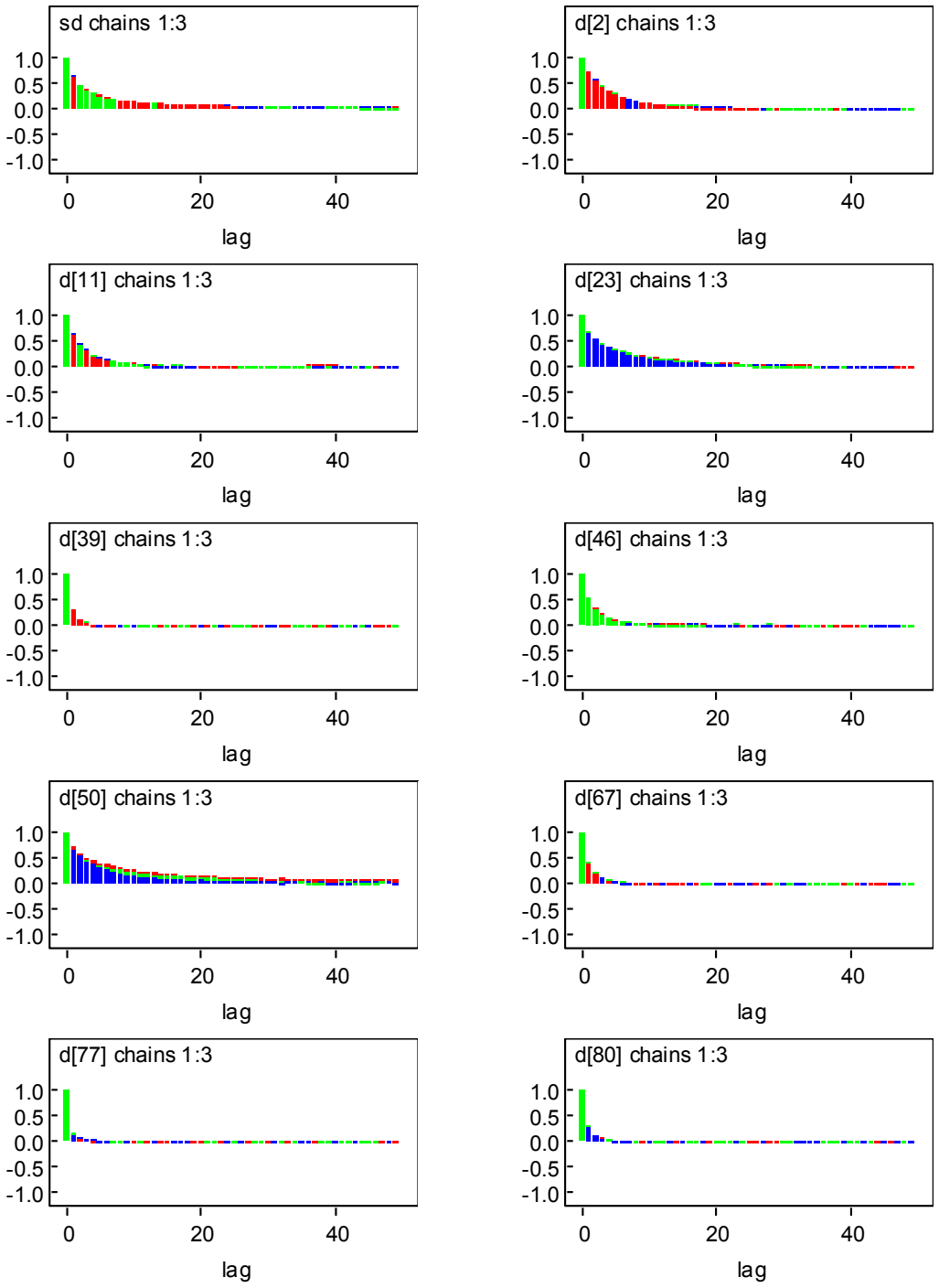


FIGURE D.5: History and trace plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes

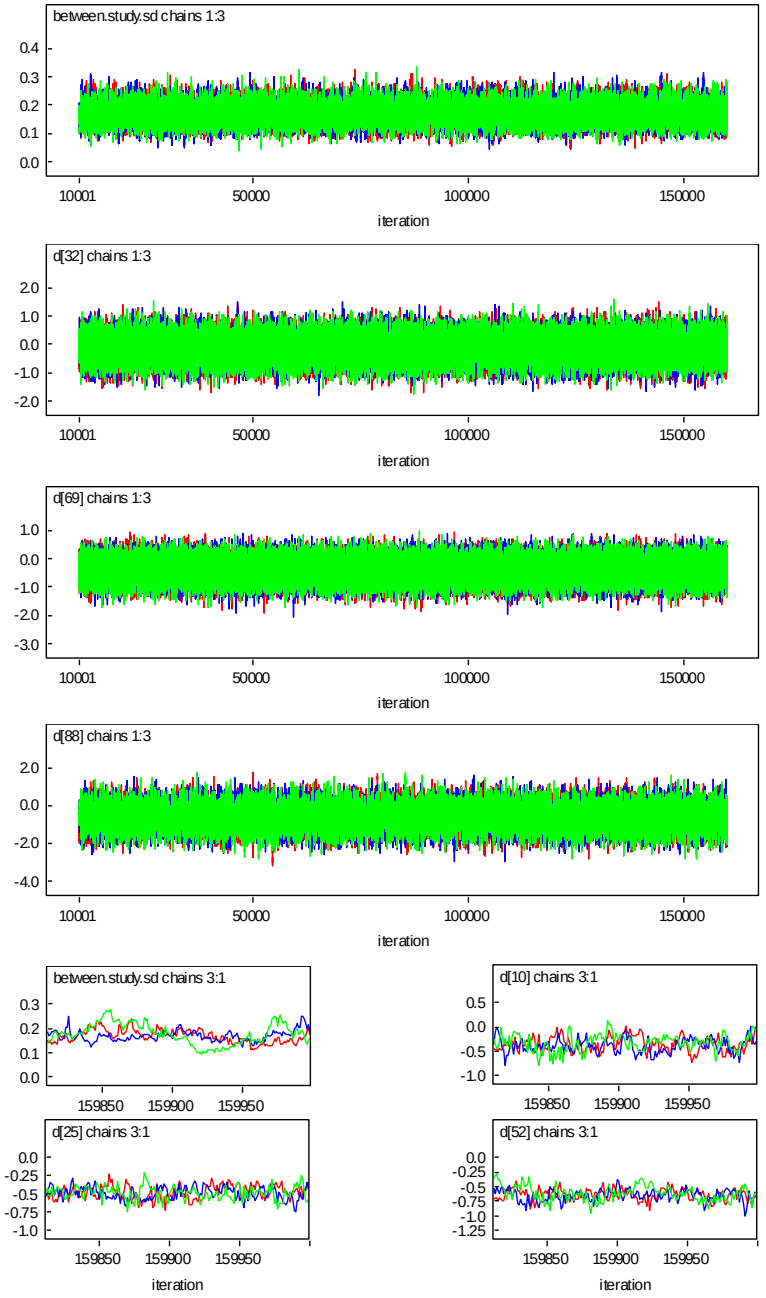
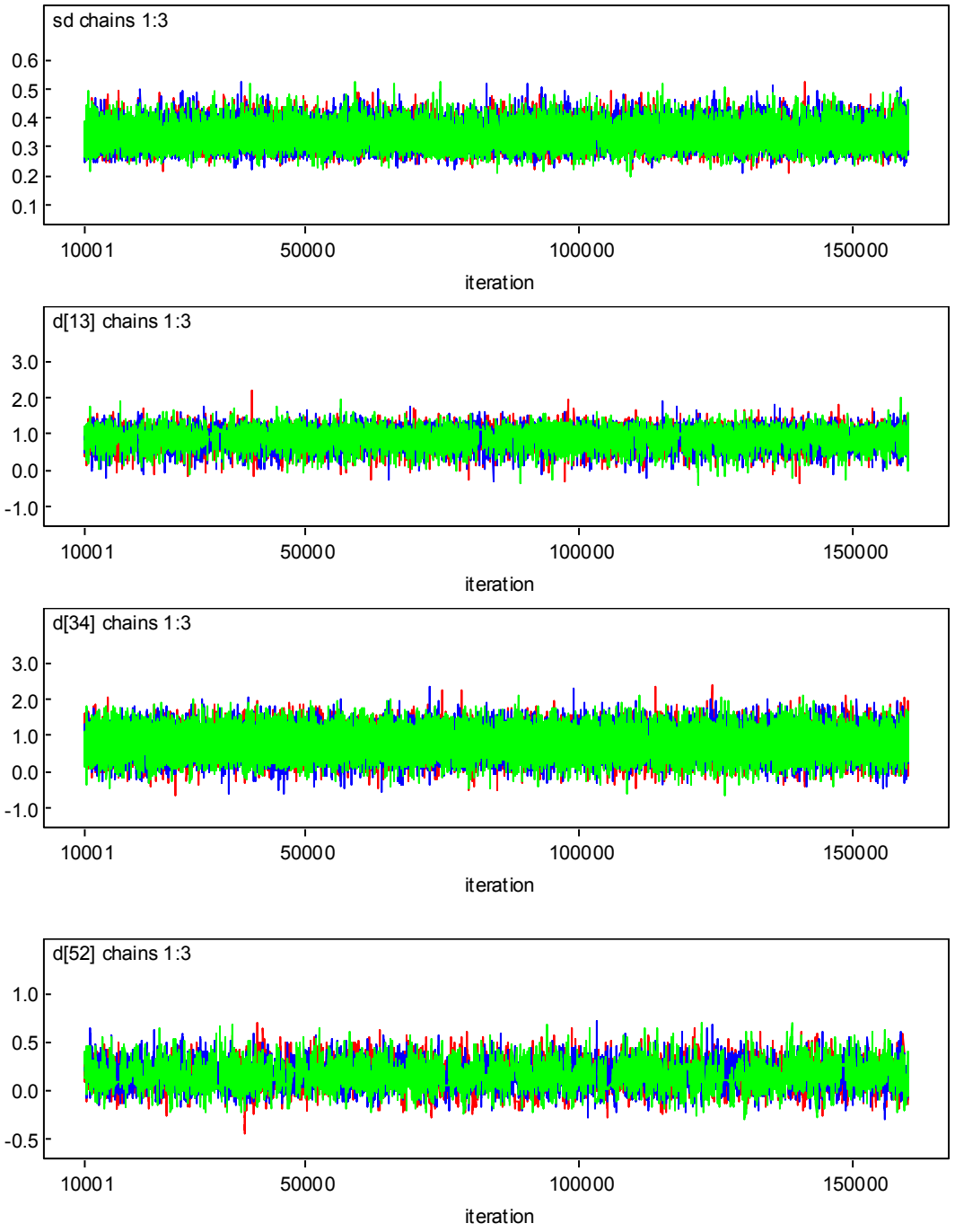
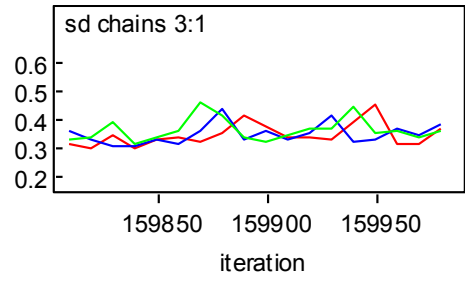
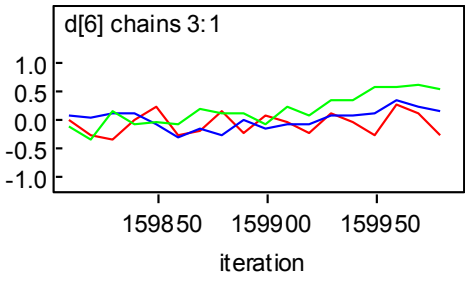
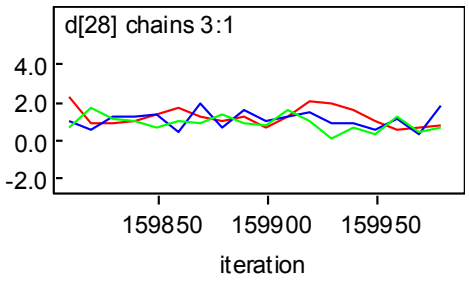


FIGURE D.6: History and trace plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events




sd chains 3:1
0.6
0.5
0.4
0.3
0.2
159850
159900
159950
iteration


d[6] chains 3:1
1.0
0.5
0.0
-0.5
-1.0
159850
159900
159950
iteration


d[28] chains 3:1
4.0
2.0
0.0
-2.0
159850
159900
159950
iteration

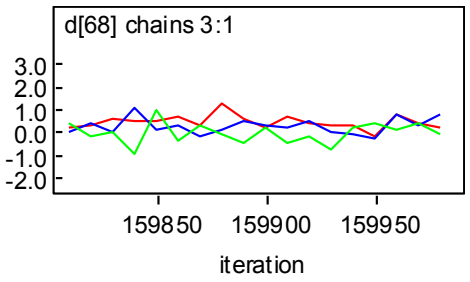

d[68] chains 3:1
3.0
2.0
1.0
0.0
-1.0
-2.0
159850
159900
159950
iteration

FIGURE D.7: Density plots obtained from hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence episodes

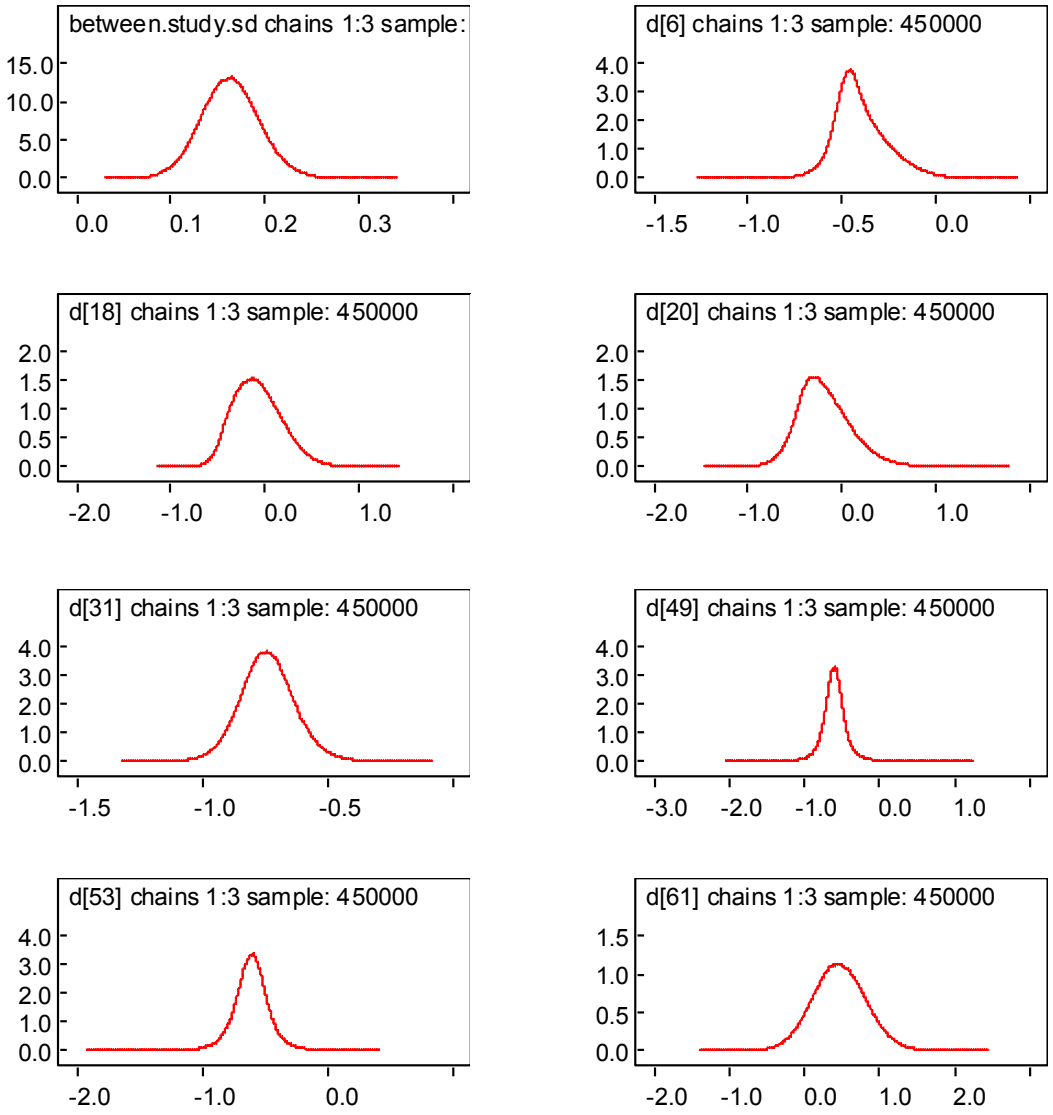
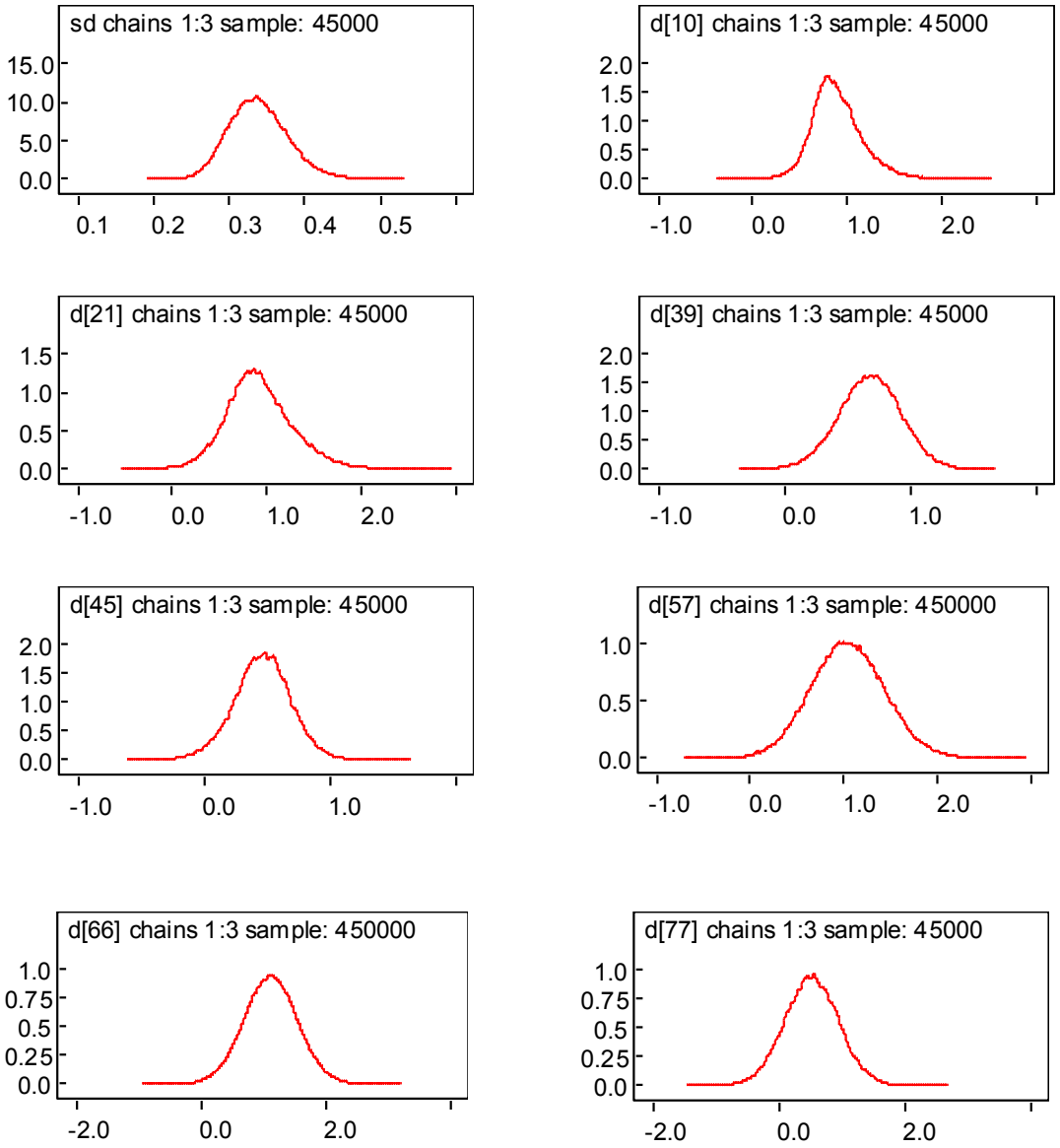


FIGURE D.8: Density plots obtained from hierarchical network meta-analysis evaluating the number of patients with adverse events



D.8 Sensitivity analysis

TABLE D.6: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for hierarchical network meta-analysis

Treament	**Code**	**Median difference† (95%CrI)**	**Median difference†† (95%CrI)**	**Median difference††† (95%CrI)**
Sacral nerve stimulation	[81]	-9.08 (-11.76,-6.52)	-9 (-11.55,-6.36)	-9 (-11.62,-6.5)
OnaBoNT-A 200U trigone sparing	[73]	-2.08 (-2.86,-1.45)	-2.07 (-2.86,-1.44)	-2.08 (-2.87,-1.46)
Electrostimulation + PFMT + BT	[97]	-2.16 (-3.11,-1.2)	-2.17 (-3.13,-1.22)	-2.16 (-3.13,-1.21)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-2.01 (-3.81,-0.24)	-2.01 (-3.8,-0.25)	-2.06 (-3.81,-0.2)
OnaBoNT-A 100u trigone sparing	[72]	-1.93 (-2.34,-1.52)	-1.93 (-2.33,-1.5)	-1.93 (-2.35,-1.51)
Solifenacin/trospium + placebo injection	[100]	-2.03 (-3.01,-1.05)	-2.02 (-3,-1.06)	-2.03 (-3,-1.05)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.89 (-2.61,-0.98)	-1.89 (-2.61,-1)	-1.89 (-2.61,-0.97)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.93 (-3.09,-0.4)	-1.92 (-3.04,-0.46)	-1.92 (-3,-0.37)
Tolterodine ER 4mg q.d + Neurostimulation	[96]	-1.29 (-1.69,-0.89)	-1.29 (-1.68,-0.89)	-1.29 (-1.69,-0.89)
Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	-1.05 (-1.94,-0.17)	-1.06 (-1.93,-0.17)	-1.06 (-1.96,-0.17)
Solifenacin ER 10mg q.d	[30]	-0.81 (-1.06,-0.61)	-0.82 (-1.06,-0.61)	-0.82 (-1.06,-0.61)
Tolterodine ER 2mg b.i.d + Estrogen 0.625mg 2xwk	[99]	-0.75 (-1.24,-0.25)	-0.75 (-1.23,-0.27)	-0.75 (-1.24,-0.26)
Solifenacin ER 5 - 10mg q.d	[31]	-0.74 (-0.95,-0.52)	-0.74 (-0.95,-0.52)	-0.74 (-0.95,-0.52)
Solifenacin ER 5mg q.d	[29]	-0.73 (-0.93,-0.51)	-0.73 (-0.93,-0.52)	-0.73 (-0.93,-0.51)
Solifenacin ER 5 - 15mg q.d	[34]	-0.73 (-1.01,-0.35)	-0.73 (-1.01,-0.35)	-0.73 (-1.01,-0.35)
Oxybutynin IR 3mg t.i.d	[19]	-0.71 (-1.07,-0.36)	-0.71 (-1.07,-0.37)	-0.71 (-1.07,-0.36)
Tolterodine ER 4mg q.d + BT	[87]	-0.69 (-1.37,-0.07)	-0.69 (-1.36,-0.06)	-0.7 (-1.36,-0.08)
Terodiline 25mg b.i.d	[28]	-0.67 (-1.22,-0.11)	-0.66 (-1.21,-0.12)	-0.66 (-1.2,-0.11)
Imidafenacin 0.25mg b.i.d	[37]	-0.67 (-1.32,-0.16)	-0.67 (-1.32,-0.15)	-0.67 (-1.31,-0.15)
Fesoterodine ER 8mg q.d	[26]	-0.66 (-0.86,-0.45)	-0.65 (-0.86,-0.45)	-0.66 (-0.86,-0.45)
Oxybutynin ER 10mg q.d	[8]	-0.66 (-1.15,-0.26)	-0.66 (-1.15,-0.27)	-0.66 (-1.15,-0.27)
Mirabegron IR 100mg b.i.d	[48]	-0.64 (-1.11,-0.38)	-0.64 (-1.11,-0.39)	-0.64 (-1.12,-0.39)
Darifenacin ER 30mg q.d	[38]	-0.64 (-1.23,-0.12)	-0.63 (-1.22,-0.1)	-0.63 (-1.22,-0.11)
Trospium chloride ER 60mg q.d	[44]	-0.63 (-1.05,-0.2)	-0.63 (-1.05,-0.21)	-0.63 (-1.06,-0.21)
Mirabegron ER 25mg q.d	[50]	-0.63 (-0.89,-0.42)	-0.63 (-0.87,-0.42)	-0.63 (-0.88,-0.42)
Propiverine ER 30mg q.d	[42]	-0.62 (-1.16,-0.26)	-0.63 (-1.17,-0.26)	-0.62 (-1.16,-0.26)
Mirabegron ER 100mg q.d	[52]	-0.62 (-0.81,-0.44)	-0.62 (-0.8,-0.44)	-0.62 (-0.81,-0.44)
Mirabegron IR 150mg b.i.d	[49]	-0.61 (-0.95,-0.26)	-0.6 (-0.93,-0.25)	-0.61 (-0.95,-0.26)
Mirabegron ER 200mg q.d	[53]	-0.61 (-0.91,-0.31)	-0.61 (-0.9,-0.32)	-0.61 (-0.91,-0.32)
Mirabegron ER 50mg q.d	[51]	-0.59 (-0.77,-0.42)	-0.59 (-0.76,-0.42)	-0.59 (-0.77,-0.42)
Solabegron IR 125mg b.i.d	[55]	-0.59 (-0.92,-0.25)	-0.59 (-0.92,-0.25)	-0.59 (-0.93,-0.24)
Cizolirtine citrate 400mg b.i.d	[57]	-0.59 (-1.12,-0.06)	-0.6 (-1.13,-0.07)	-0.59 (-1.13,-0.06)
Oxybutynin IR 5mg t.i.d	[7]	-0.57 (-0.95,-0.24)	-0.57 (-0.95,-0.25)	-0.57 (-0.96,-0.24)
Elocalcitol 75mg	[70]	-0.57 (-1.17,0.02)	-0.57 (-1.17,0.01)	-0.58 (-1.17,0)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	-0.51 (-1.03,-0.09)	-0.52 (-1.03,-0.09)	-0.51 (-1.02,-0.08)
Darifenacin ER 15mg q.d.	[40]	-0.51 (-0.95,-0.09)	-0.51 (-0.94,-0.08)	-0.51 (-0.94,-0.08)
Propiverine IR 15mg b.i.d	[43]	-0.51 (-0.91,-0.19)	-0.51 (-0.91,-0.19)	-0.51 (-0.91,-0.19)
Oxybutynin intravesically 5mg t.i.d	[14]	-0.49 (-1.13,0.01)	-0.5 (-1.15,0)	-0.49 (-1.14,0.01)
Tolterodine ER 4mg q.d	[4]	-0.5 (-0.6,-0.4)	-0.5 (-0.6,-0.4)	-0.5 (-0.6,-0.4)
Oxybutynin gel 56mg/day	[135]	-0.48 (-1.01,-0.03)	-0.48 (-1.01,-0.03)	-0.48 (-1,-0.02)
Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	-0.48 (-0.79,-0.17)	-0.48 (-0.78,-0.18)	-0.48 (-0.79,-0.17)
PFMT + BT	[89]	-0.49 (-1.03,0.05)	-0.49 (-1.02,0.05)	-0.49 (-1.04,0.05)
Fesoterodine ER 4mg q.d	[25]	-0.46 (-0.65,-0.29)	-0.46 (-0.64,-0.29)	-0.46 (-0.65,-0.29)
Tolterodine IR 2mg b.i.d	[5]	-0.45 (-0.57,-0.31)	-0.45 (-0.57,-0.31)	-0.45 (-0.57,-0.31)
Tolterodine IR 2mg b.i.d + PFMT	[95]	-0.45 (-1.07,0.17)	-0.45 (-1.07,0.17)	-0.45 (-1.07,0.17)
Oxybutynin chloride topical gel 1g/day	[13]	-0.44 (-0.83,-0.07)	-0.44 (-0.83,-0.07)	-0.44 (-0.83,-0.07)
Cizolirtine citrate 200mg b.i.d	[56]	-0.44 (-1.66,0.88)	-0.47 (-1.69,0.9)	-0.46 (-1.67,0.87)
Propiverine ER 20mg q.d	[41]	-0.43 (-0.62,-0.23)	-0.43 (-0.62,-0.23)	-0.43 (-0.62,-0.23)
Tolterodine IR 1mg b.i.d	[6]	-0.43 (-0.64,-0.09)	-0.43 (-0.64,-0.09)	-0.43 (-0.64,-0.08)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.43 (-1.5,0.48)	-0.45 (-1.5,0.46)	-0.46 (-1.54,0.46)
Imidafenacin 0.1mg b.i.d	[36]	-0.41 (-0.71,-0.13)	-0.41 (-0.7,-0.12)	-0.41 (-0.71,-0.12)
Elocalcitol 150mg	[69]	-0.42 (-1.03,0.19)	-0.43 (-1.03,0.18)	-0.43 (-1.03,0.18)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.4 (-0.83,0.03)	-0.4 (-0.83,0.04)	-0.4 (-0.83,0.03)
Oxybutynin trandermal 3.9mg/day	[10]	-0.39 (-0.67,-0.11)	-0.39 (-0.66,-0.11)	-0.39 (-0.67,-0.11)

TABLE D.6: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for hierarchical network meta-analysis (cont.)

Oxybutynin vaginal ring 6mg q.d	[17]	-0.39 (-0.81,0.03)	-0.4 (-0.81,0.03)	-0.39 (-0.81,0.03)
Darifenacin ER 7.5mg q.d	[39]	-0.37 (-0.87,0.19)	-0.37 (-0.86,0.19)	-0.37 (-0.87,0.19)
Oxbutynin patch 73.5mg	[15]	-0.34 (-0.63,-0.04)	-0.35 (-0.63,-0.05)	-0.35 (-0.63,-0.05)
Oxybutynin ER 15mg q.d	[9]	-0.34 (-0.85,0.22)	-0.35 (-0.86,0.22)	-0.35 (-0.86,0.21)
Oxybutynin ER 5 - 30mg/day	[22]	-0.33 (-0.72,0.14)	-0.33 (-0.71,0.13)	-0.34 (-0.73,0.12)
Fesoterodine ER 4 - 8mg q.d	[27]	-0.33 (-0.56,-0.09)	-0.33 (-0.56,-0.09)	-0.32 (-0.56,-0.09)
Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	-0.34 (-0.89,0.22)	-0.34 (-0.89,0.21)	-0.34 (-0.91,0.22)
Oxybutynin gel 84mg/day	[134]	-0.32 (-0.77,0.18)	-0.33 (-0.77,0.17)	-0.33 (-0.77,0.17)
Bladder Training (BT)/Behaviour Therapy	[85]	-0.33 (-0.74,0.11)	-0.32 (-0.72,0.11)	-0.33 (-0.73,0.1)
Oxybutynin IR 5 - 20mg	[23]	-0.31 (-0.82,0.34)	-0.32 (-0.83,0.34)	-0.32 (-0.82,0.33)
Imidafenacin 0.05mg b.i.d	[35]	-0.31 (-0.77,0.19)	-0.31 (-0.77,0.19)	-0.31 (-0.77,0.19)
Oxybutynin transdermal 1.3mg/day	[11]	-0.29 (-0.69,0.2)	-0.29 (-0.69,0.19)	-0.29 (-0.68,0.2)
Duloxetine 40mg b.i.d	[65]	-0.26 (-0.76,0.24)	-0.26 (-0.75,0.24)	-0.26 (-0.76,0.24)
Oxybutynin ER 2.5mg q.d + BT	[92]	-0.25 (-1.22,0.71)	-0.25 (-1.23,0.7)	-0.25 (-1.21,0.7)
Oxybutynin ER 2.5mg q.d	[20]	-0.22 (-0.68,0.44)	-0.22 (-0.67,0.43)	-0.22 (-0.67,0.43)
Solabegron IR 50mg b.i.d	[54]	-0.22 (-0.56,0.12)	-0.21 (-0.55,0.11)	-0.21 (-0.56,0.12)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.21 (-0.84,0.43)	-0.22 (-0.84,0.44)	-0.21 (-0.84,0.44)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.17 (-0.46,0.12)	-0.17 (-0.45,0.12)	-0.17 (-0.46,0.12)
Oxybutynin transdermal 2.6mg/day	[12]	-0.15 (-0.56,0.41)	-0.14 (-0.55,0.4)	-0.15 (-0.56,0.4)
Tarafenacin 0.4mg q.d	[82]	-0.11 (-0.82,0.58)	-0.11 (-0.82,0.59)	-0.1 (-0.81,0.6)
ZD0947IL 25mg/day	[58]	-0.11 (-0.96,0.76)	-0.1 (-0.96,0.76)	-0.1 (-0.97,0.77)
Oxybutynin IR 5mg b.i.d	[18]	-0.1 (-0.52,0.46)	-0.09 (-0.52,0.46)	-0.11 (-0.52,0.46)
Lipo-BoNTA 200U	[138]	-0.06 (-0.94,0.82)	-0.06 (-0.95,0.82)	-0.05 (-0.94,0.84)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.02 (-0.57,0.54)	-0.02 (-0.58,0.54)	-0.03 (-0.58,0.53)
Placebo	[1]	NA	NA	NA
Pregabalin 150mg b.i.d	[62]	0 (-0.88,0.9)	0.01 (-0.87,0.88)	0 (-0.89,0.88)
Estradiol 25mg	[68]	0 (-0.38,0.38)	0 (-0.37,0.37)	0 (-0.38,0.38)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	0.01 (-0.5,0.51)	0.01 (-0.49,0.52)	0.01 (-0.5,0.52)
Electrostimulation	[80]	0.02 (-0.32,0.36)	0.02 (-0.32,0.37)	0.02 (-0.32,0.37)
Percutaneous tibial nerve stimulation	[83]	0.01 (-1.03,1.15)	0 (-1.02,1.17)	0.01 (-1.05,1.19)
Control	[2]	0.04 (-0.81,0.86)	0.05 (-0.79,0.89)	0.04 (-0.79,0.87)
Tarafenacin 0.2mg q.d	[90]	0.07 (-0.65,0.8)	0.06 (-0.65,0.8)	0.08 (-0.64,0.81)
ONO-8539 100mg b.i.d	[60]	0.06 (-0.71,0.76)	0.06 (-0.72,0.76)	0.05 (-0.72,0.77)
Reflexology	[71]	0.05 (-0.91,0.97)	0.05 (-0.89,0.98)	0.04 (-0.91,0.98)
Trospium chloride IR 15mg t.i.d	[46]	0.21 (-0.66,1.07)	0.19 (-0.64,1.08)	0.19 (-0.68,1.08)
Flavoxate chloride 200mg q.d	[64]	0.32 (-0.36,1.01)	0.32 (-0.36,1)	0.33 (-0.36,1.02)
Emepronium bromide ER 200mg q.d	[63]	0.34 (-0.35,1.01)	0.33 (-0.34,1)	0.34 (-0.33,1.03)
Vaginal oestrogen cream 1.25mg/day	[132]	0.32 (-0.19,0.83)	0.32 (-0.19,0.83)	0.32 (-0.2,0.83)
Oxybutynin ER 5-30mg q.d + BT	[86]	0.39 (-0.62,1.47)	0.39 (-0.62,1.46)	0.39 (-0.63,1.46)
ONO-8539 30mg b.i.d	[59]	0.48 (-0.21,1.23)	0.49 (-0.19,1.21)	0.48 (-0.2,1.22)
ONO-8539 300mg b.i.d	[61]	0.45 (-0.24,1.17)	0.45 (-0.23,1.17)	0.45 (-0.23,1.17)
Sham therapy	[3]	0.61 (-0.48,1.86)	0.6 (-0.46,1.88)	0.62 (-0.49,1.88)
Resiniferatoxin 50nM	[67]	0.59 (-1.02,2.17)	0.56 (-1.08,2.18)	0.56 (-1.05,2.24)

† Between-study standard deviation based on an Uniform(0,5) prior distribution on the standard deviation scale

†† Between-study standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Between-study standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE D.7: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on class-specific variance parameters for hierarchical network meta-analysis

Treatment	**Code**	**Median difference† (95% CrI)**	**Median difference†† (95% CrI)**	**Median difference††† (95% CrI)**
Sacral nerve stimulation	[81]	-8.72 (-11.33,-6.09)	-8.9 (-11.41,-6.26)	-8.92 (-11.49,-6.19)
OnaBoNT-A 200U trigone sparing	[73]	-2.3 (-3.16,-1.42)	-2.01 (-2.65,-1.49)	-2.04 (-2.77,-1.47)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-2.2 (-4.06,-0.36)	-2.03 (-3.78,-0.26)	-1.99 (-3.84,-0.18)
Anticholinergic + saline injection	[100]	-1.97 (-2.96,-1.01)	-2.04 (-3,-1.07)	-2.02 (-2.99,-1.04)
Electrostimulation + PFMT + BT	[97]	-1.93 (-2.94,-0.91)	-2.2 (-3.2,-1.24)	-2.18 (-3.12,-1.23)
OnaBoNT-A 100u trigone sparing	[72]	-1.88 (-2.31,-1.45)	-1.94 (-2.36,-1.53)	-1.93 (-2.34,-1.53)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.63 (-2.73,-0.54)	-1.93 (-2.51,-1.22)	-1.91 (-2.57,-1.08)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.39 (-4,1.06)	-1.95 (-2.67,-1.06)	-1.94 (-2.87,-0.78)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.29 (-1.7,-0.89)	-1.28 (-1.7,-0.87)	-1.29 (-1.69,-0.88)
Oxybutynin intravesically 5mg t.i.d	[14]	-1.19 (-2.49,0.1)	-0.47 (-1.04,-0.03)	-0.49 (-1.13,0)
Mirabegron IR 100mg b.i.d	[48]	-1.12 (-1.98,-0.25)	-0.63 (-0.98,-0.39)	-0.64 (-1.09,-0.37)
Oxybutynin ER 10mg q.d	[8]	-0.98 (-1.55,-0.42)	-0.6 (-1.09,-0.25)	-0.66 (-1.15,-0.27)
Oxybutynin IR 3mg t.i.d	[19]	-0.9 (-1.28,-0.52)	-0.65 (-1.03,-0.33)	-0.7 (-1.07,-0.36)
Solifenacin ER 10mg q.d	[30]	-0.88 (-1.14,-0.63)	-0.8 (-1.03,-0.6)	-0.81 (-1.06,-0.61)
Imidafenacin 0.25mg b.i.d	[37]	-0.82 (-1.43,-0.21)	-0.55 (-1.17,-0.13)	-0.63 (-1.27,-0.16)
Propiverine ER 30mg q.d	[42]	-0.79 (-1.31,-0.28)	-0.54 (-1.02,-0.25)	-0.6 (-1.13,-0.26)
Tolterodine ER 4mg q.d + BT	[87]	-0.78 (-1.45,-0.12)	-0.58 (-1.25,-0.01)	-0.64 (-1.31,-0.04)
Trospium chloride IR 15mg t.i.d + Physiotherapy	[91]	-0.78 (-1.71,0.16)	-1.18 (-2.05,-0.27)	-1.12 (-1.99,-0.24)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	-0.75 (-1.45,-0.01)	-0.49 (-0.96,-0.12)	-0.51 (-1.03,-0.09)
Darifenacin ER 30mg q.d	[38]	-0.74 (-1.36,-0.12)	-0.57 (-1.1,-0.1)	-0.62 (-1.2,-0.11)
Tolterodine ER 2mg b.i.d + Estrogen 0.625mg 2xwk	[99]	-0.75 (-1.24,-0.25)	-0.75 (-1.25,-0.25)	-0.75 (-1.24,-0.26)
Solifenacin ER 5 - 10mg q.d	[31]	-0.73 (-1,-0.45)	-0.74 (-0.95,-0.53)	-0.74 (-0.95,-0.52)
Solifenacin ER 5mg q.d	[29]	-0.71 (-0.96,-0.46)	-0.73 (-0.93,-0.53)	-0.73 (-0.93,-0.52)
Oxybutynin IR 5mg t.i.d	[7]	-0.71 (-1.18,-0.25)	-0.54 (-0.92,-0.24)	-0.57 (-0.95,-0.24)
Fesoterodine ER 8mg q.d	[26]	-0.69 (-0.89,-0.5)	-0.63 (-0.84,-0.43)	-0.65 (-0.86,-0.45)
Oxybutynin gel 56mg/day	[135]	-0.69 (-1.59,0.18)	-0.46 (-0.95,-0.07)	-0.48 (-1,-0.03)
Mirabegron ER 25mg q.d	[50]	-0.67 (-0.99,-0.35)	-0.62 (-0.87,-0.41)	-0.63 (-0.88,-0.41)
Terodiline 25mg b.i.d	[28]	-0.66 (-1.21,-0.1)	-0.66 (-1.22,-0.1)	-0.66 (-1.2,-0.1)
Trospium chloride ER 60mg q.d	[44]	-0.66 (-1.09,-0.24)	-0.58 (-1.02,-0.16)	-0.6 (-1.02,-0.18)
Cizilirtine citrate 400mg b.i.d	[57]	-0.63 (-1.17,-0.08)	-0.57 (-1.11,-0.03)	-0.58 (-1.12,-0.05)
Mirabegron ER 100mg q.d	[52]	-0.62 (-0.83,-0.41)	-0.61 (-0.8,-0.43)	-0.62 (-0.8,-0.43)
Solifenacin ER 5 - 15mg q.d	[34]	-0.61 (-1.11,-0.1)	-0.74 (-0.99,-0.43)	-0.73 (-1,-0.36)
Solabegron IR 125mg b.i.d	[55]	-0.6 (-0.94,-0.26)	-0.55 (-0.9,-0.19)	-0.58 (-0.91,-0.23)
Elocalcitol 75mg	[70]	-0.6 (-1.2,0.01)	-0.55 (-1.13,0.01)	-0.56 (-1.15,0.01)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.6 (-1.66,0.46)	-0.25 (-1.26,0.46)	-0.32 (-1.32,0.48)
Propiverine IR 15mg b.i.d	[43]	-0.57 (-1,-0.16)	-0.49 (-0.83,-0.2)	-0.51 (-0.9,-0.2)
Mirabegron ER 200mg q.d	[53]	-0.57 (-1.13,-0.01)	-0.61 (-0.87,-0.35)	-0.61 (-0.9,-0.31)
Mirabegron ER 50mg q.d	[51]	-0.57 (-0.76,-0.38)	-0.59 (-0.77,-0.41)	-0.59 (-0.77,-0.41)
Darifenacin ER 15mg q.d.	[40]	-0.51 (-0.97,-0.05)	-0.51 (-0.91,-0.09)	-0.52 (-0.94,-0.09)
Mirabegron IR 150mg b.i.d	[49]	-0.51 (-1.57,0.54)	-0.6 (-0.89,-0.32)	-0.6 (-0.93,-0.25)
Oxybutynin chloride topical gel 1g/day	[13]	-0.5 (-1.02,0.02)	-0.43 (-0.8,-0.09)	-0.44 (-0.83,-0.07)
Tolterodine ER 4mg q.d	[4]	-0.51 (-0.61,-0.4)	-0.49 (-0.59,-0.39)	-0.5 (-0.6,-0.4)
Fesoterodine ER 4mg q.d	[25]	-0.47 (-0.66,-0.29)	-0.46 (-0.64,-0.29)	-0.46 (-0.64,-0.29)
Tolterodine 2mg + Pilocarpine 9mg b.i.d	[101]	-0.48 (-0.79,-0.17)	-0.48 (-0.8,-0.16)	-0.48 (-0.79,-0.17)
Tolterodine IR 2mg b.i.d	[5]	-0.45 (-0.6,-0.3)	-0.45 (-0.57,-0.32)	-0.45 (-0.58,-0.31)
Tolterodine IR 2mg b.i.d + PFMT	[95]	-0.45 (-1.07,0.18)	-0.45 (-1.08,0.18)	-0.45 (-1.07,0.18)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.44 (-1.11,0.23)	-0.4 (-0.8,-0.01)	-0.4 (-0.82,0.03)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.43 (-1.08,0.22)	-0.4 (-0.79,0)	-0.39 (-0.82,0.03)
Propiverine ER 20mg q.d	[41]	-0.41 (-0.61,-0.2)	-0.44 (-0.64,-0.25)	-0.43 (-0.63,-0.24)
Imidafenacin 0.1mg b.i.d	[36]	-0.4 (-0.7,-0.11)	-0.42 (-0.72,-0.13)	-0.42 (-0.71,-0.14)
Cizilirtine citrate 200mg b.i.d	[56]	-0.4 (-1.83,1.03)	-0.52 (-1.41,0.59)	-0.5 (-1.52,0.62)
Elocalcitol 150mg	[69]	-0.4 (-1.03,0.22)	-0.47 (-1.05,0.12)	-0.45 (-1.04,0.15)
Oxybutynin trandermal 3.9mg/day	[10]	-0.33 (-0.67,0)	-0.4 (-0.66,-0.13)	-0.4 (-0.67,-0.11)
Tolterodine IR 1mg b.i.d	[6]	-0.33 (-0.69,0.02)	-0.44 (-0.63,-0.16)	-0.43 (-0.64,-0.1)
Oxbutynin patch 73.5mg	[15]	-0.31 (-0.67,0.05)	-0.36 (-0.63,-0.07)	-0.35 (-0.63,-0.05)

TABLE D.7: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on class-specific variance parameters for hierarchical network meta-analysis (cont.)

Fesoterodine ER 4 - 8mg q.d	[27]	-0.28 (-0.52,-0.05)	-0.37 (-0.58,-0.12)	-0.34 (-0.56,-0.09)
Oxybutynin gel 84mg/day	[134]	-0.29 (-1.13,0.52)	-0.35 (-0.74,0.12)	-0.33 (-0.77,0.17)
Darifenacin ER 7.5mg q.d	[39]	-0.27 (-0.84,0.3)	-0.43 (-0.87,0.09)	-0.4 (-0.88,0.15)
Oxybutynin ER 15mg q.d	[9]	-0.28 (-1.41,0.86)	-0.36 (-0.81,0.15)	-0.35 (-0.85,0.21)
Duloxetine 40mg b.i.d	[65]	-0.26 (-0.75,0.25)	-0.26 (-0.77,0.24)	-0.26 (-0.76,0.24)
PFMT + BT	[89]	-0.26 (-0.87,0.37)	-0.53 (-1.07,0.02)	-0.51 (-1.04,0.03)
Imidafenacin 0.05mg b.i.d	[35]	-0.24 (-0.77,0.27)	-0.36 (-0.76,0.11)	-0.33 (-0.77,0.16)
Solabegron IR 50mg b.i.d	[54]	-0.2 (-0.54,0.14)	-0.25 (-0.61,0.1)	-0.23 (-0.58,0.11)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.14 (-0.81,0.52)	-0.32 (-0.89,0.34)	-0.27 (-0.86,0.37)
Tarafenacin 0.4mg q.d	[82]	-0.13 (-0.85,0.61)	-0.06 (-0.76,0.63)	-0.07 (-0.79,0.61)
ONO-8539 100mg b.i.d	[60]	-0.11 (-0.85,0.63)	0.2 (-0.54,0.85)	0.1 (-0.64,0.78)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.08 (-0.39,0.23)	-0.22 (-0.49,0.08)	-0.18 (-0.46,0.11)
ZD0947IL 25mg/day	[58]	-0.1 (-0.94,0.76)	-0.1 (-0.97,0.76)	-0.1 (-0.97,0.75)
Pelvic Floor Muscle Training (PFMT)/Physiotherapy	[84]	-0.08 (-0.72,0.55)	-0.42 (-0.98,0.15)	-0.38 (-0.93,0.18)
Lipo-BoNTA 200U	[138]	-0.06 (-0.96,0.83)	-0.07 (-0.95,0.85)	-0.06 (-0.95,0.82)
Oxybutynin transdermal 1.3mg/day	[11]	-0.03 (-0.66,0.61)	-0.33 (-0.68,0.13)	-0.3 (-0.69,0.19)
Placebo	[1]	NA	NA	NA
Estradiol 25mg	[68]	0 (-0.38,0.38)	0 (-0.39,0.39)	0 (-0.37,0.38)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	0 (-0.57,0.56)	-0.06 (-0.6,0.49)	-0.04 (-0.58,0.51)
Pregabalin 150mg b.i.d	[62]	0 (-0.88,0.9)	0.01 (-0.89,0.9)	0 (-0.89,0.89)
Bladder Training (BT)/Behaviour Therapy	[85]	0.02 (-0.59,0.62)	-0.35 (-0.74,0.07)	-0.33 (-0.73,0.1)
Oxybutynin ER 5 - 30mg/day	[22]	0.09 (-0.62,0.78)	-0.37 (-0.73,0.06)	-0.35 (-0.74,0.12)
Electrostimulation + Estrogen	[133]	0.07 (-0.45,0.59)	-0.02 (-0.54,0.49)	0 (-0.51,0.51)
Electrostimulation	[80]	0.08 (-0.28,0.44)	-0.01 (-0.35,0.33)	0.01 (-0.34,0.35)
Tarafenacin 0.2mg q.d	[90]	0.11 (-0.63,0.86)	0.04 (-0.66,0.74)	0.07 (-0.64,0.78)
Percutaneous tibial nerve stimulation	[83]	0.18 (-1.16,1.54)	-0.08 (-0.88,0.63)	-0.06 (-0.91,0.8)
Oxybutynin ER 2.5mg q.d + BT	[92]	0.22 (-0.94,1.4)	-0.14 (-1.09,0.73)	-0.16 (-1.11,0.76)
Oxybutynin transdermal 2.6mg/day	[12]	0.29 (-0.37,0.95)	-0.22 (-0.57,0.33)	-0.15 (-0.56,0.4)
Oxybutynin IR 5mg b.i.d	[18]	0.35 (-0.3,0.99)	-0.18 (-0.54,0.38)	-0.11 (-0.52,0.46)
Control	[2]	0.33 (-0.67,1.31)	0.15 (-0.68,0.89)	0.12 (-0.68,0.9)
Emepronium bromide ER 200mg q.d	[63]	0.34 (-0.33,1.01)	0.34 (-0.35,1.02)	0.34 (-0.33,1.03)
Flavoxate chloride 200mg q.d	[64]	0.33 (-0.35,1.01)	0.34 (-0.37,1.02)	0.34 (-0.34,1.02)
Reflexology	[71]	0.33 (-0.76,1.41)	0.15 (-0.79,1.03)	0.11 (-0.8,1.02)
Estrogen	[132]	0.38 (-0.15,0.91)	0.29 (-0.23,0.8)	0.31 (-0.21,0.82)
Oxybutynin IR 5 - 20mg	[23]	0.46 (-0.9,1.84)	-0.35 (-0.79,0.23)	-0.32 (-0.82,0.33)
Oxybutynin ER 2.5mg q.d	[20]	0.52 (-0.5,1.55)	-0.27 (-0.65,0.34)	-0.22 (-0.66,0.46)
Trospium chloride IR 15mg t.i.d	[46]	0.52 (-0.41,1.41)	-0.06 (-0.79,0.96)	0.06 (-0.72,0.96)
ONO-8539 300mg b.i.d	[61]	0.51 (-0.23,1.25)	0.41 (-0.22,1.09)	0.43 (-0.22,1.13)
ONO-8539 30mg b.i.d	[59]	0.58 (-0.18,1.33)	0.43 (-0.19,1.11)	0.46 (-0.2,1.15)
Resiniferatoxin 50nM	[67]	0.58 (-1.08,2.28)	0.61 (-1,2.24)	0.62 (-1.06,2.22)
Sham therapy	[3]	0.83 (-0.62,2.27)	0.45 (-0.36,1.33)	0.5 (-0.38,1.48)
Oxybutynin ER 5-30mg q.d + BT	[86]	0.92 (-0.31,2.12)	0.19 (-0.69,1.27)	0.26 (-0.68,1.27)

† Class-specific standard deviation based on an Uniform(0,5) prior distribution on the standard deviation scale

†† Class-specific standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Class-specific standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE D.8: Sensitivity analyses assessing the impact of different choices of prior distributions on variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events

Treatment	**Code**	**Odds Ratio† (95%CrI)**	**Odds Ratio†† (95%CrI)**	**Odds Ratio††† (95%CrI)**
Darifenacin ER 30mg q.d	[38]	6.29 (3.49,11.44)	6.29 (3.53,11.46)	6.28 (3.49,11.56)
Imidafenacin 0.25mg b.i.d	[37]	4.62 (2.27,9.36)	4.32 (2.17,8.65)	4.36 (2.16,8.64)
Solifenacin ER 20mg q.d	[33]	3.28 (1.57,7.43)	3.58 (1.97,7.74)	3.62 (1.98,7.94)
Tarafenacin 0.4mg q.d	[82]	3.43 (1.56,7.7)	3.45 (1.58,7.71)	3.42 (1.57,7.65)
Terodiline 25mg b.i.d	[28]	3.09 (1.15,8.77)	3.08 (1.15,8.97)	3.1 (1.15,8.74)
Propiverine IR 45mg t.i.d	[118]	2.5 (1.63,7.87)	3.08 (1.89,8.52)	3.05 (1.82,8.26)
Propiverine ER 60mg q.d	[119]	2.51 (1.62,7.96)	3.04 (1.98,8.31)	3.04 (1.92,8.14)
Duloxetine 60mg b.i.d	[66]	3.05 (1.6,5.88)	3.05 (1.6,5.91)	3.05 (1.6,5.91)
Oxybutynin IR 5mg t.i.d	[7]	2.56 (1.76,3.66)	2.91 (2.23,4.27)	2.91 (2.25,4.27)
Cizilirtine citrate 400mg b.i.d	[57]	2.44 (1.22,4.99)	2.7 (1.4,5.49)	2.72 (1.37,5.38)
Oxbutynin patch 73.5mg	[15]	2.67 (1.97,3.77)	2.67 (1.99,3.7)	2.67 (2,3.67)
Oxybutynin vaginal ring 6mg q.d	[17]	2.48 (1.57,3.53)	2.64 (1.83,3.66)	2.66 (1.85,3.66)
Oxybutynin intravesically 5mg t.i.d	[14]	2.57 (1.45,4.47)	2.61 (1.56,4.34)	2.63 (1.61,4.33)
Oxybutynin IR 5mg b.i.d	[18]	2.97 (2.22,4.46)	2.69 (2.09,3.65)	2.69 (2.1,3.64)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	2.49 (1.7,3.52)	2.55 (1.78,3.55)	2.57 (1.8,3.5)
Darifenacin ER 15mg q.d.	[40]	2.33 (1.54,3.54)	2.47 (1.73,3.67)	2.47 (1.72,3.68)
Oxybutynin chloride topical gel 1g/day	[13]	2.36 (1.45,3.3)	2.44 (1.52,3.33)	2.47 (1.54,3.31)
Fesoterodine ER 8mg q.d	[26]	2.45 (1.87,3.18)	2.4 (1.88,3.08)	2.4 (1.88,3.1)
Oxybutynin vaginal ring 4mg q.d	[16]	2.38 (1.44,3.35)	2.34 (1.44,3.14)	2.38 (1.45,3.15)
Fesoterodine IR 12mg b.i.d	[122]	1.98 (1.04,3.94)	2.29 (1.44,4.45)	2.27 (1.45,4.42)
Propiverine IR 30mg b.i.d	[117]	2.25 (1.41,5.5)	2.31 (1.45,5.19)	2.31 (1.44,4.99)
Propiverine ER 30mg q.d	[42]	1.96 (1.22,2.84)	2.31 (1.82,3.14)	2.31 (1.77,3.16)
Solifenacin ER 5 - 15mg q.d	[34]	2.19 (1.32,3.71)	2.2 (1.32,3.74)	2.2 (1.32,3.77)
Propiverine IR 15mg t.i.d	[116]	2.18 (1.46,3.56)	2.21 (1.43,3.53)	2.21 (1.42,3.52)
Propiverine ER 20mg q.d	[41]	2.15 (1.7,2.8)	2.11 (1.7,2.66)	2.11 (1.67,2.68)
Solifenacin ER 10mg q.d	[30]	1.61 (1.06,2.46)	2.01 (1.48,2.82)	2.01 (1.48,2.84)
Tolterodine IR 4mg b.i.d	[140]	1.45 (0.97,5.75)	2 (1.22,7.51)	1.99 (1.21,6.97)
OnaBoNTA 300u trigone sparing	[76]	1.75 (0.94,3.61)	1.92 (1.06,3.79)	1.86 (1.04,3.5)
Darifenacin ER 7.5mg q.d	[39]	2.09 (1.35,3.23)	1.94 (1.32,2.83)	1.93 (1.3,2.83)
Cizilirtine citrate 200mg b.i.d	[56]	2.33 (0.88,6.25)	1.95 (0.79,4.26)	1.93 (0.78,4.03)
Trospium chloride IR 15mg t.i.d	[46]	1.61 (0.6,2.88)	1.88 (1.32,3.43)	1.89 (1.34,3.45)
Fesoterodine IR 8mg b.id	[121]	1.85 (0.9,3.28)	1.92 (1.1,3.03)	1.9 (1.1,3.02)
Trospium chloride IR 45mg t.i.d	[47]	1.83 (1.22,3.1)	1.85 (1.27,3.09)	1.86 (1.27,3.1)
Fesoterodine ER 4 - 8mg q.d	[27]	1.7 (1.28,2.21)	1.87 (1.54,2.3)	1.88 (1.53,2.3)
Propiverine IR 15mg b.i.d	[43]	1.91 (1.26,2.65)	1.87 (1.36,2.41)	1.86 (1.37,2.43)
Solifenacin ER 5 - 10mg q.d	[31]	2.18 (1.4,3.43)	1.74 (1.28,2.39)	1.74 (1.29,2.4)
OnaBoNT-A 200U trigone sparing	[73]	1.56 (0.89,2.73)	1.7 (0.98,3.12)	1.65 (0.97,2.86)
Fesoterodine ER 4mg q.d	[25]	1.82 (1.45,2.26)	1.7 (1.39,2.07)	1.7 (1.38,2.07)
Trospium chloride ER 60mg q.d	[44]	1.64 (1.16,2.28)	1.66 (1.19,2.28)	1.68 (1.19,2.3)
Solifenacin ER 5 - 10mg q.d + BT	[77]	2.07 (1.04,4.13)	1.66 (0.9,3.04)	1.65 (0.89,3.06)
ZD0947IL 25mg/day	[58]	1.61 (0.78,3.39)	1.6 (0.77,3.36)	1.61 (0.77,3.43)
Fesoterodine IR 4mg b.i.d	[120]	1.88 (0.93,3.38)	1.63 (0.8,2.44)	1.63 (0.81,2.45)
Imidafenacin 0.1mg b.i.d	[36]	1.63 (1.16,2.3)	1.62 (1.17,2.28)	1.62 (1.16,2.28)
Trospium chloride IR 20mg b.i.d	[45]	1.74 (1.24,2.52)	1.63 (1.14,2.26)	1.64 (1.15,2.25)
OnaBoNTA 150u trigone sparing	[75]	1.75 (0.93,3.63)	1.59 (0.9,2.94)	1.53 (0.89,2.65)
Tarafenacin 0.2mg q.d	[90]	1.47 (0.68,3.13)	1.45 (0.69,3.09)	1.46 (0.67,3.11)
OnaBoNT-A 100u trigone sparing	[72]	1.73 (0.94,3.45)	1.46 (0.81,2.76)	1.41 (0.81,2.48)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1.3 (0.72,2.38)	1.38 (0.8,2.51)	1.37 (0.77,2.5)
Pregabalin 150mg b.i.d	[62]	1.38 (0.74,2.56)	1.37 (0.73,2.58)	1.37 (0.73,2.55)
Tolterodine IR 2mg b.i.d	[5]	1.32 (1.05,1.71)	1.36 (1.07,1.77)	1.36 (1.06,1.77)
Netupitant 200mg q.d	[112]	1.1 (0.45,2.89)	1.26 (0.55,3.31)	1.34 (0.55,3.41)
OnaBoNTA 50u trigone sparing	[74]	1.56 (0.78,2.99)	1.29 (0.64,2.52)	1.24 (0.63,2.26)
Solifenacin ER 5mg q.d	[29]	1.23 (0.87,1.75)	1.23 (0.88,1.68)	1.23 (0.88,1.68)
Imidafenacin 0.05mg b.i.d	[35]	1.48 (0.77,2.83)	1.23 (0.7,1.91)	1.23 (0.69,1.9)
Tolterodine ER 4mg q.d	[4]	1.24 (1.08,1.42)	1.24 (1.07,1.42)	1.23 (1.06,1.42)

TABLE D.8: Sensitivity analyses assessing the impact of different choices of prior distributions on variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events (cont.)

Mirabegron ER 200mg q.d	[53]	1.1 (0.86,1.53)	1.22 (0.99,1.74)	1.22 (0.98,1.75)
Anticholinergic (not specified) + saline injection	[100]	1.42 (0.56,3.75)	1.2 (0.49,3.02)	1.14 (0.47,2.82)
ONO-8539 300mg b.i.d	[61]	1.11 (0.63,2.05)	1.21 (0.71,2.18)	1.2 (0.71,2.19)
Mirabegron IR 150mg b.i.d	[49]	1.08 (0.8,1.5)	1.15 (0.84,1.69)	1.14 (0.83,1.7)
Tolterodine IR 2mg b.i.d + BT	[93]	1.13 (0.65,1.97)	1.16 (0.67,2.03)	1.16 (0.67,2.05)
Mirabegron ER 100mg q.d	[52]	1.09 (0.9,1.35)	1.13 (0.94,1.37)	1.12 (0.94,1.37)
Tolterodine IR 1mg b.i.d	[6]	1.17 (0.79,1.57)	1.13 (0.8,1.5)	1.14 (0.79,1.51)
Lipo-BoNTA 200U	[138]	1.11 (0,245.3)	1.1 (0,434.7)	1.09 (0,291.5)
Mirabegron ER 50mg q.d	[51]	1.06 (0.9,1.26)	1.07 (0.9,1.25)	1.06 (0.89,1.25)
Placebo	[1]	NA	NA	NA
Mirabegron IR 100mg b.i.d	[48]	1.06 (0.71,1.38)	1.01 (0.61,1.31)	1 (0.6,1.34)
Mirabegron ER 25mg q.d	[50]	1.07 (0.87,1.34)	1.01 (0.81,1.21)	1 (0.8,1.19)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1.05 (0.56,1.91)	0.98 (0.54,1.75)	0.97 (0.53,1.73)
Solabegron IR 125mg b.i.d	[55]	0.87 (0.43,1.77)	0.98 (0.5,1.95)	0.99 (0.52,1.97)
Netupitant 100mg q.d	[111]	0.94 (0.37,2.44)	0.95 (0.42,2.4)	1.01 (0.43,2.44)
ONO-8539 100mg b.i.d	[60]	0.88 (0.48,1.55)	0.96 (0.57,1.59)	0.95 (0.58,1.58)
Tolterodine IR 0.5mg b.i.d	[141]	1.19 (0.49,2.01)	0.88 (0.37,1.36)	0.88 (0.36,1.32)
ONO-8539 30mg b.i.d	[59]	0.96 (0.54,1.7)	0.83 (0.47,1.4)	0.82 (0.47,1.38)
Percutaneous tibial nerve stimulation	[83]	0.72 (0.28,1.84)	0.76 (0.3,1.94)	0.73 (0.28,1.83)
Solifenacin ER 2.5mg q.d	[32]	0.82 (0.3,1.82)	0.76 (0.3,1.35)	0.77 (0.3,1.36)
Solabegron IR 50mg b.i.d	[54]	0.89 (0.44,1.8)	0.79 (0.4,1.55)	0.8 (0.41,1.56)
Netupitant 50mg q.d	[110]	0.89 (0.34,2.28)	0.73 (0.29,1.89)	0.77 (0.29,1.97)
Electrostimulation	[80]	0.34 (0.02,2.01)	0.18 (0.01,1.88)	0.34 (0.02,2.07)
Sham therapy	[3]	0.13 (0.01,0.83)	0.07 (0,0.68)	0.13 (0.01,0.82)
Control	[2]	0.09 (0,1.72)	0.06 (0,2.13)	0.09 (0,1.48)
Bladder Training (BT)/Behaviour Therapy	[85]	0.04 (0,0.28)	0 (0,0.14)	0.04 (0,0.28)

† Between-study standard deviation based on an Uniform(0,2) prior distribution on the standard deviation scale

†† Between-study standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Between-study standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE D.9: Sensitivity analyses assessing the impact of different choices of prior distributions on class-specific variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events

Treatment	**Code**	**Odds Ratio† (95%CrI)**	**Odds Ratio†† (95%CrI)**	**Odds Ratio††† (95%CrI)**
Darifenacin ER 30mg q.d	[38]	6.29 (3.49,11.44)	6.06 (3.3,11.39)	6.12 (3.41,11.21)
Imidafenacin 0.25mg b.i.d	[37]	4.62 (2.27,9.36)	4.41 (2.12,9.1)	4.48 (2.18,9)
Solifenacin ER 20mg q.d	[33]	3.28 (1.57,7.43)	3.07 (1.54,7.06)	3.16 (1.56,6.96)
Tarafenacin 0.4mg q.d	[82]	3.43 (1.56,7.7)	3.31 (1.49,7.56)	3.35 (1.51,7.51)
Terodiline 25mg b.i.d	[28]	3.09 (1.15,8.77)	3.09 (1.15,8.75)	3.13 (1.17,8.65)
Propiverine IR 45mg t.i.d	[118]	2.5 (1.63,7.87)	2.53 (1.6,6.37)	2.47 (1.63,6.9)
Propiverine ER 60mg q.d	[119]	2.51 (1.62,7.96)	2.53 (1.6,6.49)	2.46 (1.62,7.11)
Duloxetine 60mg b.i.d	[66]	3.05 (1.6,5.88)	3.05 (1.59,5.9)	3.03 (1.59,5.87)
Oxybutynin IR 5mg t.i.d	[7]	2.56 (1.76,3.66)	2.55 (1.75,3.69)	2.57 (1.75,3.65)
Cizilirtine citrate 400mg b.i.d	[57]	2.44 (1.22,4.99)	2.44 (1.22,4.96)	2.45 (1.21,4.96)
Oxbutynin patch 73.5mg	[15]	2.67 (1.97,3.77)	2.68 (1.95,3.77)	2.67 (1.97,3.75)
Oxybutynin vaginal ring 6mg q.d	[17]	2.48 (1.57,3.53)	2.46 (1.59,3.56)	2.49 (1.59,3.53)
Oxybutynin intravesically 5mg t.i.d	[14]	2.57 (1.45,4.47)	2.55 (1.46,4.41)	2.58 (1.45,4.37)
Oxybutynin IR 5mg b.i.d	[18]	2.97 (2.22,4.46)	3.05 (2.25,4.46)	2.96 (2.22,4.44)
Oxybutynin IR 2.5 - 5mg b.i.d	[24]	2.49 (1.7,3.52)	2.47 (1.7,3.51)	2.5 (1.71,3.49)
Darifenacin ER 15mg q.d.	[40]	2.33 (1.54,3.54)	2.35 (1.56,3.57)	2.34 (1.55,3.55)
Oxybutynin chloride topical gel 1g/day	[13]	2.36 (1.45,3.3)	2.31 (1.47,3.3)	2.37 (1.47,3.28)
Fesoterodine ER 8mg q.d	[26]	2.45 (1.87,3.18)	2.43 (1.87,3.14)	2.44 (1.85,3.17)
Oxybutynin vaginal ring 4mg q.d	[16]	2.38 (1.44,3.35)	2.33 (1.46,3.36)	2.39 (1.45,3.34)
Fesoterodine IR 12mg b.i.d	[122]	1.98 (1.04,3.94)	1.98 (1.12,3.58)	1.98 (1.09,3.7)
Propiverine IR 30mg b.i.d	[117]	2.25 (1.41,5.5)	2.24 (1.38,4.73)	2.23 (1.41,5.05)
Propiverine ER 30mg q.d	[42]	1.96 (1.22,2.84)	1.96 (1.27,2.88)	1.97 (1.25,2.83)
Solifenacin ER 5 - 15mg q.d	[34]	2.19 (1.32,3.71)	2.14 (1.32,3.61)	2.17 (1.33,3.66)
Propiverine IR 15mg t.i.d	[116]	2.18 (1.46,3.56)	2.2 (1.47,3.49)	2.18 (1.47,3.48)
Propiverine ER 20mg q.d	[41]	2.15 (1.7,2.8)	2.17 (1.71,2.81)	2.15 (1.7,2.79)
Solifenacin ER 10mg q.d	[30]	1.61 (1.06,2.46)	1.63 (1.07,2.47)	1.62 (1.06,2.45)
Tolterodine IR 4mg b.i.d	[140]	1.45 (0.97,5.75)	1.46 (0.94,4.53)	1.42 (0.97,5.14)
OnaBoNTA 300u trigone sparing	[76]	1.75 (0.94,3.61)	1.75 (0.91,3.56)	1.73 (0.94,3.61)
Darifenacin ER 7.5mg q.d	[39]	2.09 (1.35,3.23)	2.12 (1.37,3.3)	2.11 (1.36,3.27)
Cizilirtine citrate 200mg b.i.d	[56]	2.33 (0.88,6.25)	2.36 (0.93,5.99)	2.36 (0.92,6.16)
Trospium chloride IR 15mg t.i.d	[46]	1.61 (0.6,2.88)	1.62 (0.72,2.81)	1.64 (0.69,2.83)
Fesoterodine IR 8mg b.id	[121]	1.85 (0.9,3.28)	1.86 (1.02,3.16)	1.86 (0.95,3.22)
Trospium chloride IR 45mg t.i.d	[47]	1.83 (1.22,3.1)	1.83 (1.21,2.99)	1.83 (1.22,3)
Fesoterodine ER 4 - 8mg q.d	[27]	1.7 (1.28,2.21)	1.7 (1.29,2.22)	1.7 (1.28,2.22)
Propiverine IR 15mg b.i.d	[43]	1.91 (1.26,2.65)	1.88 (1.28,2.65)	1.91 (1.28,2.64)
Solifenacin ER 5 - 10mg q.d	[31]	2.18 (1.4,3.43)	2.16 (1.4,3.38)	2.17 (1.4,3.4)
OnaBoNT-A 200U trigone sparing	[73]	1.56 (0.89,2.73)	1.55 (0.88,2.74)	1.57 (0.9,2.78)
Fesoterodine ER 4mg q.d	[25]	1.82 (1.45,2.26)	1.82 (1.45,2.26)	1.82 (1.46,2.26)
Trospium chloride ER 60mg q.d	[44]	1.64 (1.16,2.28)	1.64 (1.17,2.27)	1.65 (1.17,2.27)
Solifenacin ER 5 - 10mg q.d + BT	[77]	2.07 (1.04,4.13)	2.05 (1.03,4.07)	2.05 (1.03,4.06)
ZD0947IL 25mg/day	[58]	1.61 (0.78,3.39)	1.6 (0.76,3.35)	1.61 (0.77,3.4)
Fesoterodine IR 4mg b.i.d	[120]	1.88 (0.93,3.38)	1.89 (1.05,3.23)	1.88 (0.97,3.34)
Imidafenacin 0.1mg b.i.d	[36]	1.63 (1.16,2.3)	1.63 (1.17,2.32)	1.63 (1.17,2.31)
Trospium chloride IR 20mg b.i.d	[45]	1.74 (1.24,2.52)	1.75 (1.24,2.51)	1.74 (1.24,2.5)
OnaBoNTA 150u trigone sparing	[75]	1.75 (0.93,3.63)	1.74 (0.91,3.58)	1.73 (0.94,3.62)
Tarafenacin 0.2mg q.d	[90]	1.47 (0.68,3.13)	1.51 (0.7,3.26)	1.51 (0.71,3.25)
OnaBoNT-A 100u trigone sparing	[72]	1.73 (0.94,3.45)	1.72 (0.92,3.41)	1.71 (0.94,3.45)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	1.3 (0.72,2.38)	1.28 (0.71,2.33)	1.28 (0.72,2.35)
Pregabalin 150mg b.i.d	[62]	1.38 (0.74,2.56)	1.37 (0.73,2.56)	1.37 (0.73,2.54)
Tolterodine IR 2mg b.i.d	[5]	1.32 (1.05,1.71)	1.33 (1.05,1.72)	1.32 (1.06,1.71)
Netupitant 200mg q.d	[112]	1.1 (0.45,2.89)	1.09 (0.44,2.85)	1.08 (0.45,2.8)
OnaBoNTA 50u trigone sparing	[74]	1.56 (0.78,2.99)	1.54 (0.78,2.99)	1.56 (0.8,3.01)
Solifenacin ER 5mg q.d	[29]	1.23 (0.87,1.75)	1.25 (0.88,1.77)	1.24 (0.88,1.76)
Imidafenacin 0.05mg b.i.d	[35]	1.48 (0.77,2.83)	1.49 (0.77,2.84)	1.49 (0.78,2.82)
Tolterodine ER 4mg q.d	[4]	1.24 (1.08,1.42)	1.24 (1.07,1.42)	1.24 (1.08,1.42)

TABLE D.9: Sensitivity analyses assessing the impact of different choices of prior distributions on class-specific variance parameters for hierarchical network meta-analysis evaluating the number of patients experiencing adverse events (cont.)

Mirabegron ER 200mg q.d	[53]	1.1 (0.86,1.53)	1.12 (0.83,1.61)	1.1 (0.86,1.53)
Anticholinergic (not specified) + saline injection	[100]	1.42 (0.56,3.75)	1.41 (0.55,3.75)	1.42 (0.57,3.78)
ONO-8539 300mg b.i.d	[61]	1.11 (0.63,2.05)	1.1 (0.63,1.98)	1.11 (0.64,2.01)
Mirabegron IR 150mg b.i.d	[49]	1.08 (0.8,1.5)	1.09 (0.76,1.58)	1.08 (0.81,1.5)
Tolterodine IR 2mg b.i.d + BT	[93]	1.13 (0.65,1.97)	1.14 (0.66,1.99)	1.14 (0.66,1.96)
Mirabegron ER 100mg q.d	[52]	1.09 (0.9,1.35)	1.11 (0.9,1.39)	1.1 (0.91,1.35)
Tolterodine IR 1mg b.i.d	[6]	1.17 (0.79,1.57)	1.16 (0.81,1.59)	1.18 (0.81,1.57)
Lipo-BoNTA 200U	[138]	1.11 (0,245.3)	1.31 (0.01,1084)	0.95 (0,204.7)
Mirabegron ER 50mg q.d	[51]	1.06 (0.9,1.26)	1.06 (0.89,1.26)	1.06 (0.9,1.26)
Placebo	[1]	NA	NA	NA
Mirabegron IR 100mg b.i.d	[48]	1.06 (0.71,1.38)	1.04 (0.68,1.44)	1.06 (0.72,1.38)
Mirabegron ER 25mg q.d	[50]	1.07 (0.87,1.34)	1.08 (0.84,1.38)	1.07 (0.87,1.33)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	1.05 (0.56,1.91)	1.06 (0.58,1.92)	1.06 (0.57,1.9)
Solabegron IR 125mg b.i.d	[55]	0.87 (0.43,1.77)	0.87 (0.43,1.74)	0.86 (0.43,1.75)
Netupitant 100mg q.d	[111]	0.94 (0.37,2.44)	0.96 (0.38,2.5)	0.95 (0.37,2.38)
ONO-8539 100mg b.i.d	[60]	0.88 (0.48,1.55)	0.9 (0.5,1.56)	0.9 (0.5,1.56)
Tolterodine IR 0.5mg b.i.d	[141]	1.19 (0.49,2.01)	1.17 (0.56,1.92)	1.2 (0.54,1.97)
ONO-8539 30mg b.i.d	[59]	0.96 (0.54,1.7)	0.97 (0.55,1.68)	0.97 (0.55,1.69)
Percutaneous tibial nerve stimulation	[83]	0.72 (0.28,1.84)	0.74 (0.28,1.84)	0.73 (0.29,1.87)
Solifenacin ER 2.5mg q.d	[32]	0.82 (0.3,1.82)	0.91 (0.33,1.9)	0.86 (0.32,1.85)
Solabegron IR 50mg b.i.d	[54]	0.89 (0.44,1.8)	0.89 (0.45,1.77)	0.88 (0.45,1.78)
Netupitant 50mg q.d	[110]	0.89 (0.34,2.28)	0.91 (0.35,2.37)	0.89 (0.35,2.24)
Electrostimulation	[80]	0.34 (0.02,2.01)	0.39 (0.02,2.02)	0.42 (0.03,1.99)
Sham therapy	[3]	0.13 (0.01,0.83)	0.13 (0.01,0.84)	0.14 (0.02,0.82)
Control	[2]	0.09 (0,1.72)	0.09 (0,1.38)	0.11 (0,1.25)
Bladder Training (BT)/Behaviour Therapy	[85]	0.04 (0,0.28)	0.04 (0,0.28)	0.03 (0,0.26)

† Class-specific standard deviation based on an Uniform(0,2) prior distribution on the standard deviation scale

†† Class-specific standard deviation based on a Gamma(0.001,0.001) prior distribution on the precision scale

††† Class-specific standard deviation based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

Appendix E

Multivariate Hierarchical Network Meta-Analysis of Randomised Controlled Trials in Overactive Bladder

E.1 Example WinBUGS code for multivariate hierarchical NMA

```
Model {

# Generating missing SEs
#=======================
        for(i in 1:N1){
        for (m in 1:no){

change_var[i,m] <- ((pow(b_sd_star[i,m],2) + pow(f_sd_star[i,m],2) -
2*rho.star[i,m]*b_sd_star[i,m]*f_sd_star[i,m])*equals(ind_c_miss[i,m],1)) +
(pow(c_sd[i,m],2)*(equals(ind_c_miss[i,m],0)))

change_sd[i,m] <- sqrt(change_var[i,m])
se[i,m] <- change_sd[i,m]/sqrt(numinclanalysis[i,m])

b_sd_star[i,m] <- 0*equals(ind_b_miss[i,m],1) +
(b_sd[i,m])*(equals(ind_b_miss[i,m],0))

b_sd[i,m] ~ dunif(0,15)
f_sd[i,m] ~ dunif(0, 15)
c_sd[i,m] ~ dunif(0,25)

z[i,m]~ dnorm(z.star[i,m],z.prec[m])
rho.star[i,m]<- (exp(2*z[i,m])-1)/(exp(2*z[i,m])+1)
}

f_sd_star[i,1] <- (-0.011+(0.835*b_sd[i,1]))*(equals(ind_f_miss[i,1],1)) +
(f_sd[i,1])*(equals(ind_f_miss[i,1],0))
f_sd_star[i,2] <- (1.44+(0.42*b_sd[i,2]))*(equals(ind_f_miss[i,2],1)) +
(f_sd[i,2])*(equals(ind_f_miss[i,2],0))
f_sd_star[i,3] <- (0.161+(0.832*b_sd[i,3]))*(equals(ind_f_miss[i,3],1)) +
(f_sd[i,3])*(equals(ind_f_miss[i,3],0))

z.star[i,1] ~ dnorm(0.67, 12.76)
z.star[i,2] ~ dnorm(0.74, 10.41)
z.star[i,3] ~ dnorm(0.51, 10.41)
}

z.se[1] <- 1/(sqrt(49-3))
z.prec[1]<- pow(z.se[1],-2)
z.se[2] <- 1/(sqrt(44-3))
z.prec[2] <- pow(z.se[2],-2)
z.se[3] <- 1/(sqrt(21-3))
z.prec[3] <- pow(z.se[2],-2)
```

```
#Likelihood for arm level data
#======================
for(i in 1:N1){
#tmp1[i] <- studyid[i]                    # study id not used in the model
# multivariate likelihood
y[i,1:3] ~ dmnorm(mean.y[study[i],arm[i],1:3],omega[i,,])
omega[i,1:3,1:3] <- inverse(cov.mat[i,,])          # within-study precision matrix

    #elements of within-study covariance matrix
    cov.mat[i,1,1] <-  pow(se[i,1],2)
    cov.mat[i,2,2] <-  pow(se[i,2],2)
    cov.mat[i,3,3] <-  pow(se[i,3],2)
    cov.mat[i,1,2] <-  se[i,1]*se[i,2]*0.4564
    cov.mat[i,1,3] <-  se[i,1]*se[i,3]*0.6178
    cov.mat[i,2,3] <-  se[i,2]*se[i,3]*0.6763
    cov.mat[i,2,1] <-  cov.mat[i,1,2]
    cov.mat[i,3,1] <-  cov.mat[i,1,3]
    cov.mat[i,3,2] <-  cov.mat[i,2,3]
            }

for(j in 1:ns){
  for(k in 1:na2[j]) {
    for(m in 1:no){
mean.y[j,k,m] <- mu[j,m] + delta[j,k,m]      # define study-specific treatment effects
      }
   }
 }

 #Random effects between-study model
 #================================
for(j in 1:ns)  {
    tmp3[j] <- s[j]

for(m in 1:no)      {
      delta[j,1,m] <-0              #delta in control arm set to zero for all outcomes
      w[j,1,m] <-0                  #multi-arm adjustment in control group set to zero
       }

for(k in 2:na2[j])     {
#trial specific treatment effects drawn from multivariate normal distribution
       delta[j,k,1:no] ~ dmnorm(md[j,k,1:no],precBK[j,k,1:no,1:no])
 for(m in 1:no){
md[j,k,m] <-  (d[m,t[j,k]] - d[m,t[j,1]])+ sw[j,k,m]        #consistency equations
w[j,k,m] <- delta[j,k,m] - (d[m,t[j,k]] - d[m,t[j,1]])      #multi-arm adjustment
sw[j,k,m] <- sum(w[j,1:k-1,m])/(k-1)
```

```
    for(mm in 1:no)   {
                        precBK[j,k,m,mm] <- prec[m,mm]*2*(k-1)/k
                    }
}}}

# Reference treatment effect set to zero
 d[1,1] <- 0
 d[2,1] <- 0
 d[3,1] <- 0

 # Between-study covariance matrix
prec[1:no,1:no] <- inverse(sigma[,])
 sd.se~ dunif(0, 2)
 #prec.se.star~dgamma(0.01,0.01)
 #sd.se<-1/sqrt(prec.se.star)
 #sd.se~dnorm(0,1)I(0,)

  for(m in 1:no) {
    prec.se[m] <- pow(sd.se,-2)
    sigma[m,m] <- pow(sd[m],2)
    sd[m] ~ dunif(0, 2)
  for(j in 1:ns){
      mu[j, m] ~ dnorm(0,0.001)
        }}

#spherical parameterization
      pi <- 3.1415
   for(i in 1:2) {
       for(j in (i+1):no) {
          sigma[i,j] <- rho[i,j]*sd[i]*sd[j]
          sigma[j,i] <- sigma[i,j]
          g[j,i] <- 0
          a[i,j] ~ dunif(0, pi)
          rho[i,j] <- inprod(g[,i], g[,j])
    }}

 g[1,1] <- 1
 g[1,2] <- cos(a[1,2])
 g[2,2] <- sin(a[1,2])
 g[1,3] <- cos(a[1,3])
 g[2,3] <- sin(a[1,3])*cos(a[2,3])
 g[3,3] <- sin(a[1,3])*sin(a[2,3])

# Borrowing information across outcomes
#=========================================
for(k in 2: nt){
```

```
for(m in 1:no) {
    meanD[m,k-1] <- alpha[k-1] + gamma[m]          #outcome and treatment effects
    d[m,k] ~ dnorm(meanD[m,k-1], prec.btw)}}                #treatment effects

for(m in 1:no) {gamma[m] ~ dnorm(0, 0.01) }
#for(k in 1:(nt-1)) {alpha[k] ~ dnorm(0, 0.001) }
for(k in 1:2){alpha[k]~ dnorm(D.d[1], D.d.prec[1])}
for(k in 3:5){alpha[k]~ dnorm(D.d[2], D.d.prec[2])}
for(k in 6:23){alpha[k]~ dnorm(D.d[3], D.d.prec[3])}
for(k in 24:26){alpha[k]~ dnorm(D.d[4], D.d.prec[4])}
alpha[27]~ dnorm(D.d[5], D.d.prec[5])
for(k in 28:30){alpha[k]~ dnorm(D.d[6], D.d.prec[6])}
for(k in 31:32){alpha[k]~ dnorm(D.d[7], D.d.prec[7])}
alpha[33]~ dnorm(D.d[6], D.d.prec[6])
for(k in 34:36){alpha[k]~ dnorm(D.d[8], D.d.prec[8])}
for(k in 37:39){alpha[k]~ dnorm(D.d[9], D.d.prec[9])}
for(k in 40:42){alpha[k]~ dnorm(D.d[10], D.d.prec[10])}
alpha[43]~ dnorm(D.d[11], D.d.prec[11])
alpha[44]~ dnorm(D.d[12], D.d.prec[12])
for(k in 45:46){alpha[k]~ dnorm(D.d[11], D.d.prec[11])}
for(k in 47:52){alpha[k]~ dnorm(D.d[13], D.d.prec[13])}
for(k in 53:54){alpha[k]~ dnorm(D.d[14], D.d.prec[14])}
for(k in 55:56){alpha[k]~ dnorm(D.d[15], D.d.prec[15])}
alpha[57]~ dnorm(D.d[16], D.d.prec[16])
for(k in 58:60){alpha[k]~ dnorm(D.d[17], D.d.prec[17])}
alpha[61]~ dnorm(D.d[18], D.d.prec[18])
alpha[62]~ dnorm(D.d[19], D.d.prec[19])
alpha[63]~ dnorm(D.d[20], D.d.prec[20])
alpha[64]~ dnorm(D.d[21], D.d.prec[21])
alpha[65]~ dnorm(D.d[3], D.d.prec[3])
alpha[66]~ dnorm(D.d[22], D.d.prec[22])
alpha[67]~ dnorm(D.d[23], D.d.prec[23])
for(k in 68:69){alpha[k]~ dnorm(D.d[24], D.d.prec[24])}
alpha[70]~ dnorm(D.d[25], D.d.prec[25])
for(k in 71:72){alpha[k]~ dnorm(D.d[26], D.d.prec[26])}
alpha[73]~ dnorm(D.d[3], D.d.prec[3])
alpha[74]~ dnorm(D.d[27], D.d.prec[27])
alpha[75]~ dnorm(D.d[28], D.d.prec[28])
alpha[76]~ dnorm(D.d[29], D.d.prec[29])
for(k in 77:78){alpha[k]~ dnorm(D.d[26], D.d.prec[26])}
alpha[79]~ dnorm(D.d[30], D.d.prec[30])
alpha[80]~ dnorm(D.d[31], D.d.prec[31])
alpha[81]~ dnorm(D.d[23], D.d.prec[23])
alpha[82]~ dnorm(D.d[32], D.d.prec[32])
alpha[83]~ dnorm(D.d[33], D.d.prec[33])
alpha[84]~ dnorm(D.d[34], D.d.prec[34])
```

```
alpha[85]~ dnorm(D.d[35], D.d.prec[35])
alpha[86]~ dnorm(D.d[36], D.d.prec[36])
alpha[87]~ dnorm(D.d[23], D.d.prec[23])
alpha[88]~ dnorm(D.d[37], D.d.prec[37])
alpha[89]~ dnorm(D.d[38], D.d.prec[38])
alpha[90]~ dnorm(D.d[39], D.d.prec[39])
alpha[91]~ dnorm(D.d[35], D.d.prec[35])
alpha[92]~ dnorm(D.d[36], D.d.prec[36])
alpha[93]~ dnorm(D.d[10], D.d.prec[10])
alpha[94]~ dnorm(D.d[40], D.d.prec[40])
alpha[95]~ dnorm(D.d[41], D.d.prec[41])
alpha[96]~ dnorm(D.d[42], D.d.prec[42])
alpha[97]~ dnorm(D.d[43], D.d.prec[43])
alpha[98]~ dnorm(D.d[44], D.d.prec[44])
alpha[99]~ dnorm(D.d[45], D.d.prec[45])
alpha[100]~ dnorm(D.d[46], D.d.prec[46])
for(k in 101:102){alpha[k]~ dnorm(D.d[47], D.d.prec[47])}
for(k in 103:104){alpha[k]~ dnorm(D.d[10], D.d.prec[10])}
alpha[105]~ dnorm(D.d[3], D.d.prec[3])
for(k in 106:108){alpha[k]~ dnorm(D.d[48], D.d.prec[48])}
for(k in 109:111){alpha[k]~ dnorm(D.d[49], D.d.prec[49])}
alpha[112]~ dnorm(D.d[50], D.d.prec[50])
alpha[113]~ dnorm(D.d[51], D.d.prec[51])
alpha[114]~ dnorm(D.d[52], D.d.prec[52])

for(i in 1:52){
D.d[i] ~ dnorm(0, 0.001)
D.d.prec[i]<- pow(D.d.sd[i],-2)
D.d.sd[i] ~ dunif(0,2)}

 prec.btw <- pow(sd.btw,-2)
 sd.btw ~ dunif(0, 2)
#prec.btw <- pow(sd.btw,-2)
#sd.btw ~ dnorm(0,1)I(0,)
#prec.btw ~ dgamma(0.001,0.001)
#sd.btw <-1/sqrt(prec.btw)

for (m in 1:no) {
for (c in 1:(nt-1)) {
        for (k in (c+1):nt) {
                diff[m,c,k] <- (d[m,k] - d[m,c] )}}
for (k in 1:nt) {
        rk[m,k] <- rank(d[m,],k)

        best[m,k] <- equals(rk[m,k],1)}
        }}}
```

E.2 Multivariate NMA results

TABLE E.1: Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate network meta-analysis

Treatment	**Code**	**Incontinence episodes†**	**Voiding episodes†**	**Urgency episodes†**
Sacral nerve stimulation	[81]	-7.43 (-9.59,-4.73)	-7.64 (-9.82,-4.91)	-7.96 (-10.19,-5.29)
OnaBoNT-A 200u trigone sparing	[73]	-2.38 (-3.13,-1.66)	-2.61 (-3.38,-1.9)	-2.9 (-3.76,-2.14)
Estradiol 3mg intravaginally	[128]	-1.91 (-3.6,-0.26)	-2.13 (-3.82,-0.51)	-2.44 (-4.18,-0.72)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-1.84 (-3.51,-0.17)	-2.06 (-3.73,-0.35)	-2.39 (-4.09,-0.7)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.73 (-3.26,-0.12)	-1.95 (-3.48,-0.35)	-2.25 (-3.84,-0.62)
Electrostimulation + PFE + Bladder training	[97]	-1.6 (-2.52,-0.7)	-1.8 (-2.77,-0.91)	-2.13 (-3.05,-1.16)
Solifenacin/trospium + placebo injection	[100]	-1.64 (-2.56,-0.59)	-1.85 (-2.82,-0.8)	-2.14 (-3.17,-1.12)
OnaBoNTA 100u trigone sparing	[72]	-1.58 (-1.96,-1.21)	-1.69 (-2.04,-1.36)	-2.06 (-2.53,-1.64)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.48 (-2.43,-0.48)	-1.7 (-2.67,-0.69)	-2.01 (-3.02,-1.05)
Oxybutynin intravesically 5mg t.i.d	[14]	-1.31 (-2.4,-0.07)	-1.53 (-2.63,-0.28)	-1.84 (-2.95,-0.54)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.35 (-1.76,-0.96)	-1.6 (-2.05,-1.15)	-1.9 (-2.43,-1.38)
Propiverine 30mg b.i.d	[42]	-1.3 (-3.48,0.7)	-1.53 (-3.69,0.47)	-1.87 (-4.06,0.26)
Estriol 1mg intravesival	[131]	-1.31 (-2.53,0.05)	-1.54 (-2.72,-0.19)	-1.85 (-3.11,-0.45)
Oxybutynin ER 10mg q.d	[8]	-1.06 (-1.53,-0.56)	-1.29 (-1.8,-0.78)	-1.59 (-2.18,-0.99)
Mirabegron 100mg b.i.d	[48]	-1.05 (-1.69,-0.43)	-1.26 (-1.9,-0.64)	-1.58 (-2.29,-0.83)
Solifenacin ER 10mg q.d	[30]	-0.87 (-1.09,-0.65)	-1.13 (-1.35,-0.92)	-1.41 (-1.78,-1.04)
Imidafenacin IR 0.25mg b.i.d	[37]	-0.88 (-1.49,-0.32)	-1.11 (-1.72,-0.53)	-1.42 (-2.06,-0.75)
Mirabegron 150mg b.i.d	[49]	-0.83 (-1.69,-0.08)	-1.07 (-1.92,-0.32)	-1.37 (-2.29,-0.58)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.83 (-1.57,-0.2)	-1.06 (-1.78,-0.45)	-1.36 (-2.15,-0.67)
Oxybutynin IR 3mg t.i.d	[19]	-0.83 (-1.17,-0.47)	-1.02 (-1.4,-0.64)	-1.35 (-1.81,-0.88)
Tolterodine ER 4mg q.d + Behaviour therapy	[87]	-0.78 (-1.47,-0.15)	-1 (-1.71,-0.35)	-1.32 (-2.03,-0.62)
Propiverine ER 30mg q.d	[42]	-0.76 (-1.27,-0.26)	-0.99 (-1.54,-0.43)	-1.3 (-1.92,-0.7)
Darifenacin ER 30mg q.d	[38]	-0.73 (-1.35,-0.13)	-0.95 (-1.59,-0.33)	-1.26 (-1.9,-0.57)
Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	-0.72 (-1.22,-0.22)	-0.94 (-1.5,-0.38)	-1.26 (-1.88,-0.66)
Fesoterodine ER 8mg q.d	[26]	-0.71 (-0.9,-0.53)	-0.94 (-1.14,-0.76)	-1.26 (-1.62,-0.9)
Trospium IR 15mg t.i.d + Physiotherapy	[91]	-0.7 (-1.54,0.14)	-0.95 (-1.8,-0.1)	-1.24 (-2.11,-0.39)
Solifenacin ER (5mg-10mg) q.d	[31]	-0.67 (-0.91,-0.43)	-0.85 (-1.09,-0.61)	-1.18 (-1.56,-0.78)
Solabegron 125mg b.i.d	[55]	-0.64 (-0.92,-0.38)	-0.87 (-1.16,-0.61)	-1.18 (-1.58,-0.78)
Mirabegron 25mg q.d	[50]	-0.63 (-0.9,-0.36)	-0.85 (-1.12,-0.57)	-1.14 (-1.54,-0.74)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.61 (-1.16,-0.03)	-0.86 (-1.41,-0.27)	-1.15 (-1.78,-0.53)
Solifenacin ER 5mg - 15mg q.d	[34]	-0.61 (-1.11,-0.12)	-0.82 (-1.38,-0.29)	-1.12 (-1.75,-0.54)
Trospium ER 60mg q.d	[44]	-0.6 (-0.94,-0.24)	-0.81 (-1.15,-0.45)	-1.12 (-1.58,-0.66)
Cizolirtine Citrate 400mg b.i.d	[57]	-0.58 (-1.21,-0.04)	-0.83 (-1.47,-0.24)	-1.14 (-1.81,-0.47)
Mirabegron 100mg q.d	[52]	-0.59 (-0.79,-0.4)	-0.78 (-0.97,-0.57)	-1.1 (-1.45,-0.75)
Solifenacin ER 5mg q.d	[29]	-0.57 (-0.79,-0.38)	-0.73 (-0.93,-0.56)	-1.09 (-1.46,-0.73)
Mirabegron 50mg q.d	[51]	-0.56 (-0.73,-0.4)	-0.82 (-0.99,-0.65)	-1.09 (-1.43,-0.74)
Oxybutynin IR (2.5-5mg) b.i.d	[24]	-0.53 (-1.14,0.12)	-0.74 (-1.34,-0.09)	-1.05 (-1.72,-0.35)
Oxybutynin IR 5mg t.i.d	[7]	-0.53 (-0.91,-0.14)	-0.73 (-1.12,-0.34)	-1.07 (-1.55,-0.53)
Mirabegron 200mg q.d	[53]	-0.52 (-1.11,-0.01)	-0.74 (-1.38,-0.19)	-1.06 (-1.72,-0.43)
Propiverine IR 15mg b.i.d	[43]	-0.53 (-0.94,-0.1)	-0.71 (-1.15,-0.25)	-1.04 (-1.56,-0.5)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.51 (-1.12,0.12)	-0.73 (-1.32,-0.1)	-1.04 (-1.66,-0.39)
Pregabalin 150mg b.i.d	[62]	-0.49 (-1.02,0.06)	-0.74 (-1.24,-0.25)	-1.03 (-1.61,-0.45)
Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	-0.5 (-0.78,-0.21)	-0.73 (-1.06,-0.41)	-1.03 (-1.47,-0.6)
Fesoterodine ER 4mg q.d	[25]	-0.49 (-0.67,-0.33)	-0.7 (-0.87,-0.55)	-1.02 (-1.37,-0.68)
Oxybutynin chloride topical gel 1g q.d	[13]	-0.5 (-0.93,-0.06)	-0.71 (-1.15,-0.28)	-1.03 (-1.54,-0.5)
Tolterodine ER 4mg q.d	[4]	-0.49 (-0.6,-0.39)	-0.64 (-0.75,-0.52)	-0.98 (-1.28,-0.68)
Darifenacin ER 15mg q.d	[40]	-0.47 (-0.94,0.03)	-0.68 (-1.18,-0.11)	-1 (-1.55,-0.37)
Electromagnetic stimulation	[125]	-0.45 (-2.54,1.46)	-0.67 (-2.73,1.22)	-0.98 (-3.1,0.98)
Oxybutynin gel 84mg/day	[134]	-0.45 (-1.03,0.2)	-0.71 (-1.28,-0.07)	-1.01 (-1.64,-0.25)
Tolterodine IR 2mg b.i.d	[5]	-0.45 (-0.58,-0.31)	-0.69 (-0.84,-0.53)	-0.97 (-1.3,-0.64)
Tolterodine IR 2mg b.i.d + PFE	[95]	-0.44 (-1.06,0.22)	-0.65 (-1.29,0.06)	-0.96 (-1.62,-0.23)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.41 (-0.96,0.11)	-0.67 (-1.23,-0.13)	-0.96 (-1.57,-0.37)
Propiverine ER 20mg q.d	[41]	-0.41 (-0.59,-0.22)	-0.65 (-0.84,-0.45)	-0.95 (-1.3,-0.59)
Trospium chloride 45mg t.i.d	[47]	-0.41 (-1.12,0.4)	-0.63 (-1.31,0.13)	-0.94 (-1.69,-0.08)
Elocalcitol 75mg	[70]	-0.39 (-0.91,0.21)	-0.56 (-1.08,0.07)	-0.9 (-1.53,-0.25)
Propiverine 45mg t.i.d	[118]	-0.38 (-2.55,1.76)	-0.61 (-2.75,1.58)	-0.93 (-3.13,1.32)
Fesoterodine ER (4mg-8mg) q.d	[27]	-0.36 (-0.57,-0.13)	-0.64 (-0.84,-0.44)	-0.91 (-1.27,-0.55)

TABLE E.1: Estimated posterior median difference (and 95% credible intervals) in change from baseline for urinary incontinence, voiding and urgency episodes obtained from multivariate network meta-analysis (cont.)

Tolterodine IR 1mg b.i.d	[6]	-0.34 (-0.67,-0.01)	-0.57 (-0.92,-0.21)	-0.88 (-1.35,-0.42)
Imidafenacin IR 0.1mg b.i.d	[36]	-0.35 (-0.62,-0.1)	-0.53 (-0.8,-0.28)	-0.86 (-1.26,-0.44)
Terodiline IR 25mg b.i.d	[28]	-0.35 (-0.81,0.09)	-0.51 (-0.96,-0.09)	-0.86 (-1.37,-0.28)
Oxybutynin transdermal 3.9mg/day	[10]	-0.32 (-0.64,-0.03)	-0.55 (-0.87,-0.24)	-0.84 (-1.28,-0.44)
Oxybutynin gel 56mg/day	[135]	-0.32 (-0.95,0.33)	-0.48 (-1.1,0.15)	-0.81 (-1.53,-0.08)
Oxbutynin patch 73.5mg	[15]	-0.31 (-0.65,0.05)	-0.53 (-0.93,-0.11)	-0.84 (-1.31,-0.34)
Imidafenacin IR 0.05mg b.i.d	[35]	-0.29 (-0.73,0.17)	-0.53 (-0.99,-0.05)	-0.82 (-1.36,-0.26)
Darifenacin ER 7.5mg q.d	[39]	-0.28 (-0.85,0.26)	-0.5 (-1.09,0.06)	-0.83 (-1.46,-0.2)
Elocalcitol 150mg	[69]	-0.28 (-0.82,0.33)	-0.5 (-1.05,0.12)	-0.8 (-1.45,-0.13)
Duloxetine IR 40mg b.i.d	[65]	-0.27 (-0.77,0.24)	-0.5 (-1.06,0.06)	-0.8 (-1.43,-0.2)
Solabegron 50mg b.i.d	[54]	-0.24 (-0.52,0.03)	-0.47 (-0.75,-0.21)	-0.77 (-1.18,-0.38)
Lipo-BoNTA	[138]	-0.21 (-1.06,0.64)	-0.46 (-1.35,0.42)	-0.77 (-1.68,0.19)
Tarafenacin 0.4mg q.d	[82]	-0.2 (-0.87,0.51)	-0.45 (-1.1,0.27)	-0.74 (-1.45,0.07)
Serlopitant 0.25mg q.d	[107]	-0.21 (-0.73,0.37)	-0.42 (-0.91,0.09)	-0.72 (-1.36,-0.12)
Serlopitant 4mg q.d	[109]	-0.19 (-0.76,0.36)	-0.41 (-0.94,0.1)	-0.71 (-1.34,-0.11)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.18 (-0.66,0.3)	-0.42 (-0.9,0.04)	-0.71 (-1.29,-0.17)
Oxybutynin 20mg intravesically q.d	[106]	-0.15 (-1.63,1.3)	-0.36 (-1.83,1.05)	-0.65 (-2.14,0.75)
PFMT + BT	[89]	-0.1 (-0.66,0.43)	-0.37 (-0.97,0.2)	-0.66 (-1.28,-0.04)
Cizolirtine citrate 200mg b.i.d	[56]	-0.12 (-1.5,1.07)	-0.33 (-1.69,0.84)	-0.63 (-2.03,0.52)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.1 (-0.41,0.21)	-0.42 (-0.9,-0.07)	-0.63 (-1.08,-0.17)
Percutaneous tibial nerve stimulation	[83]	-0.09 (-1.07,1.21)	-0.35 (-1.32,0.98)	-0.63 (-1.61,0.71)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.07 (-0.75,0.63)	-0.23 (-0.9,0.48)	-0.63 (-1.35,0.13)
Electrostimulation	[80]	-0.03 (-0.48,0.43)	-0.31 (-0.8,0.13)	-0.67 (-1.26,-0.14)
ONO-8539 100mg b.i.d	[60]	-0.02 (-0.66,0.61)	-0.22 (-0.87,0.41)	-0.53 (-1.23,0.16)
Oxybutynin ER 15mg q.d	[9]	-0.01 (-1.17,0.83)	-0.22 (-1.41,0.63)	-0.54 (-1.74,0.35)
Netupitant 200mg q.d	[112]	-0.02 (-1.19,1.09)	-0.23 (-1.42,0.82)	-0.52 (-1.75,0.53)
Estradiol 25mg	[68]	0.01 (-0.37,0.37)	-0.21 (-0.67,0.2)	-0.53 (-1.03,-0.02)
Placebo	[1]	NA	NA	NA
Netupitant 100mg q.d	[111]	0.01 (-1.1,0.96)	-0.2 (-1.29,0.71)	-0.52 (-1.62,0.4)
PFMT	[84]	0.07 (-0.63,0.64)	-0.18 (-0.85,0.43)	-0.48 (-1.12,0.18)
Oxybutynin transdermal 1.3mg/day	[11]	0.07 (-0.46,0.6)	-0.14 (-0.67,0.38)	-0.45 (-1.06,0.11)
ZD0947IL 25mg/day	[58]	0.07 (-0.81,0.98)	-0.12 (-1.03,0.82)	-0.45 (-1.39,0.52)
Serlopitant 1mg q.d	[108]	0.08 (-0.45,0.66)	-0.14 (-0.64,0.38)	-0.44 (-1.05,0.17)
Netupitant 50mg q.d	[110]	0.13 (-1.01,1.11)	-0.1 (-1.19,0.87)	-0.4 (-1.54,0.57)
Bladder Training	[85]	0.15 (-0.37,0.66)	-0.07 (-0.58,0.46)	-0.4 (-0.91,0.19)
Tarafenacin 0.2mg q.d	[90]	0.15 (-0.55,0.83)	-0.06 (-0.77,0.64)	-0.38 (-1.15,0.38)
Oxybutynin ER 2.5mg q.d + Bladder training	[92]	0.22 (-0.83,1.31)	0 (-1.05,1.11)	-0.32 (-1.34,0.84)
Oxybutynin transdermal 2.6mg/day	[12]	0.22 (-0.3,0.74)	-0.03 (-0.55,0.5)	-0.32 (-0.93,0.29)
Oxybutynin ER (5-30mg) q.d	[22]	0.23 (-0.34,0.85)	0.03 (-0.56,0.67)	-0.29 (-0.89,0.38)
Resiniferatoxin 50nM	[67]	0.27 (-0.9,1.61)	0.04 (-1.13,1.36)	-0.26 (-1.48,1.07)
Vaginal oestrogen cream 1.25mg/day	[132]	0.29 (-0.24,0.81)	0.03 (-0.5,0.55)	-0.22 (-0.81,0.41)
Flavoxate chloride 200mg q.d	[64]	0.34 (-0.36,1.05)	0.12 (-0.62,0.88)	-0.2 (-0.95,0.61)
Oxybutynin IR 5mg b.i.d	[18]	0.35 (-0.25,0.95)	0.13 (-0.5,0.75)	-0.17 (-0.88,0.49)
Emepronium bromide 200mg q.d	[63]	0.37 (-0.28,1.05)	0.15 (-0.55,0.87)	-0.16 (-0.9,0.6)
ONO-8539 300mg b.i.d	[61]	0.45 (-0.21,1.14)	0.22 (-0.44,0.91)	-0.1 (-0.78,0.65)
Propantheline Bromide 15mg t.i.d	[113]	0.45 (-0.67,2.07)	0.22 (-0.88,1.89)	-0.09 (-1.26,1.73)
Oxybutynin ER 2.5mg q.d	[20]	0.53 (-0.33,1.49)	0.29 (-0.61,1.25)	-0.01 (-0.9,0.94)
Estradiol 1mg intravaginally	[127]	0.58 (-0.74,2.02)	0.36 (-0.95,1.8)	0.04 (-1.33,1.52)
Propiverine 60mg q.d	[119]	0.59 (-1.97,2.89)	0.36 (-2.19,2.69)	0.05 (-2.45,2.45)
ONO-8539 30mg b.i.d	[59]	0.56 (-0.04,1.2)	0.34 (-0.27,0.96)	0.04 (-0.63,0.75)
Sham Therapy	[3]	0.54 (-0.55,2)	0.35 (-0.75,1.82)	0.03 (-1.02,1.54)
Trospium IR 15mg t.i.d	[46]	0.6 (-0.32,1.46)	0.43 (-0.49,1.36)	0.08 (-0.82,1.04)
Oxybutynin IR (5-20mg)	[23]	0.68 (-0.66,1.94)	0.46 (-0.91,1.73)	0.16 (-1.24,1.46)
Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	0.96 (-0.08,2.17)	0.72 (-0.3,1.94)	0.44 (-0.62,1.63)
Control	[2]	0.99 (0.22,2.01)	0.82 (0.04,1.87)	0.47 (-0.31,1.52)
Reflexology	[71]	1.03 (0.25,2.15)	0.78 (-0.05,1.88)	0.48 (-0.36,1.61)
Naftopidil 25mg q.d	[114]	3.46 (1.53,5.44)	3.24 (1.29,5.2)	2.94 (0.95,4.99)
Solifenacin succinate 5mg q.d + Naftopidil 25mg q.d	[115]	5.26 (2.92,7.79)	5.04 (2.71,7.54)	4.72 (2.43,7.26)

† median relative to a placebo intervention

E.3 Convergence diagnostics

FIGURE E.1: Brooks-Gelman-Rubin plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes

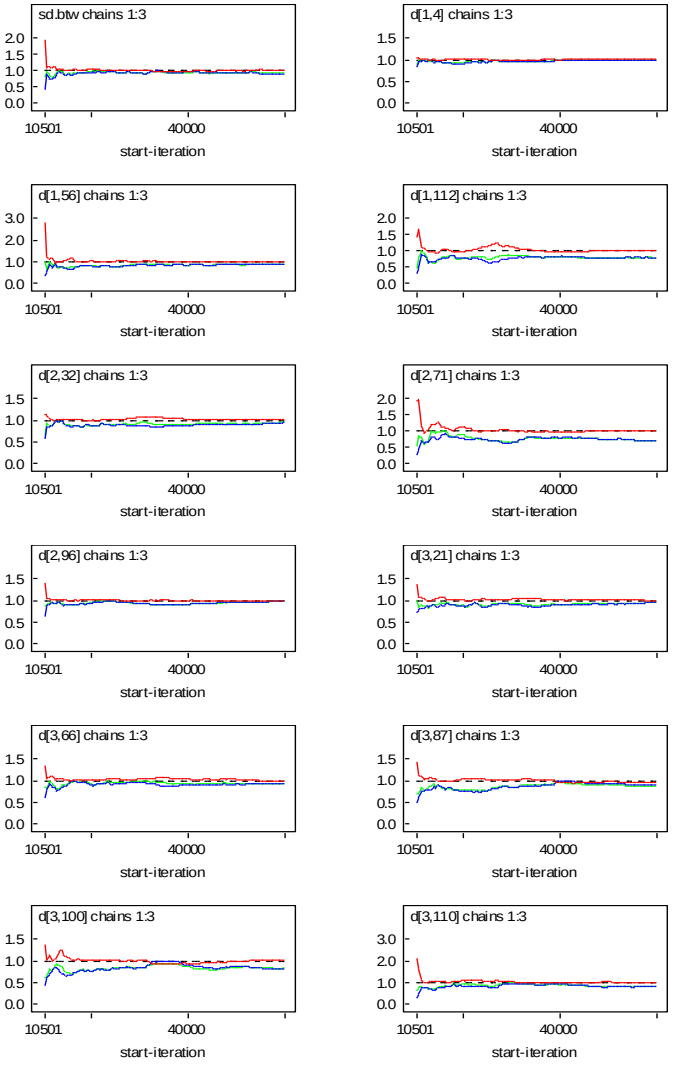


FIGURE E.2: Autocorrelation plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes

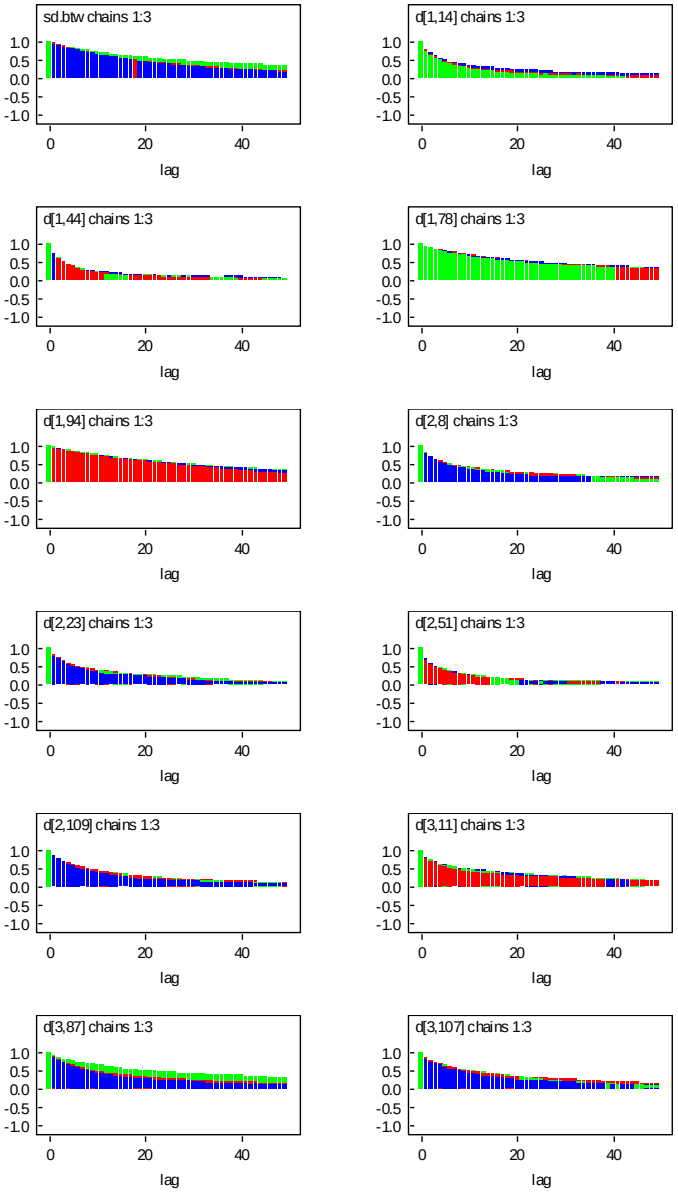
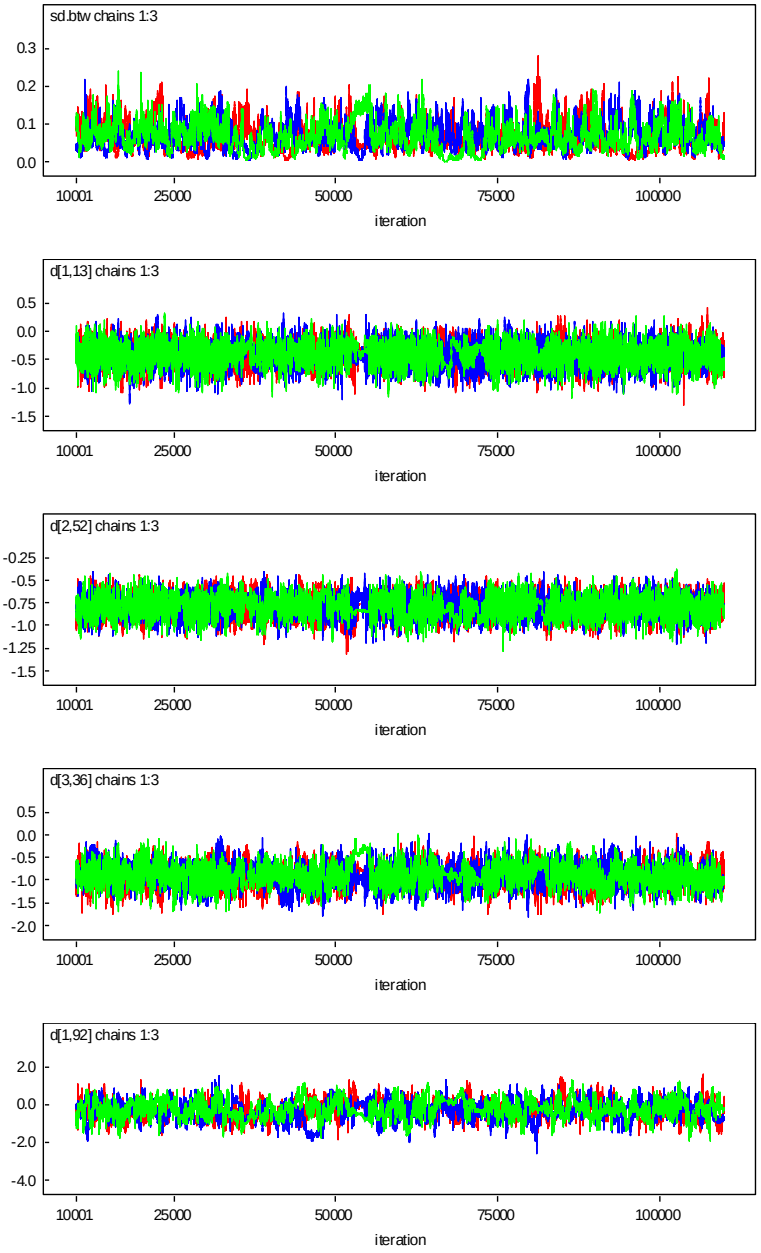
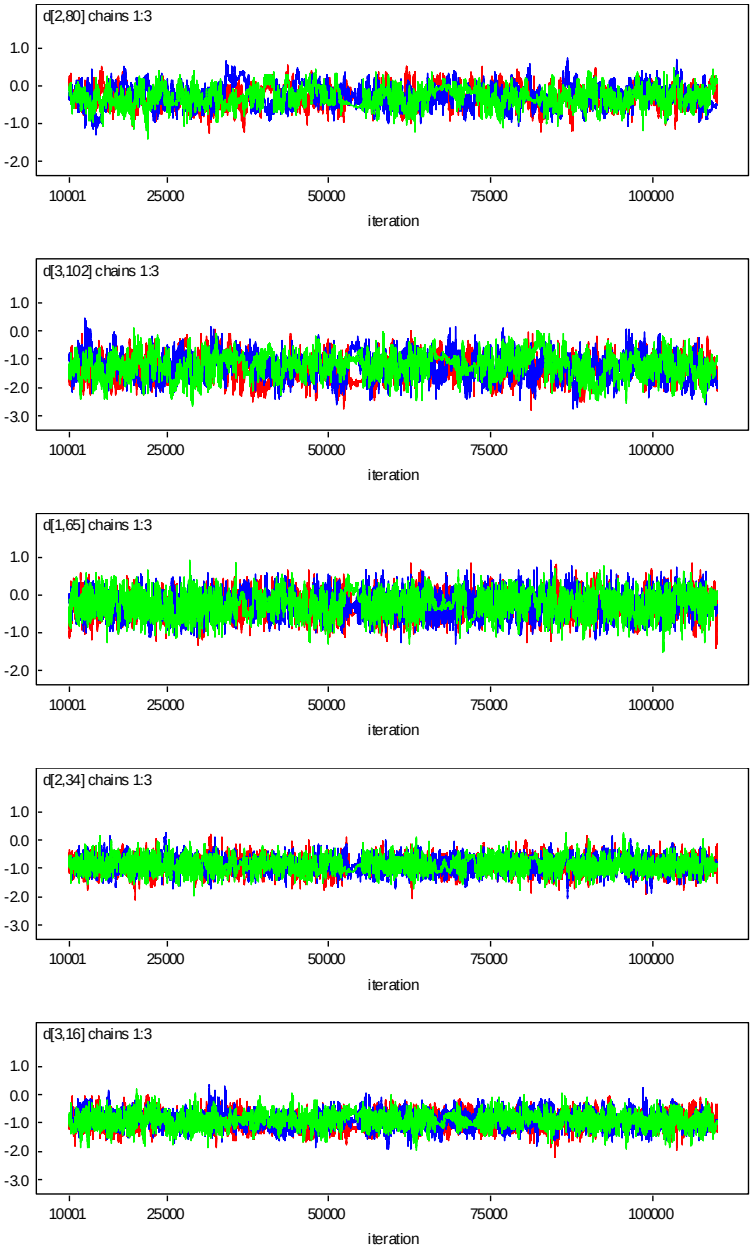


FIGURE E.3: History and trace plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes




d[2,80] chains 1:3
1.0
0.0
-1.0
-2.0
10001
25000
50000
75000
100000
iteration
d[3,102] chains 1:3
1.0
0.0
-1.0
-2.0
-3.0
10001
25000
50000
75000
100000
iteration
d[1,65] chains 1:3
1.0
0.0
-1.0
-2.0
10001
25000
50000
75000
100000
iteration
d[2,34] chains 1:3
1.0
0.0
-1.0
-2.0
-3.0
10001
25000
50000
75000
100000
iteration
d[3,16] chains 1:3
1.0
0.0
-1.0
-2.0
-3.0
10001
25000
50000
75000
100000
iteration

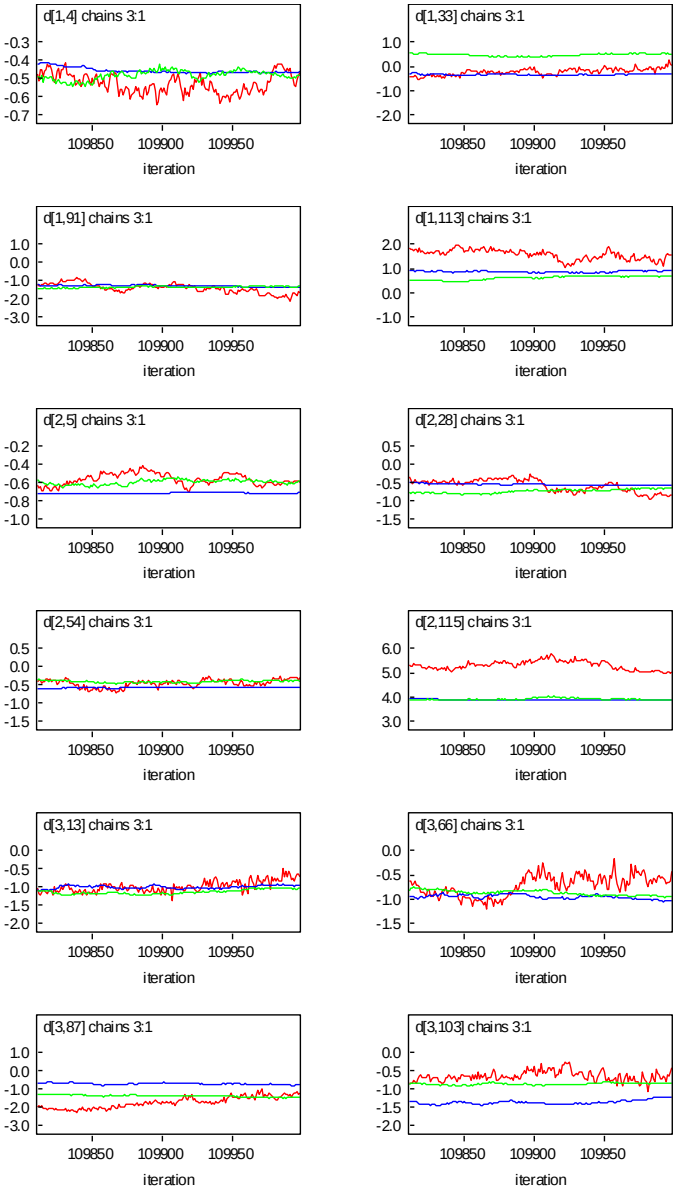
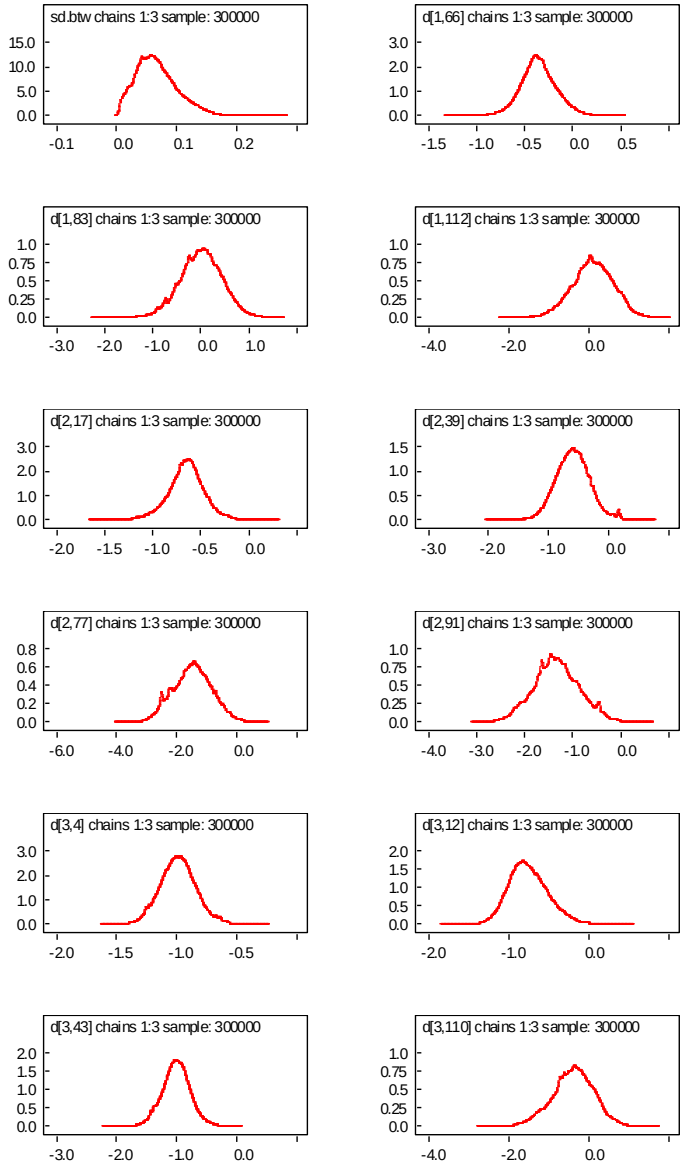

d[1,4] chains 3:1
-0.3
-0.4
-0.5
-0.6
-0.7
d[1,33] chains 3:1
1.0
0.0
-1.0
-2.0
d[1,91] chains 3:1
1.0
0.0
-1.0
-2.0
-3.0
d[1,113] chains 3:1
2.0
1.0
0.0
-1.0
d[2,5] chains 3:1
-0.2
-0.4
-0.6
-0.8
-1.0
d[2,28] chains 3:1
0.5
0.0
-0.5
-1.0
-1.5
d[2,54] chains 3:1
0.5
0.0
-0.5
-1.0
-1.5
d[2,115] chains 3:1
6.0
5.0
4.0
3.0
d[3,13] chains 3:1
0.0
-0.5
-1.0
-1.5
-2.0
d[3,66] chains 3:1
0.0
-0.5
-1.0
-1.5
d[3,87] chains 3:1
1.0
0.0
-1.0
-2.0
-3.0
d[3,103] chains 3:1
0.0
-0.5
-1.0
-1.5
-2.0
109850
109900
109950
iteration

FIGURE E.4: Density plots obtained from multivariate hierarchical network meta-analysis evaluating the mean change from baseline in urinary incontinence, voiding and urgency episodes



E.4 Assessing inconsistencies between direct and indirect information

TABLE E.2: Sensitivity analysis assessing the impact of studies A157 and A158 on the estimated posterior median differences (and 95% credible intervals) obtained from multivariate hierarchical network meta-analyses

		Incontinence episodes	Voiding episodes	Urgency episodes
Treatment	**Code**	**Median difference† (95%CrI)**	**Median difference† (95%CrI)**	**Median difference† (95%CrI)**
Sacral nerve stimulation	[81]	-8.28 (-10.23,-5.65)	-8.45 (-10.4,-5.82)	-8.69 (-10.6,-6.04)
OnaBoNT-A 200u trigone sparing	[73]	-1.95 (-2.88,-1.38)	-2.12 (-3.07,-1.56)	-2.35 (-3.29,-1.73)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-1.96 (-3.79,-0.06)	-2.15 (-3.97,-0.22)	-2.38 (-4.2,-0.45)
Electrostimulation + PFMT + BT	[97]	-1.9 (-2.68,-1.01)	-2.07 (-2.86,-1.18)	-2.31 (-3.12,-1.41)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.82 (-2.84,-0.78)	-2 (-3.02,-0.98)	-2.25 (-3.21,-1.16)
Solifenacin/trospium + placebo injection	[100]	-1.84 (-2.83,-0.86)	-2.01 (-3.05,-1.03)	-2.24 (-3.29,-1.23)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.64 (-2.3,-0.92)	-1.82 (-2.49,-1.09)	-2.07 (-2.73,-1.28)
OnaBoNTA 100u trigone sparing	[72]	-1.64 (-2,-1.3)	-1.76 (-2.1,-1.43)	-2.03 (-2.41,-1.63)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.36 (-1.76,-0.94)	-1.55 (-1.98,-1.13)	-1.78 (-2.24,-1.34)
Estriol 1mg intravesival	[131]	-1.27 (-2.68,0)	-1.44 (-2.87,-0.18)	-1.68 (-3.05,-0.38)
Estradiol 3mg intravaginally	[128]	-1.23 (-2.92,0.35)	-1.41 (-3.08,0.18)	-1.64 (-3.25,-0.04)
Trospium IR 15mg t.i.d + Physiotherapy	[91]	-1.12 (-1.86,-0.41)	-1.31 (-2.08,-0.61)	-1.55 (-2.32,-0.8)
Solifenacin ER 10mg q.d	[30]	-0.86 (-1.08,-0.64)	-1.07 (-1.28,-0.87)	-1.28 (-1.58,-1)
Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	-0.81 (-1.26,-0.3)	-0.99 (-1.46,-0.46)	-1.23 (-1.74,-0.67)
Imidafenacin IR 0.25mg b.i.d	[37]	-0.74 (-1.26,-0.27)	-0.93 (-1.45,-0.45)	-1.16 (-1.73,-0.64)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.74 (-1.38,-0.19)	-0.93 (-1.55,-0.4)	-1.15 (-1.8,-0.56)
Fesoterodine ER 8mg q.d	[26]	-0.73 (-0.93,-0.53)	-0.92 (-1.11,-0.72)	-1.17 (-1.43,-0.89)
Tolterodine ER 4mg q.d + Behaviour therapy	[87]	-0.74 (-1.78,-0.22)	-0.92 (-1.98,-0.4)	-1.16 (-2.26,-0.61)
Electromagnetic stimulation	[125]	-0.71 (-2.78,1.06)	-0.89 (-2.97,0.87)	-1.13 (-3.23,0.65)
Solabegron 125mg b.i.d	[55]	-0.69 (-0.98,-0.42)	-0.89 (-1.14,-0.63)	-1.11 (-1.45,-0.77)
Darifenacin ER 30mg q.d	[38]	-0.67 (-1.19,-0.19)	-0.85 (-1.4,-0.35)	-1.09 (-1.63,-0.56)
Solifenacin ER (5mg-10mg) q.d	[31]	-0.67 (-0.88,-0.45)	-0.83 (-1.05,-0.62)	-1.08 (-1.34,-0.79)
Solifenacin ER 5mg - 15mg q.d	[34]	-0.66 (-1.02,-0.29)	-0.85 (-1.22,-0.45)	-1.08 (-1.49,-0.66)
Oxybutynin ER 10mg q.d	[8]	-0.65 (-1.16,-0.31)	-0.83 (-1.35,-0.5)	-1.06 (-1.62,-0.67)
Mirabegron 100mg b.i.d	[48]	-0.64 (-1.04,-0.41)	-0.82 (-1.22,-0.59)	-1.05 (-1.45,-0.76)
Propiverine 30mg b.i.d	[42]	-0.63 (-1.91,0.07)	-0.83 (-2.09,-0.09)	-1.07 (-2.35,-0.38)
Mirabegron 25mg q.d	[50]	-0.61 (-0.83,-0.42)	-0.79 (-1.01,-0.6)	-1.02 (-1.28,-0.74)
Oxybutynin IR 3mg t.i.d	[19]	-0.63 (-0.99,-0.34)	-0.78 (-1.17,-0.5)	-1.03 (-1.44,-0.69)
Mirabegron 150mg b.i.d	[49]	-0.61 (-0.92,-0.38)	-0.8 (-1.13,-0.58)	-1.03 (-1.37,-0.74)
Propiverine ER 30mg q.d	[42]	-0.61 (-1.09,-0.26)	-0.79 (-1.28,-0.41)	-1.02 (-1.53,-0.62)
Mirabegron 100mg q.d	[52]	-0.6 (-0.78,-0.45)	-0.77 (-0.94,-0.6)	-1.01 (-1.26,-0.77)
Mirabegron 200mg q.d	[53]	-0.6 (-0.86,-0.35)	-0.78 (-1.06,-0.53)	-1.02 (-1.33,-0.71)
Mirabegron 50mg q.d	[51]	-0.59 (-0.74,-0.44)	-0.8 (-0.96,-0.64)	-1.01 (-1.25,-0.77)
Solifenacin ER 5mg q.d	[29]	-0.58 (-0.76,-0.38)	-0.72 (-0.9,-0.55)	-0.99 (-1.24,-0.74)
Trospium ER 60mg q.d	[44]	-0.54 (-0.9,-0.17)	-0.72 (-1.07,-0.36)	-0.96 (-1.35,-0.61)
Darifenacin ER 15mg q.d	[40]	-0.54 (-0.98,-0.11)	-0.72 (-1.18,-0.26)	-0.95 (-1.41,-0.46)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.54 (-1.05,-0.06)	-0.74 (-1.28,-0.25)	-0.97 (-1.54,-0.45)
Cizolirtine Citrate 400mg b.i.d	[57]	-0.53 (-1.06,0.02)	-0.73 (-1.25,-0.15)	-0.96 (-1.51,-0.39)
Pregabalin 150mg b.i.d	[62]	-0.51 (-1.05,0)	-0.7 (-1.23,-0.23)	-0.94 (-1.48,-0.4)
Fesoterodine ER 4mg q.d	[25]	-0.51 (-0.68,-0.36)	-0.7 (-0.85,-0.55)	-0.94 (-1.17,-0.7)
Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	-0.51 (-0.81,-0.21)	-0.7 (-1.02,-0.39)	-0.92 (-1.32,-0.58)
Propiverine 45mg t.i.d	[118]	-0.47 (-1.18,0.37)	-0.66 (-1.35,0.16)	-0.9 (-1.58,-0.07)
Tolterodine ER 4mg q.d	[4]	-0.49 (-0.59,-0.4)	-0.63 (-0.73,-0.52)	-0.89 (-1.08,-0.67)
Propiverine IR 15mg b.i.d	[43]	-0.48 (-0.82,-0.14)	-0.64 (-0.99,-0.29)	-0.9 (-1.28,-0.51)
Oxybutynin IR 5mg t.i.d	[7]	-0.49 (-0.77,-0.23)	-0.66 (-0.94,-0.4)	-0.89 (-1.24,-0.57)
Oxybutynin intravesically 5mg t.i.d	[14]	-0.48 (-1.05,-0.08)	-0.66 (-1.23,-0.27)	-0.89 (-1.47,-0.47)
Tolterodine IR 2mg b.i.d	[5]	-0.45 (-0.58,-0.33)	-0.65 (-0.79,-0.52)	-0.87 (-1.11,-0.64)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.46 (-0.83,-0.14)	-0.66 (-1.04,-0.35)	-0.88 (-1.3,-0.52)
Darifenacin ER 7.5mg q.d	[39]	-0.46 (-0.89,0.22)	-0.65 (-1.1,0.05)	-0.89 (-1.36,-0.21)
Tolterodine IR 2mg b.i.d + PFE	[95]	-0.45 (-1.08,0.15)	-0.63 (-1.29,-0.01)	-0.87 (-1.57,-0.23)
Propiverine ER 20mg q.d	[41]	-0.45 (-0.61,-0.26)	-0.64 (-0.81,-0.46)	-0.87 (-1.13,-0.62)
Oxybutynin chloride topical gel 1g q.d	[13]	-0.44 (-0.78,-0.14)	-0.62 (-0.96,-0.32)	-0.86 (-1.25,-0.52)
Oxybutynin IR (2.5-5mg) b.i.d	[24]	-0.44 (-0.79,-0.12)	-0.61 (-0.95,-0.3)	-0.86 (-1.24,-0.5)
Fesoterodine ER (4mg-8mg) q.d	[27]	-0.42 (-0.62,-0.21)	-0.64 (-0.82,-0.45)	-0.88 (-1.13,-0.6)
Tolterodine IR 1mg b.i.d	[6]	-0.42 (-0.63,-0.15)	-0.61 (-0.83,-0.34)	-0.84 (-1.1,-0.52)
PFMT + BT	[89]	-0.41 (-0.9,0.18)	-0.61 (-1.13,0)	-0.84 (-1.36,-0.26)
Oxybutynin gel 84mg/day	[134]	-0.41 (-0.79,-0.03)	-0.61 (-0.99,-0.24)	-0.83 (-1.24,-0.44)

TABLE E.2: Sensitivity analysis assessing the impact of studies A157 and A158 on the estimated posterior median differences (and 95% credible intervals) obtained from multivariate hierarchical network meta-analyses (cont.)

Oxybutynin transdermal 3.9mg/day	[10]	-0.41 (-0.65,-0.16)	-0.6 (-0.84,-0.35)	-0.84 (-1.12,-0.53)
Oxybutynin vaginal ring 4mg q.d	[16]	-0.4 (-0.77,-0.02)	-0.57 (-0.95,-0.18)	-0.81 (-1.26,-0.43)
Oxybutynin 20mg intravesically q.d	[106]	-0.39 (-0.81,0.11)	-0.56 (-0.99,-0.09)	-0.8 (-1.25,-0.32)
Imidafenacin IR 0.1mg b.i.d	[36]	-0.37 (-0.63,-0.13)	-0.53 (-0.79,-0.29)	-0.77 (-1.11,-0.46)
Oxbutynin patch 73.5mg	[15]	-0.37 (-0.63,-0.07)	-0.55 (-0.86,-0.24)	-0.79 (-1.13,-0.44)
Oxybutynin gel 56mg/day	[135]	-0.36 (-0.7,0.09)	-0.52 (-0.85,-0.07)	-0.76 (-1.14,-0.32)
Oxybutynin IR (5-20mg)	[23]	-0.36 (-0.75,0.32)	-0.55 (-0.95,0.13)	-0.78 (-1.22,-0.12)
Oxybutynin ER 15mg q.d	[9]	-0.35 (-0.76,0.22)	-0.53 (-0.92,0.04)	-0.78 (-1.2,-0.16)
Terodiline IR 25mg b.i.d	[28]	-0.35 (-0.8,0.06)	-0.51 (-0.94,-0.1)	-0.76 (-1.22,-0.29)
PFMT	[84]	-0.34 (-0.83,0.16)	-0.54 (-1.03,-0.02)	-0.77 (-1.29,-0.23)
Elocalcitol 75mg	[70]	-0.35 (-0.94,0.13)	-0.5 (-1.13,0)	-0.75 (-1.41,-0.23)
Imidafenacin IR 0.05mg b.i.d	[35]	-0.33 (-0.73,0.11)	-0.52 (-0.93,-0.08)	-0.74 (-1.2,-0.26)
Elocalcitol 150mg	[69]	-0.33 (-0.92,0.19)	-0.51 (-1.1,0.01)	-0.75 (-1.34,-0.19)
Oxybutynin ER (5-30mg) q.d	[22]	-0.34 (-0.67,0.19)	-0.52 (-0.84,0.02)	-0.76 (-1.14,-0.19)
Solabegron 50mg b.i.d	[54]	-0.31 (-0.59,-0.05)	-0.5 (-0.76,-0.26)	-0.73 (-1.06,-0.41)
Trospium chloride 45mg t.i.d	[47]	-0.31 (-0.74,0.12)	-0.5 (-0.9,-0.08)	-0.72 (-1.2,-0.28)
Oxybutynin transdermal 1.3mg/day	[11]	-0.31 (-0.6,0.22)	-0.48 (-0.78,0.04)	-0.72 (-1.05,-0.16)
Oxybutynin ER 2.5mg q.d	[20]	-0.29 (-0.69,0.36)	-0.5 (-0.9,0.16)	-0.7 (-1.17,-0.03)
Duloxetine IR 40mg b.i.d	[65]	-0.29 (-0.76,0.21)	-0.47 (-0.97,0.07)	-0.71 (-1.23,-0.14)
Propiverine 60mg q.d	[119]	-0.28 (-0.91,0.99)	-0.46 (-1.09,0.84)	-0.67 (-1.37,0.61)
Bladder Training	[85]	-0.27 (-0.63,0.17)	-0.45 (-0.81,0.02)	-0.68 (-1.09,-0.22)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.25 (-0.66,0.19)	-0.44 (-0.86,-0.01)	-0.67 (-1.12,-0.2)
Oxybutynin transdermal 2.6mg/day	[12]	-0.24 (-0.55,0.32)	-0.43 (-0.72,0.12)	-0.66 (-1.02,-0.07)
Serlopitant 0.25mg q.d	[107]	-0.22 (-0.7,0.24)	-0.41 (-0.86,0.03)	-0.65 (-1.11,-0.14)
Tarafenacin 0.4mg q.d	[82]	-0.22 (-0.83,0.32)	-0.41 (-1.02,0.14)	-0.64 (-1.28,-0.07)
Serlopitant 4mg q.d	[109]	-0.21 (-0.7,0.23)	-0.39 (-0.86,0.02)	-0.63 (-1.09,-0.16)
Oxybutynin IR 5mg b.i.d	[18]	-0.2 (-0.56,0.42)	-0.39 (-0.74,0.24)	-0.61 (-1.02,0)
Cizolirtine citrate 200mg b.i.d	[56]	-0.21 (-1.32,0.87)	-0.39 (-1.5,0.71)	-0.62 (-1.74,0.45)
Lipo-BoNTA	[138]	-0.17 (-0.98,0.63)	-0.36 (-1.19,0.45)	-0.57 (-1.43,0.26)
Oxybutynin ER 2.5mg q.d + Bladder training	[92]	-0.16 (-0.97,0.67)	-0.34 (-1.16,0.49)	-0.56 (-1.42,0.22)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.12 (-0.96,0.88)	-0.28 (-1.1,0.74)	-0.58 (-1.44,0.45)
Trospium IR 15mg t.i.d	[46]	-0.1 (-0.62,1.41)	-0.26 (-0.78,1.23)	-0.5 (-1.07,0.93)
Percutaneous tibial nerve stimulation	[83]	-0.1 (-1,0.7)	-0.29 (-1.21,0.49)	-0.54 (-1.49,0.33)
Serlopitant 1mg q.d	[108]	-0.06 (-0.55,0.43)	-0.23 (-0.69,0.23)	-0.48 (-0.98,0.06)
Electrostimulation	[80]	-0.07 (-0.8,0.81)	-0.26 (-1,0.61)	-0.46 (-1.2,0.41)
Estradiol 25mg	[68]	0 (-0.38,0.36)	-0.18 (-0.6,0.2)	-0.41 (-0.87,-0.03)
Placebo	[1]	NA	NA	NA
Tarafenacin 0.2mg q.d	[90]	0.03 (-0.62,0.62)	-0.15 (-0.79,0.45)	-0.37 (-1.07,0.25)
ZD0947IL 25mg/day	[58]	0.05 (-0.65,0.9)	-0.12 (-0.82,0.75)	-0.36 (-1.08,0.55)
Netupitant 100mg q.d	[111]	0.05 (-0.67,0.87)	-0.13 (-0.84,0.68)	-0.35 (-1.13,0.42)
Netupitant 50mg q.d	[110]	0.1 (-0.61,0.91)	-0.08 (-0.8,0.72)	-0.32 (-1.06,0.48)
Netupitant 200mg q.d	[112]	0.14 (-0.88,0.94)	-0.05 (-1.05,0.74)	-0.26 (-1.33,0.54)
ONO-8539 100mg b.i.d	[60]	0.18 (-0.51,0.77)	0.01 (-0.69,0.6)	-0.23 (-0.94,0.4)
Vaginal oestrogen cream 1.25mg/day	[132]	0.28 (-0.59,1.21)	0.08 (-0.81,1.02)	-0.1 (-0.97,0.82)
Emepronium bromide 200mg q.d	[63]	0.31 (-0.54,1.05)	0.12 (-0.74,0.89)	-0.11 (-1.04,0.7)
Flavoxate chloride 200mg q.d	[64]	0.31 (-0.31,1.09)	0.12 (-0.53,0.94)	-0.12 (-0.78,0.75)
Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	0.28 (-0.69,1.25)	0.1 (-0.89,1.06)	-0.13 (-1.11,0.86)
Resiniferatoxin 50nM	[67]	0.36 (-0.68,1.49)	0.18 (-0.85,1.28)	-0.06 (-1.05,1.09)
ONO-8539 300mg b.i.d	[61]	0.44 (-0.07,1.04)	0.26 (-0.27,0.85)	0.02 (-0.54,0.61)
ONO-8539 30mg b.i.d	[59]	0.48 (-0.07,1.14)	0.3 (-0.25,0.95)	0.05 (-0.51,0.76)
Estradiol 1mg intravaginally	[127]	0.51 (-0.98,1.75)	0.33 (-1.15,1.57)	0.1 (-1.32,1.35)
Propantheline Bromide 15mg t.i.d	[113]	0.57 (-0.87,1.58)	0.37 (-1.04,1.39)	0.16 (-1.34,1.19)
Control	[2]	0.51 (-0.14,1.72)	0.37 (-0.3,1.55)	0.11 (-0.56,1.24)
Sham Therapy	[3]	0.55 (-0.4,1.5)	0.38 (-0.6,1.33)	0.14 (-0.9,1.14)
Reflexology	[71]	0.57 (-0.24,1.94)	0.37 (-0.45,1.75)	0.15 (-0.66,1.47)
Naftopidil 25mg q.d	[114]	3.15 (1.46,5.32)	2.96 (1.28,5.19)	2.72 (1.05,4.93)
Solifenacin succinate 5mg q.d + Naftopidil 25mg q.d	[115]	5.23 (3.25,7.61)	5.05 (3.08,7.45)	4.82 (2.82,7.17)

† median relative to a placebo intervention

E.5 Sensitivity analysis

TABLE E.3: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis

Treatment	**Code**	**Median difference† (95%CrI)**	**Median difference†† (95%CrI)**	**Median difference§ (95%CrI)**	**Median difference§§ (95%CrI)**
Sacral nerve stimulation	[81]	-8.58 (-10.95,-6.13)	-7.87 (-9.75,-6.08)	-8.53 (-10.46,-5.77)	-7.95 (-10.38,-5.62)
OnaBoNT-A 200u trigone sparing	[73]	-2.09 (-2.98,-1.42)	-2 (-2.95,-1.39)	-1.98 (-3.09,-1.39)	-2.03 (-2.96,-1.42)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-1.87 (-3.9,0.08)	-2.08 (-3.69,-0.06)	-1.99 (-3.6,-0.26)	-1.65 (-3.37,0.36)
Electrostimulation + PFE + Bladder training	[97]	-1.99 (-2.8,-1.06)	-1.82 (-2.79,-0.94)	-1.93 (-2.86,-0.67)	-1.86 (-2.83,-0.93)
Solifenacin/trospium + placebo injection	[100]	-1.78 (-2.86,-0.87)	-1.67 (-2.63,-0.78)	-1.7 (-2.68,-0.75)	-1.71 (-2.66,-0.86)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.74 (-3.11,-0.17)	-1.67 (-3.27,-0.47)	-1.74 (-2.82,-0.79)	-1.77 (-2.88,-0.96)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.65 (-2.44,-0.84)	-1.54 (-2.42,-0.87)	-1.65 (-2.33,-0.82)	-1.64 (-2.35,-0.85)
OnaBoNTA 100u trigone sparing	[72]	-1.69 (-2.17,-1.31)	-1.58 (-2.06,-1.28)	-1.62 (-2.05,-1.25)	-1.64 (-2.03,-1.29)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.33 (-1.75,-0.92)	-1.32 (-1.74,-0.95)	-1.36 (-1.77,-0.94)	-1.33 (-1.73,-0.93)
Estriol 1mg intravesival	[131]	-1.21 (-4.51,16.41)	-1.51 (-12.74,-0.35)	-1.12 (-2.41,-0.06)	-1.24 (-2.52,-0.02)
Trospium IR 15mg t.i.d + Physiotherapy	[91]	-1.07 (-1.95,-0.07)	-0.94 (-1.98,-0.04)	-1.02 (-1.86,-0.12)	-1.07 (-1.92,-0.25)
Estradiol 3mg intravaginally	[128]	-1.25 (-17.22,0.4)	-0.87 (-3.19,1.97)	-0.84 (-2.65,0.65)	-1.13 (-2.95,0.22)
Solifenacin ER 10mg q.d	[30]	-0.84 (-1.07,-0.62)	-0.86 (-1.06,-0.63)	-0.85 (-1.07,-0.63)	-0.84 (-1.09,-0.62)
Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	-0.74 (-1.24,-0.24)	-0.78 (-1.26,-0.26)	-0.7 (-1.21,-0.22)	-0.74 (-1.24,-0.22)
Tolterodine ER 4mg q.d + Behaviour therapy	[87]	-0.69 (-1.35,-0.1)	-0.67 (-1.34,-0.01)	-0.7 (-1.35,-0.15)	-0.71 (-1.32,-0.1)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.68 (-1.46,0.09)	-0.84 (-1.46,-0.02)	-0.73 (-1.44,-0.11)	-0.73 (-1.44,-0.09)
Fesoterodine ER 8mg q.d	[26]	-0.69 (-0.88,-0.51)	-0.69 (-0.88,-0.51)	-0.7 (-0.89,-0.52)	-0.69 (-0.88,-0.5)
Imidafenacin IR 0.25mg b.i.d	[37]	-0.76 (-1.34,-0.22)	-0.76 (-1.29,-0.21)	-0.75 (-1.27,-0.26)	-0.75 (-1.34,-0.23)
Solifenacin ER (5mg-10mg) q.d	[31]	-0.68 (-0.92,-0.44)	-0.67 (-0.91,-0.45)	-0.68 (-0.9,-0.45)	-0.67 (-0.89,-0.45)
Solifenacin ER 5mg - 15mg q.d	[34]	-0.65 (-1.05,-0.23)	-0.65 (-1.06,-0.22)	-0.66 (-1.03,-0.26)	-0.65 (-1.01,-0.23)
Mirabegron 100mg b.i.d	[48]	-0.68 (-1.55,-0.37)	-0.66 (-1.34,-0.39)	-0.65 (-1.04,-0.39)	-0.65 (-1.09,-0.39)
Solabegron 125mg b.i.d	[55]	-0.63 (-0.92,-0.32)	-0.63 (-0.9,-0.34)	-0.63 (-0.9,-0.34)	-0.63 (-0.91,-0.34)
Propiverine ER 30mg q.d	[42]	-0.62 (-1.15,-0.2)	-0.64 (-1.21,-0.21)	-0.57 (-1.08,-0.25)	-0.58 (-1.14,-0.22)
Darifenacin ER 30mg q.d	[38]	-0.64 (-1.25,-0.07)	-0.65 (-1.22,-0.15)	-0.61 (-1.2,-0.11)	-0.66 (-1.28,-0.13)
Mirabegron 25mg q.d	[50]	-0.63 (-0.9,-0.4)	-0.62 (-0.86,-0.41)	-0.63 (-0.86,-0.41)	-0.62 (-0.86,-0.41)
Mirabegron 150mg b.i.d	[49]	-0.6 (-1.16,0.08)	-0.62 (-1.05,-0.11)	-0.61 (-0.97,-0.31)	-0.61 (-1.02,-0.31)
Mirabegron 100mg q.d	[52]	-0.61 (-0.79,-0.43)	-0.6 (-0.78,-0.43)	-0.6 (-0.78,-0.43)	-0.6 (-0.78,-0.43)
Mirabegron 200mg q.d	[53]	-0.6 (-0.97,-0.19)	-0.61 (-0.92,-0.27)	-0.61 (-0.9,-0.3)	-0.6 (-0.89,-0.31)
Oxybutynin ER 10mg q.d	[8]	-0.68 (-1.33,-0.27)	-0.64 (-1.22,-0.28)	-0.61 (-1.13,-0.26)	-0.63 (-1.07,-0.27)
Propiverine 30mg b.i.d	[42]	-0.69 (-20.04,0.22)	-0.81 (-14.1,-0.16)	-0.58 (-1.76,-0.09)	-0.64 (-2.03,-0.1)
Oxybutynin IR 3mg t.i.d	[19]	-0.66 (-1.14,-0.32)	-0.61 (-1.06,-0.33)	-0.6 (-0.96,-0.31)	-0.62 (-1,-0.32)
Mirabegron 50mg q.d	[51]	-0.58 (-0.74,-0.41)	-0.58 (-0.74,-0.42)	-0.58 (-0.75,-0.42)	-0.58 (-0.75,-0.42)
Solifenacin ER 5mg q.d	[29]	-0.6 (-0.82,-0.39)	-0.59 (-0.8,-0.4)	-0.58 (-0.79,-0.41)	-0.6 (-0.79,-0.39)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.47 (-1.06,0.25)	-0.51 (-1.1,0.19)	-0.51 (-1.01,-0.03)	-0.54 (-1.03,-0.04)
Cizolirtine Citrate 400mg b.i.d	[57]	-0.54 (-1.1,-0.01)	-0.6 (-1.08,-0.08)	-0.5 (-1.06,0)	-0.54 (-1.04,0.01)
Trospium ER 60mg q.d	[44]	-0.58 (-0.97,-0.19)	-0.55 (-0.93,-0.22)	-0.56 (-0.92,-0.2)	-0.56 (-0.91,-0.21)
Propiverine 45mg t.i.d	[118]	-0.44 (-4.47,7.68)	-0.43 (-1.67,7.71)	-0.47 (-1.27,0.21)	-0.49 (-1.45,0.41)
Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	-0.49 (-0.8,-0.18)	-0.49 (-0.8,-0.2)	-0.51 (-0.79,-0.21)	-0.5 (-0.81,-0.21)
Fesoterodine ER 4mg q.d	[25]	-0.5 (-0.66,-0.33)	-0.5 (-0.67,-0.34)	-0.5 (-0.66,-0.34)	-0.5 (-0.66,-0.33)
Pregabalin 150mg b.i.d	[62]	-0.43 (-0.97,0.53)	-0.48 (-1.02,0.45)	-0.47 (-1.06,0.05)	-0.49 (-1.08,0.1)
Darifenacin ER 15mg q.d	[40]	-0.49 (-0.91,-0.03)	-0.5 (-0.91,-0.11)	-0.5 (-0.91,-0.08)	-0.5 (-0.93,-0.08)
Propiverine IR 15mg b.i.d	[43]	-0.48 (-0.87,-0.12)	-0.52 (-0.86,-0.19)	-0.48 (-0.83,-0.15)	-0.48 (-0.85,-0.14)
Tolterodine ER 4mg q.d	[4]	-0.49 (-0.6,-0.39)	-0.49 (-0.59,-0.38)	-0.49 (-0.59,-0.39)	-0.49 (-0.59,-0.39)
Oxybutynin IR 5mg t.i.d	[7]	-0.48 (-0.87,-0.19)	-0.5 (-0.82,-0.22)	-0.46 (-0.75,-0.19)	-0.47 (-0.75,-0.2)
Propiverine ER 20mg q.d	[41]	-0.42 (-0.61,-0.23)	-0.44 (-0.62,-0.24)	-0.43 (-0.61,-0.25)	-0.43 (-0.61,-0.25)
Tolterodine IR 2mg b.i.d	[5]	-0.44 (-0.57,-0.31)	-0.43 (-0.57,-0.31)	-0.44 (-0.56,-0.31)	-0.44 (-0.57,-0.31)
Propiverine 60mg q.d	[119]	-0.24 (-0.91,48.73)	-0.41 (-1.07,25.03)	-0.41 (-1.07,0.93)	-0.39 (-0.99,0.75)
Oxybutynin intravesically 5mg t.i.d	[14]	-0.49 (-1.85,-0.03)	-0.44 (-1.61,0)	-0.45 (-1.04,-0.01)	-0.44 (-0.99,-0.03)
Oxybutynin IR (2.5-5mg) b.i.d	[24]	-0.47 (-1.3,-0.05)	-0.42 (-1.01,-0.08)	-0.44 (-0.86,-0.05)	-0.43 (-0.79,-0.09)
Oxybutynin chloride topical gel 1g q.d	[13]	-0.43 (-0.86,-0.08)	-0.43 (-0.79,-0.11)	-0.41 (-0.73,-0.1)	-0.41 (-0.75,-0.11)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.41 (-0.88,0.01)	-0.42 (-0.81,-0.03)	-0.42 (-0.83,-0.07)	-0.41 (-0.78,-0.06)
Tolterodine IR 2mg b.i.d + PFE	[95]	-0.43 (-1.01,0.18)	-0.4 (-1.13,0.18)	-0.4 (-1.03,0.29)	-0.44 (-1.04,0.23)
PFMT + BT	[89]	-0.41 (-0.91,0.21)	-0.37 (-0.9,0.31)	-0.38 (-0.94,0.24)	-0.38 (-0.92,0.16)
Tolterodine IR 1mg b.i.d	[6]	-0.39 (-0.65,-0.08)	-0.39 (-0.62,-0.1)	-0.41 (-0.63,-0.07)	-0.4 (-0.63,-0.11)
Fesoterodine ER (4mg-8mg) q.d	[27]	-0.37 (-0.59,-0.13)	-0.4 (-0.59,-0.14)	-0.39 (-0.59,-0.18)	-0.4 (-0.59,-0.17)
Oxybutynin gel 84mg/day	[134]	-0.37 (-0.85,0.19)	-0.34 (-0.79,0.06)	-0.38 (-0.73,-0.01)	-0.38 (-0.71,0)
Oxybutynin transdermal 3.9mg/day	[10]	-0.38 (-0.65,-0.09)	-0.37 (-0.63,-0.12)	-0.39 (-0.63,-0.14)	-0.39 (-0.64,-0.14)

TABLE E.3: Sensitivity analyses assessing the impact in change from baseline in incontinence episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis (cont.)

Oxybutynin vaginal ring 4mg q.d	[16]	-0.38 (-0.86,0.05)	-0.36 (-0.79,0.03)	-0.38 (-0.74,-0.03)	-0.37 (-0.72,0.01)
Imidafenacin IR 0.1mg b.i.d	[36]	-0.38 (-0.64,-0.12)	-0.37 (-0.64,-0.1)	-0.37 (-0.63,-0.12)	-0.38 (-0.63,-0.13)
Terodiline IR 25mg b.i.d	[28]	-0.41 (-1,0.06)	-0.36 (-1.01,0.07)	-0.37 (-0.77,0.04)	-0.37 (-0.81,0.08)
Darifenacin ER 7.5mg q.d	[39]	-0.34 (-0.85,0.19)	-0.39 (-0.87,0.15)	-0.36 (-0.81,0.22)	-0.39 (-0.89,0.1)
Oxybutynin gel 56mg/day	[135]	-0.4 (-1.17,0.07)	-0.37 (-1.12,0.06)	-0.36 (-0.74,0.02)	-0.36 (-0.73,0.05)
Oxbutynin patch 73.5mg	[15]	-0.33 (-0.62,-0.02)	-0.35 (-0.6,-0.05)	-0.34 (-0.61,-0.03)	-0.34 (-0.61,-0.06)
Elocalcitol 75mg	[70]	-0.38 (-1,0.2)	-0.42 (-0.94,0.14)	-0.31 (-0.88,0.21)	-0.37 (-0.91,0.16)
Oxybutynin 20mg intravesically q.d	[106]	-0.38 (-20.48,6.04)	-0.32 (-2.66,5.3)	-0.33 (-0.75,0.14)	-0.36 (-0.89,0.14)
Imidafenacin IR 0.05mg b.i.d	[35]	-0.32 (-0.8,0.16)	-0.29 (-0.71,0.12)	-0.33 (-0.73,0.12)	-0.34 (-0.77,0.09)
Oxybutynin ER 15mg q.d	[9]	-0.3 (-0.82,0.47)	-0.28 (-0.76,0.32)	-0.31 (-0.69,0.19)	-0.29 (-0.67,0.18)
Oxybutynin IR (5-20mg)	[23]	-0.27 (-0.71,1.34)	-0.28 (-0.7,1.3)	-0.32 (-0.71,0.28)	-0.31 (-0.72,0.25)
Oxybutynin ER (5-30mg) q.d	[22]	-0.27 (-0.62,0.56)	-0.26 (-0.61,0.61)	-0.29 (-0.63,0.16)	-0.31 (-0.63,0.13)
Trospium chloride 45mg t.i.d	[47]	-0.22 (-0.81,64.6)	-0.23 (-0.74,39.26)	-0.3 (-0.81,0.3)	-0.3 (-0.81,0.25)
Cizolirtine citrate 200mg b.i.d	[56]	-0.32 (-1.37,0.84)	-0.39 (-1.27,1.04)	-0.25 (-1.28,0.93)	-0.24 (-1.17,0.83)
PFMT	[84]	-0.27 (-0.83,0.41)	-0.18 (-0.82,0.39)	-0.25 (-0.83,0.33)	-0.28 (-0.84,0.29)
Oxybutynin transdermal 1.3mg/day	[11]	-0.25 (-0.66,0.35)	-0.28 (-0.6,0.28)	-0.25 (-0.58,0.19)	-0.27 (-0.6,0.16)
Elocalcitol 150mg	[69]	-0.29 (-0.89,0.32)	-0.32 (-0.85,0.2)	-0.24 (-0.8,0.28)	-0.32 (-0.87,0.22)
Oxybutynin ER 2.5mg q.d	[20]	-0.2 (-0.61,1.17)	-0.24 (-0.69,1.04)	-0.27 (-0.65,0.28)	-0.25 (-0.61,0.3)
Duloxetine IR 40mg b.i.d	[65]	-0.26 (-0.78,0.24)	-0.27 (-0.77,0.21)	-0.24 (-0.8,0.29)	-0.24 (-0.77,0.26)
Bladder Training	[85]	-0.23 (-0.63,0.45)	-0.21 (-0.64,0.47)	-0.24 (-0.62,0.17)	-0.25 (-0.63,0.16)
Solabegron 50mg b.i.d	[54]	-0.24 (-0.55,0.06)	-0.27 (-0.52,0.04)	-0.26 (-0.55,0.03)	-0.25 (-0.53,0.03)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.17 (-0.47,0.19)	-0.2 (-0.47,0.14)	-0.21 (-0.49,0.09)	-0.21 (-0.48,0.14)
Oxybutynin transdermal 2.6mg/day	[12]	-0.14 (-0.52,0.64)	-0.17 (-0.52,0.53)	-0.19 (-0.52,0.33)	-0.19 (-0.52,0.31)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.16 (-0.67,0.35)	-0.18 (-0.67,0.31)	-0.22 (-0.68,0.22)	-0.22 (-0.71,0.31)
Oxybutynin ER 2.5mg q.d + Bladder training	[92]	-0.14 (-1.02,0.94)	-0.21 (-1.03,0.97)	-0.29 (-1.4,0.55)	-0.16 (-1.05,0.81)
Oxybutynin IR 5mg b.i.d	[18]	-0.11 (-0.5,0.7)	-0.12 (-0.56,0.59)	-0.16 (-0.51,0.39)	-0.15 (-0.52,0.42)
Lipo-BoNTA	[138]	-0.15 (-1.02,0.75)	-0.12 (-1.05,0.84)	-0.15 (-0.98,0.61)	-0.18 (-1.02,0.75)
Serlopitant 0.25mg q.d	[107]	-0.05 (-0.58,31.41)	-0.19 (-1.84,2.8)	-0.09 (-0.62,0.35)	-0.14 (-0.7,0.35)
Serlopitant 4mg q.d	[109]	-0.11 (-4.43,15.12)	-0.19 (-13.84,0.31)	-0.1 (-0.67,0.34)	-0.14 (-0.69,0.35)
Tarafenacin 0.4mg q.d	[82]	-0.19 (-0.84,0.44)	-0.23 (-1.1,0.4)	-0.2 (-0.81,0.4)	-0.14 (-0.75,0.43)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.02 (-0.7,0.67)	-0.05 (-0.69,0.61)	-0.1 (-0.76,0.59)	-0.1 (-0.78,0.67)
Electrostimulation	[80]	0 (-0.48,0.51)	-0.05 (-0.51,0.44)	-0.08 (-0.52,0.39)	-0.08 (-0.53,0.43)
Serlopitant 1mg q.d	[108]	0.08 (-0.51,24.38)	0 (-11.79,5.32)	0.05 (-0.45,0.63)	0.01 (-0.5,0.56)
Estradiol 25mg	[68]	0 (-0.39,0.4)	0.01 (-0.38,0.37)	-0.01 (-0.41,0.4)	-0.01 (-0.4,0.38)
Placebo	[1]	NA	NA	NA	NA
Netupitant 200mg q.d	[112]	-0.09 (-21.62,1.02)	0.08 (-46.75,1.27)	0.04 (-0.91,1.13)	-0.06 (-1.13,0.97)
Netupitant 100mg q.d	[111]	0.17 (-13.54,7.35)	0.23 (-1.46,3.13)	0.11 (-0.89,1.06)	-0.06 (-1.13,1.02)
Tarafenacin 0.2mg q.d	[90]	0.05 (-0.57,0.72)	0.02 (-0.56,0.7)	-0.01 (-0.62,0.64)	0.09 (-0.5,0.68)
Trospium IR 15mg t.i.d	[46]	0.09 (-0.69,1.15)	0.1 (-0.64,1.13)	0.1 (-0.61,1.11)	0.02 (-0.65,0.88)
ZD0947IL 25mg/day	[58]	0.06 (-0.79,0.95)	0.05 (-0.81,1.1)	0.11 (-0.69,0.98)	0.14 (-0.65,1.11)
Netupitant 50mg q.d	[110]	0.26 (-6.15,14.21)	0.1 (-25.33,1.2)	0.11 (-0.89,1.27)	0.04 (-0.95,1.11)
Electromagnetic stimulation	[125]	-0.02 (-2.12,29.45)	0.14 (-1.99,48.32)	-0.28 (-1.82,1.51)	-0.39 (-1.87,1.47)
Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	0.36 (-0.51,1.7)	0.47 (-0.57,1.8)	0.37 (-0.62,1.25)	0.28 (-0.59,1.31)
ONO-8539 100mg b.i.d	[60]	0.13 (-0.65,0.73)	0.12 (-0.56,0.77)	0.11 (-0.6,0.79)	0.13 (-0.51,0.72)
Percutaneous tibial nerve stimulation	[83]	-0.03 (-1.07,0.91)	0.06 (-1.04,0.84)	0.02 (-0.86,0.92)	0.06 (-0.88,0.93)
Vaginal oestrogen cream 1.25mg/day	[132]	0.31 (-0.23,0.9)	0.24 (-0.3,0.83)	0.23 (-0.28,0.75)	0.25 (-0.33,0.79)
Flavoxate chloride 200mg q.d	[64]	0.35 (-0.32,1.02)	0.36 (-0.35,1.12)	0.36 (-0.32,1.06)	0.36 (-0.34,1.02)
Resiniferatoxin 50nM	[67]	0.38 (-0.82,2.02)	0.43 (-1.08,1.46)	0.43 (-0.63,1.61)	0.37 (-0.64,1.67)
Emepronium bromide 200mg q.d	[63]	0.38 (-0.26,1.02)	0.39 (-0.34,1.03)	0.36 (-0.31,1.2)	0.35 (-0.34,1.06)
ONO-8539 300mg b.i.d	[61]	0.48 (-0.15,1.08)	0.48 (-0.16,1.01)	0.42 (-0.16,0.99)	0.43 (-0.15,1)
Propantheline Bromide 15mg t.i.d	[113]	0.38 (-9.42,1.6)	0.34 (-31.19,1.48)	0.62 (-0.53,1.92)	0.49 (-0.61,1.63)
Estradiol 1mg intravaginally	[127]	0.46 (-52.47,2.32)	0.42 (-18.4,2.01)	0.6 (-0.76,1.91)	0.51 (-0.89,1.9)
ONO-8539 30mg b.i.d	[59]	0.51 (-0.12,1.14)	0.46 (-0.1,1.07)	0.5 (-0.13,1.08)	0.5 (-0.11,1.11)
Control	[2]	0.52 (-0.22,1.44)	0.68 (-0.11,1.47)	0.58 (-0.16,1.29)	0.59 (-0.28,1.3)
Reflexology	[71]	0.54 (-0.31,1.53)	0.72 (-0.17,1.62)	0.62 (-0.22,1.39)	0.62 (-0.3,1.41)
Sham Therapy	[3]	0.59 (-0.59,1.6)	0.66 (-0.51,1.54)	0.68 (-0.31,1.64)	0.71 (-0.28,1.64)
Naftopidil 25mg q.d	[114]	3.38 (-43.89,6.7)	3.95 (0.79,12.92)	3.55 (1.11,6.03)	3.86 (2.21,5.77)
Solifenacin succinate 5mg q.d + Naftopidil 25mg q.d	[115]	6.41 (0.96,25.19)	6.08 (3.56,28.61)	5.15 (2.89,7.53)	5.95 (4.1,8.26)

† Elements of $\boldsymbol{V^{1/2}}$ based on a Gamma(0.001,0.001) prior distribution on the precision scale

†† Elements of $\boldsymbol{V^{1/2}}$ based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

§ Deviance of treatment effect profiles across outcomes, τ, based on a Gamma(0.01,0.01) prior distribution on the precision scale

§§ Deviance of treatment effect profiles across outcomes, τ, based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE E.4: Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis

Treatment	**Code**	**Median difference† (95%CrI)**	**Median difference†† (95%CrI)**	**Median difference§ (95%CrI)**	**Median difference§§ (95%CrI)**
Sacral nerve stimulation	[81]	-8.62 (-11.32,-3.88)	-8.07 (-9.87,-5.77)	-8.75 (-10.67,-5.98)	-8.16 (-10.62,-5.81)
OnaBoNT-A 200u trigone sparing	[73]	-2.34 (-3.74,-1.63)	-2.22 (-3.6,-1.6)	-2.19 (-3.32,-1.61)	-2.25 (-3.18,-1.61)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-2.06 (-10.46,2.26)	-2.38 (-15.9,-0.26)	-2.21 (-3.83,-0.47)	-1.86 (-3.64,0.15)
Electrostimulation + PFE + Bladder training	[97]	-2.3 (-4.23,-1.33)	-2.09 (-3.64,-1.18)	-2.13 (-3.07,-0.88)	-2.05 (-3.02,-1.13)
Solifenacin/trospium + placebo injection	[100]	-1.93 (-12.38,6.14)	-1.91 (-11.17,-0.96)	-1.91 (-2.93,-0.94)	-1.92 (-2.93,-1.07)
OnaBoNT-A 100u bladder base + trigone	[79]	-1.96 (-3.15,-0.72)	-1.88 (-3,-0.67)	-1.96 (-3.02,-1.02)	-1.99 (-3.1,-1.19)
OnaBoNT-A 100u bladder body + trigone	[78]	-1.83 (-2.66,-0.65)	-1.74 (-2.68,-0.97)	-1.85 (-2.55,-1.04)	-1.86 (-2.57,-1.05)
OnaBoNTA 100u trigone sparing	[72]	-1.72 (-2.09,-1.23)	-1.7 (-2.04,-1.32)	-1.75 (-2.14,-1.4)	-1.77 (-2.1,-1.42)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.61 (-2.4,-1.12)	-1.58 (-2.36,-1.16)	-1.6 (-2.03,-1.14)	-1.56 (-2.03,-1.13)
Estriol 1mg intravesival	[131]	-1.48 (-2.79,-0.36)	-1.65 (-3.09,-0.54)	-1.34 (-2.58,-0.29)	-1.45 (-2.72,-0.25)
Trospium IR 15mg t.i.d + Physiotherapy	[91]	-1.29 (-2.24,2.5)	-1.14 (-2.28,1.74)	-1.26 (-2.13,-0.33)	-1.33 (-2.18,-0.46)
Estradiol 3mg intravaginally	[128]	-1.38 (-3.31,0.15)	-1.19 (-3.15,0.9)	-1.05 (-2.87,0.47)	-1.35 (-3.17,-0.03)
Solifenacin ER 10mg q.d	[30]	-1.11 (-1.35,-0.9)	-1.11 (-1.33,-0.9)	-1.1 (-1.31,-0.89)	-1.1 (-1.33,-0.88)
Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	-0.96 (-5.9,6)	-0.97 (-1.51,20.96)	-0.92 (-1.47,-0.4)	-0.96 (-1.49,-0.39)
Tolterodine ER 4mg q.d + Behaviour therapy	[87]	-1 (-13.65,-0.32)	-0.93 (-16.8,-0.22)	-0.91 (-1.6,-0.33)	-0.93 (-1.57,-0.29)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-0.97 (-1.69,-0.33)	-1.1 (-1.66,-0.45)	-0.96 (-1.66,-0.35)	-0.96 (-1.66,-0.33)
Fesoterodine ER 8mg q.d	[26]	-0.93 (-1.14,-0.73)	-0.91 (-1.11,-0.73)	-0.92 (-1.11,-0.74)	-0.92 (-1.1,-0.73)
Imidafenacin IR 0.25mg b.i.d	[37]	-1.02 (-1.74,-0.43)	-0.99 (-1.69,-0.44)	-0.98 (-1.5,-0.47)	-0.97 (-1.59,-0.45)
Solifenacin ER (5mg-10mg) q.d	[31]	-0.84 (-1.07,-0.56)	-0.85 (-1.07,-0.61)	-0.86 (-1.09,-0.64)	-0.85 (-1.07,-0.62)
Solifenacin ER 5mg - 15mg q.d	[34]	-0.93 (-20.61,-0.45)	-0.84 (-1.3,4.3)	-0.87 (-1.28,-0.45)	-0.87 (-1.26,-0.43)
Mirabegron 100mg b.i.d	[48]	-0.88 (-1.62,-0.56)	-0.87 (-1.44,-0.59)	-0.85 (-1.25,-0.6)	-0.86 (-1.3,-0.6)
Solabegron 125mg b.i.d	[55]	-0.87 (-1.18,-0.57)	-0.86 (-1.13,-0.59)	-0.85 (-1.11,-0.57)	-0.86 (-1.14,-0.57)
Propiverine ER 30mg q.d	[42]	-0.75 (-2.14,10.02)	-0.78 (-1.37,11.6)	-0.77 (-1.34,-0.44)	-0.78 (-1.4,-0.4)
Darifenacin ER 30mg q.d	[38]	-0.9 (-16.62,-0.26)	-0.87 (-14.62,-0.32)	-0.81 (-1.44,-0.28)	-0.87 (-1.5,-0.31)
Mirabegron 25mg q.d	[50]	-0.83 (-1.07,-0.57)	-0.83 (-1.06,-0.59)	-0.84 (-1.07,-0.63)	-0.83 (-1.08,-0.61)
Mirabegron 150mg b.i.d	[49]	-0.86 (-1.63,-0.49)	-0.87 (-1.53,-0.56)	-0.84 (-1.21,-0.56)	-0.85 (-1.26,-0.56)
Mirabegron 100mg q.d	[52]	-0.78 (-0.96,-0.56)	-0.78 (-0.95,-0.56)	-0.79 (-0.97,-0.61)	-0.79 (-0.97,-0.61)
Mirabegron 200mg q.d	[53]	-0.78 (-1.13,11.62)	-0.81 (-1.11,19.81)	-0.82 (-1.16,-0.51)	-0.82 (-1.14,-0.51)
Oxybutynin ER 10mg q.d	[8]	-0.9 (-1.69,-0.48)	-0.86 (-1.53,-0.49)	-0.83 (-1.35,-0.47)	-0.85 (-1.3,-0.48)
Propiverine 30mg b.i.d	[42]	-0.88 (-3.05,0)	-1 (-2.71,-0.31)	-0.8 (-1.97,-0.32)	-0.88 (-2.24,-0.33)
Oxybutynin IR 3mg t.i.d	[19]	-0.8 (-1.25,-0.43)	-0.77 (-1.19,-0.47)	-0.78 (-1.16,-0.47)	-0.79 (-1.21,-0.48)
Mirabegron 50mg q.d	[51]	-0.82 (-0.99,-0.64)	-0.82 (-0.98,-0.65)	-0.82 (-0.99,-0.66)	-0.82 (-0.99,-0.66)
Solifenacin ER 5mg q.d	[29]	-0.75 (-0.95,-0.57)	-0.76 (-0.94,-0.58)	-0.75 (-0.93,-0.58)	-0.75 (-0.93,-0.57)
Tolterodine IR 2mg b.i.d + BT	[93]	-0.85 (-2.22,-0.25)	-0.81 (-2.16,-0.17)	-0.76 (-1.29,-0.27)	-0.79 (-1.33,-0.28)
Cizolirtine Citrate 400mg b.i.d	[57]	-0.82 (-2.55,-0.14)	-0.86 (-2.12,-0.29)	-0.73 (-1.34,-0.21)	-0.78 (-1.32,-0.2)
Trospium ER 60mg q.d	[44]	-0.77 (-1.13,-0.38)	-0.74 (-1.11,-0.42)	-0.76 (-1.12,-0.41)	-0.76 (-1.1,-0.41)
Propiverine 45mg t.i.d	[118]	-0.66 (-2.01,0.78)	-0.67 (-1.92,0.37)	-0.69 (-1.46,-0.02)	-0.7 (-1.65,0.18)
Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	-0.73 (-1.15,-0.36)	-0.73 (-1.14,-0.4)	-0.73 (-1.04,-0.41)	-0.73 (-1.07,-0.41)
Fesoterodine ER 4mg q.d	[25]	-0.71 (-0.87,-0.53)	-0.71 (-0.87,-0.55)	-0.7 (-0.86,-0.55)	-0.71 (-0.86,-0.55)
Pregabalin 150mg b.i.d	[62]	-0.75 (-1.26,-0.23)	-0.76 (-1.24,-0.25)	-0.71 (-1.27,-0.23)	-0.73 (-1.3,-0.18)
Darifenacin ER 15mg q.d	[40]	-0.67 (-1.23,-0.05)	-0.7 (-1.15,-0.2)	-0.71 (-1.14,-0.25)	-0.7 (-1.15,-0.25)
Propiverine IR 15mg b.i.d	[43]	-0.62 (-1.02,0.32)	-0.69 (-1.06,-0.11)	-0.66 (-1.04,-0.3)	-0.67 (-1.06,-0.27)
Tolterodine ER 4mg q.d	[4]	-0.63 (-0.74,-0.5)	-0.64 (-0.74,-0.52)	-0.63 (-0.73,-0.53)	-0.63 (-0.73,-0.52)
Oxybutynin IR 5mg t.i.d	[7]	-0.65 (-0.98,-0.33)	-0.68 (-0.99,-0.37)	-0.66 (-0.95,-0.39)	-0.66 (-0.95,-0.4)
Propiverine ER 20mg q.d	[41]	-0.67 (-0.88,-0.48)	-0.67 (-0.86,-0.49)	-0.66 (-0.85,-0.49)	-0.67 (-0.85,-0.5)
Tolterodine IR 2mg b.i.d	[5]	-0.67 (-0.83,-0.52)	-0.66 (-0.82,-0.53)	-0.66 (-0.81,-0.53)	-0.67 (-0.81,-0.52)
Propiverine 60mg q.d	[119]	-0.49 (-1.29,1.75)	-0.64 (-1.35,1.11)	-0.61 (-1.27,0.7)	-0.59 (-1.2,0.54)
Oxybutynin intravesically 5mg t.i.d	[14]	-0.69 (-3.27,-0.03)	-0.64 (-2.05,-0.09)	-0.66 (-1.26,-0.21)	-0.64 (-1.21,-0.22)
Oxybutynin IR (2.5-5mg) b.i.d	[24]	-0.63 (-1.05,-0.22)	-0.62 (-1.02,-0.24)	-0.64 (-1.04,-0.25)	-0.62 (-0.98,-0.28)
Oxybutynin chloride topical gel 1g q.d	[13]	-0.65 (-1.07,-0.3)	-0.64 (-1,-0.32)	-0.63 (-0.95,-0.31)	-0.63 (-0.95,-0.33)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.69 (-1.49,-0.28)	-0.68 (-1.24,-0.34)	-0.65 (-1.06,-0.28)	-0.65 (-1.04,-0.29)
Tolterodine IR 2mg b.i.d + PFE	[95]	-0.72 (-21.4,-0.01)	-0.58 (-1.36,18.18)	-0.62 (-1.27,0.12)	-0.65 (-1.28,0.05)
PFMT + BT	[89]	-0.78 (-2.92,-0.16)	-0.7 (-2.35,-0.1)	-0.63 (-1.26,0.02)	-0.64 (-1.22,-0.05)
Tolterodine IR 1mg b.i.d	[6]	-0.63 (-0.92,-0.29)	-0.62 (-0.9,-0.32)	-0.63 (-0.87,-0.26)	-0.62 (-0.88,-0.33)
Fesoterodine ER (4mg-8mg) q.d	[27]	-0.67 (-0.87,-0.46)	-0.66 (-0.86,-0.47)	-0.65 (-0.85,-0.46)	-0.66 (-0.85,-0.46)
Oxybutynin gel 84mg/day	[134]	-0.64 (-1.18,-0.22)	-0.61 (-1.06,-0.25)	-0.62 (-0.98,-0.25)	-0.63 (-0.97,-0.26)
Oxybutynin transdermal 3.9mg/day	[10]	-0.6 (-0.9,-0.29)	-0.59 (-0.89,-0.32)	-0.61 (-0.85,-0.36)	-0.61 (-0.87,-0.35)

TABLE E.4: Sensitivity analyses assessing the impact in change from baseline in voiding episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis (cont.)

Oxybutynin vaginal ring 4mg q.d	[16]	-0.59 (-1.15,-0.15)	-0.58 (-1.04,-0.2)	-0.59 (-0.94,-0.24)	-0.58 (-0.93,-0.2)
Imidafenacin IR 0.1mg b.i.d	[36]	-0.56 (-0.83,-0.29)	-0.55 (-0.83,-0.28)	-0.56 (-0.81,-0.32)	-0.56 (-0.81,-0.32)
Terodiline IR 25mg b.i.d	[28]	-0.5 (-0.97,-0.03)	-0.51 (-1.03,-0.05)	-0.54 (-0.93,-0.13)	-0.53 (-0.95,-0.09)
Darifenacin ER 7.5mg q.d	[39]	-0.58 (-1.44,0.23)	-0.61 (-1.14,0.13)	-0.58 (-1.05,0.03)	-0.61 (-1.15,-0.1)
Oxybutynin gel 56mg/day	[135]	-0.52 (-0.92,0.07)	-0.52 (-0.87,0.03)	-0.53 (-0.89,-0.15)	-0.53 (-0.88,-0.13)
Oxbutynin patch 73.5mg	[15]	-0.6 (-19.79,-0.16)	-0.54 (-0.85,9.54)	-0.56 (-0.86,-0.2)	-0.56 (-0.88,-0.24)
Elocalcitol 75mg	[70]	-0.44 (-1.05,0.37)	-0.54 (-1.06,0.22)	-0.48 (-1.05,0.05)	-0.54 (-1.09,0)
Oxybutynin 20mg intravesically q.d	[106]	-0.55 (-1.34,0.44)	-0.51 (-1.12,0.26)	-0.55 (-0.96,-0.07)	-0.58 (-1.1,-0.07)
Imidafenacin IR 0.05mg b.i.d	[35]	-0.59 (-1.14,-0.1)	-0.54 (-1,-0.14)	-0.56 (-0.98,-0.1)	-0.58 (-1.01,-0.13)
Oxybutynin ER 15mg q.d	[9]	-0.49 (-0.96,0.61)	-0.49 (-0.91,0.28)	-0.52 (-0.9,-0.01)	-0.5 (-0.89,-0.02)
Oxybutynin IR (5-20mg)	[23]	-0.56 (-6.21,6)	-0.49 (-0.92,13.33)	-0.54 (-0.94,0.09)	-0.53 (-0.97,0.05)
Oxybutynin ER (5-30mg) q.d	[22]	-0.5 (-0.93,0.5)	-0.47 (-0.91,0.27)	-0.49 (-0.84,-0.02)	-0.51 (-0.84,-0.04)
Trospium chloride 45mg t.i.d	[47]	-0.53 (-1.06,0.07)	-0.49 (-0.99,0.02)	-0.51 (-1,0.04)	-0.51 (-1,0)
Cizolirtine citrate 200mg b.i.d	[56]	-0.44 (-1.41,0.81)	-0.58 (-1.44,0.86)	-0.45 (-1.49,0.71)	-0.45 (-1.39,0.63)
PFMT	[84]	-0.53 (-1.14,0.62)	-0.43 (-1.12,0.3)	-0.48 (-1.11,0.13)	-0.52 (-1.12,0.1)
Oxybutynin transdermal 1.3mg/day	[11]	-0.45 (-0.83,0.24)	-0.48 (-0.8,0.13)	-0.45 (-0.8,-0.01)	-0.48 (-0.81,-0.04)
Elocalcitol 150mg	[69]	-0.46 (-1.09,0.3)	-0.52 (-1.06,0.12)	-0.46 (-1.04,0.08)	-0.54 (-1.09,0.01)
Oxybutynin ER 2.5mg q.d	[20]	-0.54 (-1.73,0.11)	-0.52 (-1.7,-0.02)	-0.52 (-0.92,0.02)	-0.5 (-0.89,0.06)
Duloxetine IR 40mg b.i.d	[65]	-0.57 (-18.72,0.04)	-0.53 (-8.7,0.06)	-0.45 (-1.05,0.11)	-0.45 (-1.03,0.08)
Bladder Training	[85]	-0.5 (-0.96,0.17)	-0.45 (-0.97,0.09)	-0.45 (-0.86,-0.02)	-0.47 (-0.87,-0.02)
Solabegron 50mg b.i.d	[54]	-0.49 (-0.82,-0.19)	-0.5 (-0.75,-0.21)	-0.48 (-0.76,-0.2)	-0.48 (-0.75,-0.19)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.6 (-2.49,-0.21)	-0.57 (-2.28,-0.2)	-0.52 (-0.92,-0.15)	-0.54 (-0.92,-0.13)
Oxybutynin transdermal 2.6mg/day	[12]	-0.39 (-0.76,0.26)	-0.41 (-0.76,0.17)	-0.41 (-0.76,0.11)	-0.42 (-0.76,0.09)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.45 (-1.01,0)	-0.43 (-0.94,-0.01)	-0.45 (-0.91,-0.03)	-0.46 (-0.93,0.07)
Oxybutynin ER 2.5mg q.d + Bladder training	[92]	-0.48 (-2.07,0.44)	-0.49 (-1.73,0.65)	-0.5 (-1.57,0.34)	-0.37 (-1.29,0.61)
Oxybutynin IR 5mg b.i.d	[18]	-0.36 (-0.81,0.7)	-0.33 (-0.81,0.48)	-0.38 (-0.76,0.18)	-0.39 (-0.77,0.2)
Lipo-BoNTA	[138]	-0.5 (-2.43,0.49)	-0.44 (-2.07,0.6)	-0.39 (-1.25,0.38)	-0.42 (-1.26,0.52)
Serlopitant 0.25mg q.d	[107]	-0.35 (-0.81,0.12)	-0.41 (-0.82,0.06)	-0.31 (-0.8,0.1)	-0.37 (-0.87,0.11)
Serlopitant 4mg q.d	[109]	-0.36 (-0.83,0.09)	-0.38 (-0.83,0.1)	-0.33 (-0.84,0.09)	-0.36 (-0.87,0.1)
Tarafenacin 0.4mg q.d	[82]	-0.47 (-1.38,0.17)	-0.48 (-1.32,0.14)	-0.43 (-1.04,0.16)	-0.37 (-0.99,0.2)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.35 (-1.59,0.47)	-0.32 (-1.59,0.42)	-0.27 (-0.93,0.46)	-0.26 (-0.94,0.52)
Electrostimulation	[80]	-0.42 (-1.49,0.1)	-0.4 (-1.53,0.08)	-0.35 (-0.78,0.12)	-0.35 (-0.82,0.18)
Serlopitant 1mg q.d	[108]	-0.18 (-0.64,0.39)	-0.21 (-0.61,0.27)	-0.15 (-0.62,0.39)	-0.2 (-0.67,0.31)
Estradiol 25mg	[68]	-0.29 (-26.23,0.22)	-0.22 (-23.51,0.26)	-0.22 (-0.67,0.25)	-0.22 (-0.67,0.21)
Placebo	[1]	NA	NA	NA	NA
Netupitant 200mg q.d	[112]	-0.18 (-1.24,0.8)	-0.08 (-1.06,1.05)	-0.17 (-1.12,0.89)	-0.28 (-1.33,0.73)
Netupitant 100mg q.d	[111]	-0.09 (-1.14,0.91)	-0.03 (-1.08,0.91)	-0.1 (-1.08,0.82)	-0.27 (-1.31,0.79)
Tarafenacin 0.2mg q.d	[90]	-0.16 (-0.93,0.53)	-0.18 (-0.79,0.5)	-0.22 (-0.84,0.44)	-0.12 (-0.71,0.49)
Trospium IR 15mg t.i.d	[46]	-0.04 (-0.88,6.44)	-0.07 (-0.83,7.02)	-0.08 (-0.8,0.97)	-0.14 (-0.84,0.73)
ZD0947IL 25mg/day	[58]	0 (-0.88,1.65)	-0.08 (-1,1.09)	-0.08 (-0.88,0.8)	-0.04 (-0.86,0.93)
Netupitant 50mg q.d	[110]	-0.03 (-1.08,0.96)	-0.05 (-0.96,0.99)	-0.1 (-1.08,1.04)	-0.18 (-1.14,0.88)
Electromagnetic stimulation	[125]	-0.42 (-2.34,1.35)	-0.23 (-2.3,1.44)	-0.5 (-2.05,1.28)	-0.61 (-2.09,1.24)
Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	0.01 (-1.33,1.29)	0.15 (-1.08,1.33)	0.15 (-0.86,1.04)	0.05 (-0.87,1.11)
ONO-8539 100mg b.i.d	[60]	-0.03 (-0.76,0.6)	-0.05 (-0.69,0.59)	-0.09 (-0.79,0.61)	-0.07 (-0.71,0.52)
Percutaneous tibial nerve stimulation	[83]	-0.31 (-1.29,0.61)	-0.2 (-1.32,0.6)	-0.22 (-1.1,0.66)	-0.18 (-1.18,0.68)
Vaginal oestrogen cream 1.25mg/day	[132]	-0.07 (-1.1,0.54)	-0.09 (-1.13,0.46)	-0.01 (-0.53,0.5)	0 (-0.6,0.55)
Flavoxate chloride 200mg q.d	[64]	0.16 (-8.81,13.16)	0.18 (-0.65,59.98)	0.15 (-0.57,0.89)	0.14 (-0.61,0.85)
Resiniferatoxin 50nM	[67]	0.12 (-1.03,1.68)	0.2 (-1.52,1.14)	0.21 (-0.82,1.4)	0.14 (-0.86,1.45)
Emepronium bromide 200mg q.d	[63]	0.29 (-0.48,31.41)	0.21 (-0.6,42.29)	0.14 (-0.57,1.01)	0.14 (-0.6,0.87)
ONO-8539 300mg b.i.d	[61]	0.24 (-0.41,0.83)	0.24 (-0.42,0.74)	0.19 (-0.39,0.77)	0.2 (-0.37,0.78)
Propantheline Bromide 15mg t.i.d	[113]	0.28 (-0.71,1.28)	0.21 (-0.94,1.32)	0.41 (-0.74,1.7)	0.27 (-0.8,1.4)
Estradiol 1mg intravaginally	[127]	0.45 (-1.24,2.12)	0.28 (-0.8,1.78)	0.39 (-0.94,1.7)	0.3 (-1.1,1.68)
ONO-8539 30mg b.i.d	[59]	0.3 (-0.31,0.97)	0.24 (-0.3,0.86)	0.29 (-0.35,0.86)	0.28 (-0.32,0.88)
Control	[2]	0.43 (-0.37,2.26)	0.54 (-0.38,1.5)	0.43 (-0.36,1.13)	0.43 (-0.48,1.15)
Reflexology	[71]	0.24 (-1.01,1.49)	0.38 (-0.89,1.36)	0.37 (-0.51,1.17)	0.37 (-0.59,1.18)
Sham Therapy	[3]	0.46 (-0.57,1.53)	0.48 (-0.55,1.52)	0.49 (-0.51,1.45)	0.53 (-0.48,1.45)
Naftopidil 25mg q.d	[114]	3.36 (0.71,6.49)	3.59 (1.16,5.47)	3.33 (0.9,5.81)	3.64 (2,5.54)
Solifenacin succinate 5mg q.d + Naftopidil 25mg q.d	[115]	5.71 (3.19,8.53)	5.63 (3.29,8.72)	4.93 (2.67,7.3)	5.73 (3.91,8.03)

† Elements of $\boldsymbol{V^{1/2}}$ based on a Gamma(0.001,0.001) prior distribution on the precision scale

†† Elements of $\boldsymbol{V^{1/2}}$ based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

§ Deviance of treatment effect profiles across outcomes, τ, based on a Gamma(0.01,0.01) prior distribution on the precision scale

§§ Deviance of treatment effect profiles across outcomes, τ, based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

TABLE E.5: Sensitivity analyses assessing the impact in change from baseline in urgency episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis

Treatment	**Code**	**Median difference† (95%CrI)**	**Median difference†† (95%CrI)**	**Median difference§ (95%CrI)**	**Median difference§§ (95%CrI)**
Sacral nerve stimulation	[81]	-9.25 (-17.35,-4.97)	-8.35 (-10.22,3.64)	-9.05 (-11.04,-6.31)	-8.46 (-10.95,-6.07)
OnaBoNT-A 200u trigone sparing	[73]	-2.52 (-3.52,-1.14)	-2.5 (-3.53,-1.48)	-2.49 (-3.7,-1.85)	-2.55 (-3.5,-1.82)
Oxybutynin IR 2.5mg b.i.d + Salivary pastilles	[98]	-2.17 (-6.43,23.36)	-2.58 (-6.49,3.42)	-2.55 (-4.21,-0.79)	-2.15 (-3.99,-0.29)
Electrostimulation + PFE + Bladder training	[97]	-2.65 (-56.88,-1.63)	-2.39 (-15.44,-0.1)	-2.45 (-3.4,-1.25)	-2.39 (-3.34,-1.4)
Solifenacin/trospium + placebo injection	[100]	-2.19 (-5.89,49.25)	-2.22 (-15.6,1.93)	-2.22 (-3.28,-1.26)	-2.22 (-3.27,-1.35)
OnaBoNT-A 100u bladder base + trigone	[79]	-2.32 (-3.98,-1.05)	-2.22 (-3.67,-1.04)	-2.27 (-3.36,-1.34)	-2.3 (-3.41,-1.49)
OnaBoNT-A 100u bladder body + trigone	[78]	-2.23 (-3.9,-1.18)	-2.13 (-3.47,-1.36)	-2.16 (-2.96,-1.37)	-2.17 (-2.9,-1.34)
OnaBoNTA 100u trigone sparing	[72]	-2.14 (-2.74,-1.47)	-2.08 (-2.61,-1.61)	-2.12 (-2.6,-1.69)	-2.13 (-2.59,-1.71)
Tolerodine ER 4mg q.d + Neurostimulation	[96]	-1.92 (-3.4,-1.32)	-1.92 (-3.08,-1.42)	-1.91 (-2.4,-1.39)	-1.86 (-2.38,-1.36)
Estriol 1mg intravesival	[131]	-1.92 (-52.53,3.51)	-2.07 (-19.29,-0.75)	-1.66 (-2.96,-0.58)	-1.76 (-3.07,-0.48)
Trospium IR 15mg t.i.d + Physiotherapy	[91]	-1.82 (-20.29,-0.63)	-1.54 (-4.23,7.33)	-1.56 (-2.46,-0.58)	-1.61 (-2.52,-0.72)
Estradiol 3mg intravaginally	[128]	-1.76 (-30.65,-0.09)	-1.42 (-3.42,2.6)	-1.39 (-3.19,0.26)	-1.68 (-3.53,-0.3)
Solifenacin ER 10mg q.d	[30]	-1.37 (-1.84,-0.9)	-1.41 (-1.82,-0.98)	-1.38 (-1.73,-1.01)	-1.36 (-1.76,-1)
Tolterodine ER 2mg b.i.d + Oestrogen 0.625mg 2xwk	[99]	-1.25 (-6.62,10.38)	-1.31 (-2.01,15)	-1.23 (-1.84,-0.66)	-1.28 (-1.84,-0.62)
Tolterodine ER 4mg q.d + Behaviour therapy	[87]	-1.16 (-3.88,34.4)	-1.22 (-5.55,2.71)	-1.24 (-1.92,-0.59)	-1.23 (-1.94,-0.56)
Pregabalin 150mg b.i.d + Tolterodine ER 4mg q.d	[102]	-1.19 (-2.04,0.06)	-1.4 (-2.03,-0.26)	-1.25 (-2.09,-0.57)	-1.24 (-2.04,-0.57)
Fesoterodine ER 8mg q.d	[26]	-1.26 (-1.83,-0.89)	-1.25 (-1.7,-0.93)	-1.24 (-1.58,-0.88)	-1.22 (-1.56,-0.89)
Imidafenacin IR 0.25mg b.i.d	[37]	-1.29 (-2.41,-0.48)	-1.28 (-2.29,-0.6)	-1.28 (-1.88,-0.67)	-1.28 (-1.92,-0.65)
Solifenacin ER (5mg-10mg) q.d	[31]	-1.16 (-1.58,-0.49)	-1.18 (-1.58,-0.58)	-1.2 (-1.54,-0.81)	-1.17 (-1.53,-0.83)
Solifenacin ER 5mg - 15mg q.d	[34]	-1.19 (-19.35,15.31)	-1.21 (-8.54,1.97)	-1.19 (-1.67,-0.71)	-1.16 (-1.63,-0.69)
Mirabegron 100mg b.i.d	[48]	-1.17 (-2.14,-0.32)	-1.19 (-1.74,-0.58)	-1.17 (-1.62,-0.79)	-1.16 (-1.66,-0.8)
Solabegron 125mg b.i.d	[55]	-1.21 (-69.74,-0.46)	-1.2 (-24.06,-0.76)	-1.16 (-1.53,-0.77)	-1.16 (-1.54,-0.75)
Propiverine ER 30mg q.d	[42]	-1.18 (-41.8,-0.48)	-1.15 (-1.78,11.39)	-1.11 (-1.7,-0.67)	-1.1 (-1.73,-0.64)
Darifenacin ER 30mg q.d	[38]	-1.16 (-12.9,3.88)	-1.15 (-3.06,11.84)	-1.12 (-1.8,-0.53)	-1.18 (-1.84,-0.59)
Mirabegron 25mg q.d	[50]	-1.09 (-1.46,0.15)	-1.13 (-1.49,0.21)	-1.14 (-1.49,-0.78)	-1.13 (-1.49,-0.78)
Mirabegron 150mg b.i.d	[49]	-1.15 (-2.27,-0.24)	-1.18 (-1.82,-0.53)	-1.15 (-1.58,-0.75)	-1.14 (-1.6,-0.77)
Mirabegron 100mg q.d	[52]	-1.1 (-1.44,-0.37)	-1.12 (-1.45,-0.48)	-1.12 (-1.45,-0.78)	-1.12 (-1.45,-0.79)
Mirabegron 200mg q.d	[53]	-1.11 (-5.29,30.81)	-1.17 (-10.55,3.46)	-1.14 (-1.53,-0.72)	-1.13 (-1.53,-0.74)
Oxybutynin ER 10mg q.d	[8]	-1.18 (-27.39,2.21)	-1.13 (-2.09,4.24)	-1.14 (-1.67,-0.66)	-1.14 (-1.65,-0.67)
Propiverine 30mg b.i.d	[42]	-1.17 (-14.27,4.51)	-1.34 (-5.3,1.16)	-1.14 (-2.36,-0.53)	-1.18 (-2.55,-0.58)
Oxybutynin IR 3mg t.i.d	[19]	-1.04 (-1.56,32.3)	-1.08 (-1.61,6.44)	-1.1 (-1.58,-0.69)	-1.1 (-1.62,-0.69)
Mirabegron 50mg q.d	[51]	-1.08 (-1.41,-0.4)	-1.11 (-1.44,-0.38)	-1.12 (-1.44,-0.78)	-1.11 (-1.44,-0.79)
Solifenacin ER 5mg q.d	[29]	-1.13 (-1.61,-0.75)	-1.14 (-1.53,-0.79)	-1.11 (-1.46,-0.76)	-1.11 (-1.45,-0.77)
Tolterodine IR 2mg b.i.d + BT	[93]	-1.09 (-2.05,-0.06)	-1.09 (-2.02,-0.25)	-1.06 (-1.64,-0.49)	-1.08 (-1.65,-0.49)
Cizolirtine Citrate 400mg b.i.d	[57]	-1.16 (-4.01,-0.51)	-1.2 (-3.28,-0.6)	-1.04 (-1.71,-0.45)	-1.08 (-1.7,-0.46)
Trospium ER 60mg q.d	[44]	-1.01 (-1.5,17.75)	-1.05 (-1.54,14.01)	-1.08 (-1.52,-0.61)	-1.07 (-1.5,-0.63)
Propiverine 45mg t.i.d	[118]	-0.97 (-22.29,10.74)	-0.98 (-2.61,4.56)	-1.01 (-1.8,-0.28)	-1.01 (-2.01,-0.07)
Tolterodine IR 2mg b.i.d + Pilocarpine 9mg b.i.d	[101]	-1.06 (-78.09,4.47)	-1.04 (-13.41,6.74)	-1.05 (-1.45,-0.6)	-1.03 (-1.46,-0.61)
Fesoterodine ER 4mg q.d	[25]	-1.04 (-1.43,-0.67)	-1.05 (-1.43,-0.73)	-1.03 (-1.34,-0.7)	-1.02 (-1.33,-0.7)
Pregabalin 150mg b.i.d	[62]	-1.02 (-1.72,-0.04)	-1.07 (-1.68,-0.3)	-1.01 (-1.69,-0.44)	-1.02 (-1.66,-0.38)
Darifenacin ER 15mg q.d	[40]	-0.97 (-1.62,0.42)	-1 (-1.63,-0.06)	-1.03 (-1.55,-0.48)	-1.01 (-1.53,-0.49)
Propiverine IR 15mg b.i.d	[43]	-0.99 (-2.26,0.06)	-1.05 (-1.65,-0.33)	-1 (-1.45,-0.57)	-1 (-1.46,-0.54)
Tolterodine ER 4mg q.d	[4]	-0.98 (-1.28,-0.67)	-1 (-1.3,-0.72)	-0.98 (-1.27,-0.67)	-0.97 (-1.26,-0.69)
Oxybutynin IR 5mg t.i.d	[7]	-1.03 (-20.9,-0.61)	-1 (-1.46,6.64)	-0.98 (-1.38,-0.58)	-0.98 (-1.37,-0.59)
Propiverine ER 20mg q.d	[41]	-0.97 (-1.46,-0.52)	-1 (-1.38,-0.61)	-0.97 (-1.3,-0.62)	-0.97 (-1.31,-0.63)
Tolterodine IR 2mg b.i.d	[5]	-0.96 (-1.44,-0.38)	-0.98 (-1.36,-0.51)	-0.97 (-1.27,-0.63)	-0.96 (-1.27,-0.63)
Propiverine 60mg q.d	[119]	-0.86 (-16.98,10.28)	-0.98 (-2.65,5.84)	-0.91 (-1.6,0.42)	-0.9 (-1.53,0.25)
Oxybutynin intravesically 5mg t.i.d	[14]	-0.99 (-35.24,7.33)	-0.98 (-6.74,1.16)	-0.99 (-1.63,-0.45)	-0.96 (-1.62,-0.44)
Oxybutynin IR (2.5-5mg) b.i.d	[24]	-1 (-23.39,-0.49)	-0.96 (-29.88,-0.56)	-0.98 (-1.43,-0.44)	-0.94 (-1.38,-0.52)
Oxybutynin chloride topical gel 1g q.d	[13]	-0.92 (-6.17,34.22)	-0.94 (-1.46,14.52)	-0.94 (-1.35,-0.52)	-0.94 (-1.34,-0.54)
Oxybutynin vaginal ring 6mg q.d	[17]	-0.91 (-1.81,19.92)	-0.96 (-1.54,5.97)	-0.96 (-1.45,-0.52)	-0.94 (-1.4,-0.52)
Tolterodine IR 2mg b.i.d + PFE	[95]	-0.92 (-1.76,0.29)	-0.95 (-1.78,0.15)	-0.93 (-1.65,-0.09)	-0.96 (-1.64,-0.22)
PFMT + BT	[89]	-0.92 (-1.57,99.58)	-0.95 (-2.53,3.14)	-0.93 (-1.57,-0.29)	-0.92 (-1.55,-0.28)
Tolterodine IR 1mg b.i.d	[6]	-0.99 (-57.01,-0.54)	-0.95 (-12.77,-0.59)	-0.94 (-1.31,-0.44)	-0.93 (-1.29,-0.52)
Fesoterodine ER (4mg-8mg) q.d	[27]	-0.97 (-1.53,-0.55)	-0.97 (-1.4,-0.6)	-0.95 (-1.3,-0.6)	-0.95 (-1.29,-0.6)
Oxybutynin gel 84mg/day	[134]	-0.87 (-1.38,23.7)	-0.88 (-1.41,23.04)	-0.92 (-1.36,-0.44)	-0.91 (-1.34,-0.48)
Oxybutynin transdermal 3.9mg/day	[10]	-0.9 (-1.47,0.07)	-0.92 (-1.5,-0.32)	-0.92 (-1.27,-0.53)	-0.91 (-1.28,-0.54)

TABLE E.5: Sensitivity analyses assessing the impact in change from baseline in urgency episodes for different choices of prior distribution on variance parameters for multivariate hierarchical network meta-analysis (cont.)

Oxybutynin vaginal ring 4mg q.d	[16]	-0.92 (-22.22,4.6)	-0.93 (-10.01,1.22)	-0.91 (-1.34,-0.45)	-0.89 (-1.34,-0.43)
Imidafenacin IR 0.1mg b.i.d	[36]	-0.87 (-1.35,-0.22)	-0.89 (-1.36,-0.39)	-0.9 (-1.27,-0.49)	-0.89 (-1.26,-0.51)
Terodiline IR 25mg b.i.d	[28]	-0.84 (-5.95,13.09)	-0.85 (-1.65,17.65)	-0.88 (-1.4,-0.36)	-0.87 (-1.36,-0.35)
Darifenacin ER 7.5mg q.d	[39]	-0.95 (-2.73,-0.27)	-0.97 (-2.15,-0.35)	-0.89 (-1.45,-0.28)	-0.93 (-1.52,-0.38)
Oxybutynin gel 56mg/day	[135]	-0.91 (-69.77,5.35)	-0.92 (-13.1,-0.43)	-0.87 (-1.32,-0.39)	-0.85 (-1.3,-0.38)
Oxbutynin patch 73.5mg	[15]	-0.82 (-1.25,102.6)	-0.89 (-1.34,63.1)	-0.88 (-1.26,-0.44)	-0.87 (-1.27,-0.45)
Elocalcitol 75mg	[70]	-0.84 (-1.68,0.15)	-0.93 (-1.65,-0.1)	-0.83 (-1.45,-0.23)	-0.88 (-1.48,-0.27)
Oxybutynin 20mg intravesically q.d	[106]	-0.83 (-2.43,13.33)	-0.88 (-7.58,4.12)	-0.87 (-1.37,-0.34)	-0.88 (-1.49,-0.31)
Imidafenacin IR 0.05mg b.i.d	[35]	-0.86 (-1.59,0.25)	-0.84 (-1.45,-0.03)	-0.87 (-1.37,-0.36)	-0.88 (-1.38,-0.36)
Oxybutynin ER 15mg q.d	[9]	-0.79 (-1.47,55.46)	-0.84 (-2.25,6.25)	-0.84 (-1.3,-0.28)	-0.81 (-1.29,-0.27)
Oxybutynin IR (5-20mg)	[23]	-0.83 (-8.22,11.46)	-0.86 (-10.63,0.54)	-0.85 (-1.34,-0.2)	-0.83 (-1.32,-0.18)
Oxybutynin ER (5-30mg) q.d	[22]	-0.77 (-1.26,47.29)	-0.81 (-1.39,12.69)	-0.82 (-1.27,-0.29)	-0.83 (-1.25,-0.28)
Trospium chloride 45mg t.i.d	[47]	-0.76 (-1.41,31.11)	-0.78 (-1.38,23.06)	-0.83 (-1.4,-0.16)	-0.83 (-1.39,-0.22)
Cizolirtine citrate 200mg b.i.d	[56]	-0.81 (-1.87,0.55)	-0.97 (-1.83,0.64)	-0.75 (-1.89,0.38)	-0.75 (-1.76,0.31)
PFMT	[84]	-0.87 (-15.47,19.03)	-0.71 (-1.45,29.63)	-0.8 (-1.49,-0.1)	-0.81 (-1.49,-0.15)
Oxybutynin transdermal 1.3mg/day	[11]	-0.74 (-1.23,14.67)	-0.89 (-31.49,-0.32)	-0.77 (-1.21,-0.26)	-0.79 (-1.23,-0.28)
Elocalcitol 150mg	[69]	-0.77 (-1.59,0.51)	-0.85 (-1.5,0.12)	-0.78 (-1.4,-0.16)	-0.83 (-1.47,-0.23)
Oxybutynin ER 2.5mg q.d	[20]	-0.8 (-2.01,0.93)	-0.85 (-1.58,0.25)	-0.82 (-1.27,-0.22)	-0.78 (-1.26,-0.18)
Duloxetine IR 40mg b.i.d	[65]	-0.86 (-39.06,0.18)	-0.83 (-6,2.1)	-0.76 (-1.45,-0.15)	-0.77 (-1.39,-0.18)
Bladder Training	[85]	-0.81 (-1.57,0.32)	-0.77 (-1.52,0.08)	-0.77 (-1.26,-0.26)	-0.78 (-1.28,-0.26)
Solabegron 50mg b.i.d	[54]	-0.74 (-1.23,261.6)	-0.84 (-11.73,7.79)	-0.8 (-1.18,-0.4)	-0.79 (-1.18,-0.38)
Oxybutynin IR 2.5mg t.i.d	[21]	-0.8 (-2.09,-0.33)	-0.83 (-1.81,-0.39)	-0.75 (-1.15,-0.32)	-0.74 (-1.16,-0.34)
Oxybutynin transdermal 2.6mg/day	[12]	-0.79 (-15.02,0.44)	-0.76 (-2.27,4.03)	-0.72 (-1.17,-0.14)	-0.72 (-1.18,-0.17)
Pregabalin 75mg b.i.d + Tolterodine ER 2mg q.d	[103]	-0.71 (-1.5,0.3)	-0.74 (-1.33,0.05)	-0.76 (-1.33,-0.22)	-0.75 (-1.35,-0.16)
Oxybutynin ER 2.5mg q.d + Bladder training	[92]	-0.74 (-2.33,0.62)	-0.8 (-1.78,0.45)	-0.81 (-1.82,0.06)	-0.68 (-1.63,0.35)
Oxybutynin IR 5mg b.i.d	[18]	-0.73 (-40.86,4.7)	-0.7 (-12.26,2.47)	-0.7 (-1.17,-0.09)	-0.69 (-1.18,-0.07)
Lipo-BoNTA	[138]	-0.76 (-2.62,0.42)	-0.73 (-2.13,0.31)	-0.69 (-1.6,0.12)	-0.71 (-1.61,0.26)
Serlopitant 0.25mg q.d	[107]	-0.58 (-1.18,26.58)	-0.7 (-1.21,15.49)	-0.63 (-1.18,-0.11)	-0.68 (-1.22,-0.13)
Serlopitant 4mg q.d	[109]	-0.6 (-1.28,31.89)	-0.74 (-7.43,2.89)	-0.65 (-1.19,-0.11)	-0.67 (-1.27,-0.13)
Tarafenacin 0.4mg q.d	[82]	-0.71 (-1.71,0.5)	-0.82 (-1.61,-0.02)	-0.73 (-1.39,-0.12)	-0.68 (-1.32,-0.02)
Electrostimulation + vaginal oestrogen cream 1.25mg/day	[133]	-0.77 (-7.23,0.11)	-0.76 (-6.74,0.02)	-0.67 (-1.34,0.12)	-0.64 (-1.38,0.13)
Electrostimulation	[80]	-0.8 (-4.85,-0.18)	-0.76 (-4.48,-0.25)	-0.7 (-1.23,-0.18)	-0.7 (-1.27,-0.19)
Serlopitant 1mg q.d	[108]	-0.56 (-25.9,2.42)	-0.55 (-11.97,1.64)	-0.48 (-1.02,0.14)	-0.52 (-1.06,0.07)
Estradiol 25mg	[68]	-0.6 (-62.48,0.04)	-0.54 (-4.97,1.42)	-0.53 (-1.05,0.03)	-0.53 (-1.04,-0.03)
Placebo	[1]	NA	NA	NA	NA
Netupitant 200mg q.d	[112]	-0.41 (-1.44,0.89)	-0.4 (-1.34,0.81)	-0.47 (-1.41,0.62)	-0.58 (-1.68,0.43)
Netupitant 100mg q.d	[111]	-0.42 (-1.59,0.69)	-0.41 (-1.48,0.55)	-0.41 (-1.38,0.52)	-0.6 (-1.6,0.49)
Tarafenacin 0.2mg q.d	[90]	-0.47 (-1.38,0.64)	-0.54 (-1.31,0.6)	-0.54 (-1.22,0.14)	-0.44 (-1.07,0.23)
Trospium IR 15mg t.i.d	[46]	-0.39 (-1.26,32.98)	-0.46 (-1.91,6.04)	-0.4 (-1.21,0.69)	-0.47 (-1.24,0.42)
ZD0947IL 25mg/day	[58]	-0.4 (-1.47,1.22)	-0.46 (-1.49,0.93)	-0.41 (-1.24,0.48)	-0.37 (-1.22,0.65)
Netupitant 50mg q.d	[110]	-0.32 (-1.42,0.84)	-0.37 (-1.28,0.72)	-0.41 (-1.4,0.74)	-0.5 (-1.43,0.57)
Electromagnetic stimulation	[125]	-0.81 (-34.11,10.2)	-0.43 (-2.64,20.71)	-0.81 (-2.42,0.94)	-0.89 (-2.45,0.91)
Oxybutynin ER 5-30mg/day + Behaviour therapy	[22]	-0.37 (-39.69,1.33)	-0.16 (-5.93,1.7)	-0.17 (-1.2,0.8)	-0.25 (-1.2,0.84)
ONO-8539 100mg b.i.d	[60]	-0.31 (-1.32,0.87)	-0.39 (-1.11,0.52)	-0.4 (-1.15,0.41)	-0.38 (-1.06,0.3)
Percutaneous tibial nerve stimulation	[83]	-0.7 (-3.2,0.34)	-0.56 (-2.93,0.3)	-0.52 (-1.5,0.39)	-0.47 (-1.54,0.43)
Vaginal oestrogen cream 1.25mg/day	[132]	-0.35 (-3.68,0.36)	-0.39 (-3.36,0.26)	-0.3 (-0.84,0.34)	-0.25 (-0.9,0.34)
Flavoxate chloride 200mg q.d	[64]	-0.31 (-53.47,0.55)	-0.12 (-0.99,34.65)	-0.17 (-0.89,0.61)	-0.17 (-0.93,0.6)
Resiniferatoxin 50nM	[67]	-0.04 (-1.54,57.63)	-0.05 (-2.24,17.83)	-0.11 (-1.14,1.13)	-0.15 (-1.24,1.18)
Emepronium bromide 200mg q.d	[63]	-0.04 (-0.81,38.91)	-0.12 (-0.96,23.73)	-0.18 (-0.91,0.69)	-0.17 (-0.94,0.63)
ONO-8539 300mg b.i.d	[61]	-0.09 (-1.14,0.67)	-0.09 (-0.99,0.62)	-0.14 (-0.77,0.53)	-0.12 (-0.74,0.52)
Propantheline Bromide 15mg t.i.d	[113]	0.01 (-6.67,36.21)	-0.21 (-22.27,0.96)	0.09 (-1.06,1.42)	-0.04 (-1.12,1.13)
Estradiol 1mg intravaginally	[127]	-0.06 (-46.43,1.83)	-0.17 (-20.47,1.46)	0.06 (-1.26,1.43)	-0.01 (-1.42,1.36)
ONO-8539 30mg b.i.d	[59]	0.03 (-0.71,1.52)	-0.06 (-0.82,1.28)	-0.02 (-0.73,0.63)	-0.02 (-0.67,0.65)
Control	[2]	0.09 (-0.77,15.33)	0.12 (-6.33,1.18)	0.09 (-0.79,0.82)	0.1 (-0.9,0.86)
Reflexology	[71]	0.12 (-0.89,15.79)	0.1 (-6.05,1.45)	0.1 (-0.87,0.9)	0.09 (-0.87,0.92)
Sham Therapy	[3]	0.22 (-0.89,54.33)	0.2 (-0.9,11.57)	0.18 (-0.93,1.13)	0.21 (-0.87,1.17)
Naftopidil 25mg q.d	[114]	3.21 (-13.5,27.33)	3.36 (0.5,13.03)	2.96 (0.53,5.47)	3.33 (1.73,5.28)
Solifenacin succinate 5mg q.d + Naftopidil 25mg q.d	[115]	5.63 (-3.65,14.63)	5.46 (2.81,17.52)	4.63 (2.27,7.02)	5.41 (3.56,7.84)

† Elements of $\boldsymbol{V^{1/2}}$ based on a Gamma(0.001,0.001) prior distribution on the precision scale

†† Elements of $\boldsymbol{V^{1/2}}$ based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

§ Deviance of treatment effect profiles across outcomes, τ, based on a Gamma(0.01,0.01) prior distribution on the precision scale

§§ Deviance of treatment effect profiles across outcomes, τ, based on a Half-Normal(0,1)I(0,) prior distribution on the standard deviation scale

Appendix F

Published Papers

This appendix contains details of the manuscripts which resulted from the work carried out in this thesis and are published in peer reviewed journals:

Owen, R. K., Abrams, K. R., Mayne, C., Slack, M., and Tincello, D. G. Comparison of the effectiveness of repeated injections of onabotulinum toxin a for refractory idiopathic detrusor overactivity: analysis of an open label extension of a randomized trial (the relax study). *Neurourology and Urodynamics*, 2016.

Owen, R. K., Abrams, K. R., Mayne, C., Slack, M., and Tincello, D. G. Patient factors associated with onabotulinum toxin a treatment outcome in women with detrusor overactivity. *Neurourology and urodynamics*, 2016.

Owen, R. K., Tincello, D. G., and Keith, R. A. Network meta-analysis: development of a three-level hierarchical modeling approach incorporating dose-related constraints. *Value in Health*, 18(1):116–126, 2015.

F.1 Comparison of the effectiveness of repeated injections of onabotulinum toxin A for refractory idiopathic detrusor overactivity

In this paper, I contributed to study design, performed all of the statistical analyses, wrote the first draft of the paper and critically revised the final manuscript. My total contribution was approximately 60% of the research paper.

Received: 25 May 2016 | Accepted: 21 July 2016

DOI 10.1002/nau.23095

ORIGINAL CLINICAL ARTICLE

WILEY Neurourology and Urodynamics ICS SUFU

Comparison of the effectiveness of repeated injections of onabotulinum toxin A for refractory idiopathic detrusor overactivity: analysis of an open label extension of a randomized trial (the RELAX study)

Rhiannon K. Owen[1] | Keith R. Abrams[1] | Christopher Mayne[2] | Mark Slack[3] | Douglas G. Tincello[1,2]*

[1] Department of Health Sciences, University of Leicester, Leicester, UK

[2] University Hospitals of Leicester, Leicester, UK

[3] Addenbrookes' Hospital, Cambridge, UK

***Correspondence**
Douglas G. Tincello, BSc, MD, FRCOG, Professor of Urogynaecology, Department of Health Sciences, University of Leicester, Centre for Medicine, University Road, Leicester LE1 7RH, UK.
Email: dgt4@le.ac.uk

Funding Information
The study was funded by grants from The Moulton Charitable Trust and Wellbeing of Women. The application was funded after external peer review and study medication (active and placebo) were provided by Allergan free of charge. Otherwise, the study design, conduct and analysis was independent of the funders or Allergan. The study was sponsored by University Hospitals of Leicester NHS Trust. The sponsor, funders and Allergan remained independent of the design, conduct and analysis of the study. This paper presents independent research partly funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Rhiannon Owen is funded by a NIHR Doctoral Research Fellowship (NI-DRF-2013-06-062).

AIMS: To assess effects of repeat treatment with onabotulinumtoxin A (onaBoNT-A) in women with refractory idiopathic detrusor overactivity (DO).
METHODS: Analysis of an open-label extension study of a large randomized placebo controlled trial of onaBoNT-A. Participants had been randomized to receive 200 IU onaBoNTA or placebo and were offered up to two further onaBoNTA injections over a 5-year period. For this analysis, the primary outcome was duration of treatment effect by patient-reported symptom return. Weibull proportional hazards regression models were fitted in a Bayesian framework to estimate missing times. Multivariable hazard regression analysis (hazard ratio, 95% credible intervals (HR, 95% CrI) compared repeated injections adjusting for differences in baseline symptom severity. Secondary outcomes included inter-injection interval, incontinence, urgency, and voiding episodes 6 weeks after injection.
RESULTS: Four hundred and forty-two active injections were administered: 228 patients had one, 155 had two, and 59 had three injections. Time to symptom return for injection number 1 and 2 was 84 (95%CI: 63, 112) and 180 (95%CI: 135, 223) days, respectively. Median inter-injection intervals for receiving second and third injection were 266 days (range: 130, 1400) and 372 days (range: 134, 1283). No statistically significant differences in symptom outcomes or time to symptom return (HR 0.88, 95% CrI 0.37, 2.07 for injection 2, HR 0.33, 95% CrI 0.09, 1.03 for injection 3) were observed.
CONCLUSIONS: Repeated onaBoNT-A injections have consistent efficacy and duration of action. There appears to be long-term placebo effects in both groups of randomized patients, with implications for open-label extension studies.

KEYWORDS
botulinum toxin, detrusor overactivity, efficacy, overactive bladder, treatment

1 | INTRODUCTION

Botulinum toxin (BoNTA) is an established treatment for overactive bladder (OAB) and detrusor overactivity (DO) when conservative and medical treatments fail.[1,2] Several randomized studies[3–8] and numerous uncontrolled reports demonstrate high efficacy with long duration of action, although it is recognized that around 10% patients have problems with voiding dysfunction or urinary tract infections. Only onabotulinum toxin is currently licensed at a dose of 100 units for non-neurogenic OAB. Outcomes of repeated injections from uncontrolled series of both 100 and 200 units have been reported, suggesting repeat injections are

Dr. Fred Milani led the peer-review process as the Associate Editor responsible for the paper.

equally effective, although the number of patients included in these reports has been low.[9–14] For all these reports, repeated injections were compared using diary data, and quality of life measures. All demonstrated comparable effects on improvements in the objective outcomes and quality of life measures for each treatment, with no evidence of a decline in efficacy over time, and stable incidences of complications. The inter-injection interval remained stable at around 300 days between injections.

We present an analysis of an open label extension of a large randomized trial of 200 units of onabotulinum toxin A (onaBoNTA)[5] evaluating the efficacy of repeated injections on objective outcome measures and also compare the duration of treatment effect defined by time to repeat injection and time to patient-reported return of symptoms. We have used Bayesian methods to account for missing data, adjusting for baseline symptoms, and any potential selection effect due to the response to first treatment. Relative treatment efficacy of repeat injections is also presented.

2 | PATIENTS AND METHODS

The RELAX trial recruited 240 women with proven DO on urodynamics within 2 years of recruitment and refractory to standard treatment, with at least eight voids and at least two "moderate" or "severe" urgency episodes per 24 h.[5] The trial and extension study received ethical approval from the Scottish Multicentre Research Ethics Committee (ref: 04/MRE10/67) and was registered on Current Controlled Trials (ISRCTN26091555) on May 26, 2005. Women were randomized on a 1:1 basis to receive 200 units of onaBoNT-A or placebo, injected in 20 sites, sparing the trigone. Blinded outcome data were collected at baseline, 6 weeks, 3 months, and 6 months. Following completion of the blinded trial, participants entered a 5 year open label extension study after further informed, written consent and were offered a maximum of two further onaBoNT-A injections, administered as per local practice at each research site. When the protocol was designed, there were few published data on the use of botulinum toxin for idiopathic DO, so the dose of 200 units of onaBoNT-A (BOTOX®, Allergan, Dublin, Ireland) was used because this was the dose currently being offered by most investigators. The dose ranging study supported by Allergan was not published until 2010,[3] when all women were already enrolled in the extension study follow up, and the randomized trial data on 100 units were not published until 2013/14.[6–8] The provision of two injections was determined by the level of drug provision support provided by Allergan. The final treatment for the final patient occurred at the end of May 2013.

Outcome data (bladder diary, urgency episode frequency (Indevus Urgency Severity Scale (IUSS),[15] International Consultation on Incontinence Questionnaire short form (ICIQ-SF),[16] Incontinence Quality of Life (IQOL) questionnaire,[17] and Patient Global Impression of Improvement (PGI-I)[18,19]) were collected by post every 6 months from the date of the first (randomized) injection throughout the extension. Data on complications (self-reported voiding difficulty or urinary tract infection) were collected on the follow-up data form at each review. The patients were in regular contact with the local continence nurse specialists and thus we adopted this simplified reporting for the extension study.

With each follow up, injection requests could be initiated by patient request in response to the question "Do you wish to have a repeat injection at this time?" Patients could also request treatment at any time between follow-up contact. Treatment duration was based solely on self-reported return of symptoms in response to the question "have your symptoms returned?" without reference to original baseline symptom frequency or severity. Patients were sent a follow-up pack 6 weeks after every subsequent injection in addition to the scheduled 6 monthly review.

2.1 | Statistical analysis

For the comparison of repeated treatments, we grouped patients according to the sequence of *active* injection (termed injection 1, 2, and 3). We analyzed time to patient-reported recurrence of symptoms, in response to the questions "have your symptoms returned?" and "when did they return?" Data on time to patient reported return of symptoms were displayed using Kaplan–Meier curves and analyzed using Weibull proportional hazards regression models. We accounted for differences in baseline symptom severity at time of injection using PGI-S,[19] treatment at randomization, potential interactions between these factors and we further accounted for the similarity of repeated events within the same individuals.

Analysis of variance (ANOVA) adjusting for differences in baseline symptoms was used to assess the variability of mean diary data at 6 weeks post-injection for each of the active injections. Logistic regression was used to assess differences in the number of individuals with urinary tract infections (UTIs) and voiding difficulty. A result was considered significant if $P < 0.05$. Statistical analyses were performed using STATA, version 13.1 (StataCorp, College Station, TX).

3 | RESULTS

A total of 240 women were enrolled and treated; 122 women were initially randomized to onaBoNT-A and 118 to placebo.[5] A total of 442 active injections were administered during the randomized study and 5 year extension period: 228

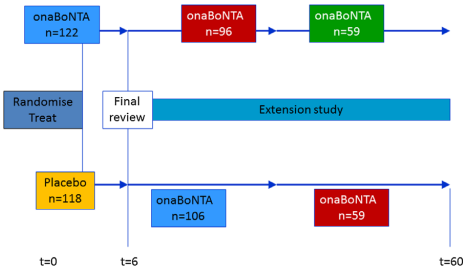


FIGURE 1 Study flowchart indicating numbers of women receiving one, two, and three injections. Injection cohorts for this analysis are shown in color: injection 1 (blue, $n = 228$), injection 2 (red, $n = 155$), and injection 3 (green, $n = 59$). Time is indicated in months by $t = n$

participants received first active, 155 received second active, and 59 received third active injections (Fig. 1).

A total of 189 (83%), 112 (72%), and 31 (53%) patients experienced symptom return after active injection 1, 2, and 3, respectively. Median time to symptom return (i.e., duration of treatment effect) for injection number 1 and 2 was 84 (95%CI: 63, 112) and 180 (95%CI: 135, 223) days from the time of injection, respectively. We were unable to calculate median time to symptom return for injection 3 as there were insufficient data on the number of events. For injection number 1 and 2, 47 (25%) and 25 (22%) patients failed to report the time of symptom return. Figure 2 displays the survival curve of patient-reported return of symptoms for each number of injections. Table 1 records the hazard ratios (HR) and credible intervals (CrI) for symptom return after each number of injections after accounting for differences in baseline severity. There was a reduced hazard of reporting symptom return for injection number 2 (HR: 0.88, 95% CrI: 0.37, 2.07) and 3 (HR: 0.33, 95% CrI: 0.09, 1.03) compared to injection 1; however, the 95% credible intervals span the point of no difference suggesting that in reality repeated injections are similar in duration of effect to the first.

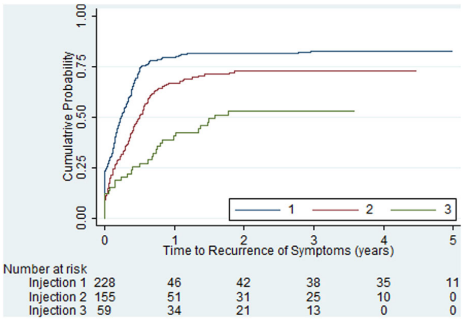


FIGURE 2 Survival curve (time to return of symptoms) for injection 1, 2, and 3. Injection cohorts are colored using the same scheme as in Figure 1

TABLE 1 Multivariable Weibull proportional hazards regression model results

		Hazard ratio	95% CrI
Active injection	1	Reference	—
	2	0.88	(0.37,2.07)
	3	0.33	(0.09,1.03)
PGI: severity	Normal	Reference	—
	Mild	1	(0.43,2.39)
	Moderate	1.21	(0.54,2.7)
	Severe	1.59	(0.68,3.92)
Randomized treatment	Placebo	Reference	—
	BoNTA	2.69	(1.66,4.68)
Active injection × PGI: severity	2 × Mild	1.02	(0.33,3.05)
	2 × Moderate	1.69	(0.62,4.92)
	2 × Severe	0.67	(0.22,1.93)
	3 × Mild	0.75	(0.17,3.04)
	3 × Moderate	0.83	(0.22,3.20)
	3 × Severe	0.61	(0.13,3.02)
Randomized treatment × active injection	BoNTA × 2	0.39	(0.21,0.71)

The final two rows show hazard ratio for the adjusted effect of injection number, baseline severity, and randomized allocation.

We found that patients randomized to onaBoNT-A in the double-blinded trial had a considerably more rapid rate of symptom return for injection 1 compared to patients initially randomized to placebo (HR: 2.69, 95% CrI: 1.66, 4.68) (Fig. 3a). However, for onaBoNT-A patients receiving their second active injection, the time to symptom return was significantly longer compared to first active injection (HR: 0.39, 95% CrI: 0.21, 0.71) which is further illustrated through the comparison of onaBoNT-A curves in Figure 3a and b.

The median inter-injection intervals for receiving second and third injection were 266 days (range: 130, 1400) and 372 days (range: 134, 1283), respectively. Tables 2 and 3 record the average symptom outcomes (Table 2) and adverse events (Table 3) at 6 weeks for all patients receiving each number of repeated onaBoNT-A injections. We observed a slight improvement in mean symptom severity with the second and third injection rounds but comparison of injections 1 and 2, and 1, 2, and 3 showed no statistically significant difference for any outcome variable (Table 2). Notably, women who opted for further injection had less severe symptoms at the preceding 6 week follow-up compared to the entire cohort, although these differences cannot be examined using statistical *t*-tests because these patients contribute to both statistics (i.e., the entire cohort of patients and the subgroup continuing treatment). The difference in symptom severity between those who did, and did not opt for further injection is particularly apparent for

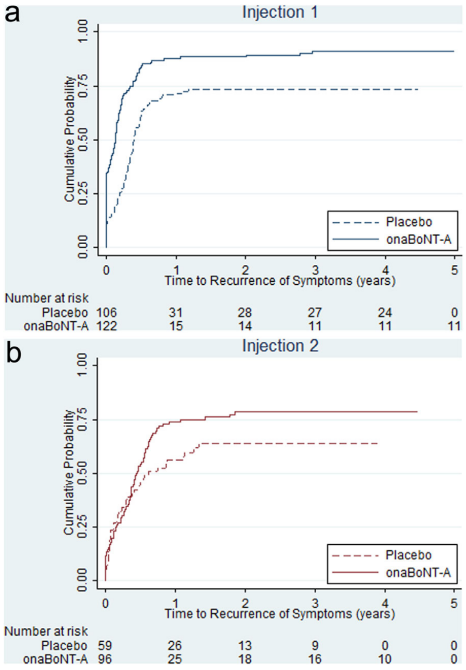


FIGURE 3 (a) Survival curve after injection 1 and (b) injection 2 grouped by randomized treatment during blinded phase

incontinence episodes. Patients ($n = 50$) receiving a third active injection experienced on average 1.71 incontinence episodes daily at 6 weeks following the first onaBoNT-A injection compared to 2.53 incontinence episodes daily for the entire cohort, suggesting that there was a clear selection effect for women continuing treatment.

There was no statistically significant difference between the number of individuals experiencing UTIs or voiding difficulty for each of the active injections (Table 3). Notably, individuals who opted for all three active injections did not experience UTIs or voiding difficulty with injection 1. Seventy-five percent of patients experienced UTIs, and 40% experienced voiding difficulty with injection 2 but continued to receive a third active injection, suggesting that in this case, occurrence of complications did not influence the decision for re-treatment. However, it should be noted that the overall number of patients who report UTI and voiding difficulty status, whether positive or negative, are particularly low with only 48 (21%) reporting UTI status, and 47 (21%) reporting voiding difficulty status out of a possible 228 individuals for injection 1. Reporting increased slightly for injection 2 and 3 with 30–40% of individuals reporting UTI and voiding difficulty status. Due to the number of missing data, these estimates should be interpreted with a degree of caution; a crude estimate of overall risk indicates about 20% of women report UTI after two or three injections, but with an estimate of overall voiding dysfunction of around 10%.

4 | DISCUSSION

Repeated injections of onaBoNT-A appear equally effective in patients with refractory DO. Median time to return of symptoms was 84 days after the first injection and 180 days after the second. Alongside this, we observed that the proportion of patients reporting symptom return was lower with second and third injections, and we observed a reduced hazard ratio of reporting return of symptoms with each injection, although there was uncertainty in these estimates so

TABLE 2 Average patient symptoms at 6 weeks following each repeated onaBoNT-A injection

				Mean episode frequency after each injection[a]			
	Injection[b]	Total no. patients	No. patients with data	1	2	3	*P**
Incontinence episodes/day	1	228	201	2.53 (3.76)	—	—	—
	2	155	133	2.35 (3.85)	1.91 (2.61)	—	0.14
	3	59	50	1.71 (3.44)	1.4 (2.60)	1.4 (2.68)	0.08
Urgency episodes/day	1	228	197	3.74 (4.15)	—	—	—
	2	155	136	3.28 (3.92)	2.77 (3.18)	—	0.66
	3	59	50	3.34 (4.02)	2.63 (2.84)	2.77 (2.47)	0.49
Voiding episodes/day	1	228	204	8.41 (3.39)	—	—	—
	2	155	138	8.24 (3.43)	7.43 (2.42)	—	0.27
	3	59	50	7.99 (2.91)	7.06 (2.02)	7.35 (1.99)	0.11

Results are presented as mean (SD).

[a]Each row includes all non-missing data for patients from the cohort receiving each number of injection.

[b]Refers to the sequence of *active* injection received, as outlined in Figure 1.

*Significance test is comparing across columns (i.e., comparing injection 3 with injection 2 and 1 in the same cohort of patients).

TABLE 3 Adverse events 6 weeks following each repeated onaBoNT-A injection

				Mean episode frequency after each injection[a]			
	Injection[b]	Total no. patients	No. patients with data	1	2	3	*P**
Urinary tract infection	1	228	48	33 (69%)	—	—	—
	2	155	61	11 (18%)	39 (64%)	—	0.48
	3	59	20	0 (0%)	15 (75%)	12 (60%)	0.87
Voiding difficulty	1	228	47	18 (38%)	—	—	—
	2	155	52	5 (10%)	15 (29%)	—	0.83
	3	59	18	0 (0%)	8 (40%)	3 (17%)	0.12

Results are presented as number of patients (%).
[a]Each row includes all non-missing data for patients from the cohort receiving each number of injection.
[b]Refers to the sequence of *active* injection received, as outlined in Figure 1.
*Significance test is comparing across columns (i.e., comparing injection 3 with injection 2 and 1 in the same cohort of patients).

this apparent difference should be interpreted carefully. These data suggest either that repeat injections have a slowly cumulative effect, or that there is a selection process whereby only those patients who observe benefit return for further treatment. Indeed, we found that participants who opt for re-injection appear to have better symptom profiles at the preceding 6 week follow up compared to the entire cohort. This would suggest that patients with less severe symptoms, or those who obtain a greater treatment benefit, choose further treatment. This finding has not, to our knowledge, been reported before and represents a potential selection effect for each subsequent injection.

Although the time between treatments in our cohort mirrors that of other papers (see below), we feel the time to reported symptom return is a more "real" measure of treatment effect. In any healthcare system, there will inevitably be a delay between the patient reporting to her carer the return of symptoms and the time of treatment. This will vary between countries and healthcare delivery systems. In our centers, there was a typical delay due to waiting time to have patients admitted or attending for treatment (during the study most procedures took place within the operating theater). Simply comparing time between treatments would overestimate the efficacy of treatment. Hence our presentation of patient-reported return of symptoms. The use of Bayesian methodology to analyze the time to patient-reported return of symptoms, accommodating potential confounding factors, provided robust data that repeat injections demonstrate consistent and similar efficacy in the duration of effect.

This result adds to findings in the current literature that suggests no difference in clinical efficacy of repeated injections based on symptom profiles. The average inter-injection interval from our data was approximately 9 months between first and second treatment, and over a year between second and third treatments, intervals which are comparable with the work of several other authors. Sahai et al.[12] reported a mean inter-injection period of 377, 378, and 256 days between injection numbers 1 and 2, 2 and 3, and 3 and 4, respectively, among 20 patients having repeat treatment of 200 units onaBoNTA for idiopathic DO. They found urinary frequency, urgency, incontinence, and quality of life assessments showed equivalent improvements after each injection, with pre-injection symptoms being similar to those at baseline. Dowson et al.[13] extended that cohort, and reported a mean inter-injection period of 322 days in 53 patients receiving a second injection of 200 units of onaBoNT-A for refractory overactive bladder with outcomes after each injection (up to the fifth analysed) that showed no difference from each other. Granese et al.[10] and Gousse et al.[11] demonstrated equivalent efficacy of 100 or 150 units in repeat injection received by 20 or 31 patients, respectively. Our data are from a larger cohort of patients (all women) where 155 women had a second injection and 59 received a third, and thus, provide more robust confirmation of these earlier published papers. After accounting for differences in baseline severity, we have also found that there was little evidence to suggest repeat injections differ in terms of patients' urinary diary data.

Using "time to return of symptoms" is not without limitations; this was an absolute response and did not allow assessment of the complexities regarding how patients process the return of a symptom, which does not yet become bothersome enough to seek repeat treatment. This threshold is likely to vary between individuals, and also between each symptom, where perhaps return of urgency is more immediately bothersome than a return of greater frequency. Thus, we have shown a shorter apparent duration of efficacy from onaBoNT-A than other workers, but a similar inter-injection interval. There is clearly more work to be done to explore this complex relationship, and to examine how different definitions of cure, efficacy, and thresholds for re-injection will impact the long-term cost-effectiveness of this treatment. There are still no long-term, robust cost-effectiveness data available in comparison to alternative treatments.

A limitation of this study was sporadic patient follow up following the end of the blinded trial. Complete data were obtained for 6 week follow up after every subsequent injection but thereafter, and over the 5 year extension study, it became increasingly difficult to monitor patients' symptom return. It was unclear whether some patients were lost to follow up or simply, their symptoms had not returned. For patients with symptom return, but missing data on time of symptom return, typically the average reported time from known events is assumed. However, this may exaggerate repeated treatment effects when patients continuing follow up are different (e.g., by having better baseline symptom profiles and/or longer duration of treatment effect) compared to patients discontinuing follow up. To help ameliorate a potential reporting bias in this situation, we used a Bayesian approach. This has several advantages, including the ability to obtain predictive estimates for missing data and associated uncertainty.[20] Given that we know the interval in which symptoms returned for each individual with missing data (i.e., it is assumed to fall between last complete follow up and date of repeat injection), we predicted estimates within this time interval for each missing observation. Models were fitted using WinBUGs 1.4.3.[20,21]

We noted an interesting finding that patient-reported duration of effect appeared to be influenced by initial treatment randomization. Patients randomized to placebo had a considerably decreased rate of symptom return for their first onaBoNT-A injection (received in the open label extension) compared to patients initially randomized to onaBoNT-A (who received their first active drug in a blinded fashion). This may represent an extended placebo effect, which has not been noted before in trials of interventions for DO or overactive bladder. It is known from migraine research that the placebo effect in randomized studies of onaBoNT-A treatment can remain at a steady rate for up to 6 months[22,23] and that the placebo effect is greater for more invasive treatments[24] but we are unaware of any literature demonstrating that patients who receive placebo initially, subsequently report greater efficacy of active treatment when they receive it. We also noted that patients who received active drug initially, subsequently reported a greater duration of effect with the second injection during the extension phase. Thus, it would seem that both groups (those randomized to both active and placebo) reported greater efficacy for the second injection, received during the extension study. This over-reporting by both groups effectively means that open label extension studies following randomization may be biased toward more positive outcomes compared to the true (randomized and blinded) effects. This observation has wide implications if this effect can be confirmed, because nearly all drug studies for medication for OAB and DO have a pooled open-label extension included to generate additional data in support of the licensing and use of the product.

5 | CONCLUSIONS

The data we present here represent the largest cohort of patients receiving two and three injections of onaBoNT-A, and have been analyzed using novel and robust statistical methods to account for real and potential biases of selection among patients choosing to continue with repeated treatment. As far as we are aware, we are the only authors to analyze the duration of treatment effect accounting for these variables when comparing efficacy of repeat treatments. Based on our data, there appears to be no loss of effect after second and third injections of onaBoNT-A, either in terms of the expected duration of action, or the magnitude of relevant urinary diary outcomes.

POTENTIAL CONFLICTS OF INTEREST

Dr. Owen reports grants from National Institute for Health Research (NIHR), during the conduct of the study. Dr. Abrams reports grants from National Institute for Health Research (NIHR), from null, during the conduct of the study; other from Allergan, other from Astellas, outside the submitted work. Dr. Tincello reports grants from Wellbeing of Women, grants from Moulton Charitable Trust, during the conduct of the study; other from Allergan honorarium for teaching, other from Allergan advisory board, outside the submitted work. Dr. Mayne reports grants from Wellbeing of Women—Charitable Trust, during the conduct of the study. Dr. Slack reports grants from Wellbeing of Women, from null, during the conduct of the study.

REFERENCES

1. Smith A, Bevan D, Douglas HR, et al. Management of urinary incontinence in women: summary of updated NICE guidance. *Brit Med J*. 2013;347:f5170.
2. Gormley EA, Lightner DJ, Burgio KL, et al. Diagnosis and treatment of overactive bladder (non-neurogenic) in adults: AUA/SUFU guideline. *J Urol*. 2012;188:2455–2463.
3. Dmochowski R, Chapple C, Nitti VW, et al. Efficacy and safety of onabotulinumtoxinA for idiopathic overactive bladder: a double-blind, placebo controlled, randomized, dose ranging trial. *J Urol*. 2010;184:2416–2422.
4. Brubaker L, Richter HE, Visco A, et al. Refractory idiopathic urge urinary incontinence and botulinum A injection. *J Urol*. 2008;180:217–222.
5. Tincello DG, Kenyon S, Abrams KR, et al. Botulinum toxin A versus placebo for refractory detrusor overactivity in women: a randomised blinded placebo-controlled trial of 240 women (the RELAX study). *Eur Urol*. 2012;62:507–514.
6. Nitti VW, Dmochowski R, Herschorn S, et al. OnabotulinumtoxinA for the treatment of patients with overactive bladder and urinary incontinence: results of a phase 3, randomized, placebo controlled trial. *J Urol*. 2013;189:2186–2193.
7. Chapple C, Sievert KD, MacDiarmid S, et al. OnabotulinumtoxinA 100U significantly improves all idiopathic overactive bladder

symptoms and quality of life in patients with overactive bladder and urinary incontinence: a randomised, double-blind, placebo-controlled trial. *Eur Urol*. 2013;64:249–256.

8. Denys P, Le NL, Ghout I, et al. Efficacy and safety of low doses of onabotulinumtoxinA for the treatment of refractory idiopathic overactive bladder: a multicentre, double-blind, randomised, placebo-controlled dose-ranging study. *Eur Urol*. 2012;61: 520–529.
9. Khan S, Kessler TM, Apostolidis A, et al. What a patient with refractory idiopathic detrusor overactivity should know about botulinum neurotoxin type A injection. *J Urol*. 2009;181: 1773–1778.
10. Granese R, Adile G, Gugliotta G, et al. Botox((R)) for idiopathic overactive bladder: efficacy, duration and safety. Effectiveness of subsequent injection. *Arch Gynecol Obstet*. 2012;286:923–929.
11. Gousse AE, Kanagarajah P, Ayyathurai R, et al. Repeat intradetrusor injections of onabotulinum toxin a for refractory idiopathic overactive bladder patients: a single-center experience. *Female Pelvic Med Reconstr Surg*. 2011;17:253–257.
12. Sahai A, Dowson C, Khan MS, et al. Repeated injections of botulinum toxin-A for idiopathic detrusor overactivity. *Urology*. 2010;75:552–558.
13. Dowson C, Watkins J, Khan MS, et al. Repeated botulinum toxin type A injections for refractory overactive bladder: medium-term outcomes, safety profile, and discontinuation rates. *Eur Urol*. 2012;61:834–839.
14. Game X, Khan S, Panicker JN, et al. Comparison of the impact on health-related quality of life of repeated detrusor injections of botulinum toxin in patients with idiopathic or neurogenic detrusor overactivity. *BJU Int*. 2011;107:1786–1792.
15. Nixon A, Colman S, Sabounjian L, et al. A validated patient reported measure of urinary urgency severity in overactive bladder for use in clinical trials. *J Urol*. 2005;174:604–607.
16. Avery K, Donovan J, Peters TJ, et al. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourol Urodyn*. 2004;23:322–330.
17. Patrick DL, Martin ML, Bushnell DM, et al. Quality of life of women with urinary incontinence: further development of the Incontinence Quality of Life Instrument (I-QOL). *Urology*. 1999;53:71–76.
18. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. *Am J Obstet Gynecol*. 2003;189:98–101.
19. Tincello DG, Owen RK, Slack MC, et al. Validation of the Patient Global Impression scales for use in detrusor overactivity: secondary analysis of the RELAX study. *BJOG*. 2013;120:212–216.
20. Spiegelhalter D, Abrams P, Myles J. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Chichester UK: John Wiley & Sons Ltd; 2003.
21. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo Methods in Practice*. London, UK: Chapman & Hall; 1996.
22. Mathew NT, Frishberg BM, Gawel M, et al. Botulinum toxin type A (BOTOX) for the prophylactic treatment of chronic daily headache: a randomized, double-blind, placebo-controlled trial. *Headache*. 2005;45:293–307.
23. Silberstein SD, Gobel H, Jensen R, et al. Botulinum toxin type A in the prophylactic treatment of chronic tension-type headache: a multicentre, double-blind, randomized, placebo-controlled, parallel-group study. *Cephalalgia*. 2006;26:790–800.
24. Diener HC. Placebo effects in treating migraine and other headaches. *Curr Opin Investig Drugs*. 2010;11:735–739.

F.2 Patient factors associated with onabotulinum toxin A treatment outcome in women with detrusor overactivity

In this paper, I performed all of the statistical analyses, contributed to the first draft of the paper and critically revised the final manuscript. My total contribution was approximately 40% of the research paper.

Neurourology and Urodynamics

Patient Factors Associated With Onabotulinum Toxin A Treatment Outcome in Women With Detrusor Overactivity

Rhiannon K. Owen,[1] Keith R. Abrams,[1] Christopher Mayne,[2] Mark Slack,[3] and Douglas G. Tincello[1,2]*
[1]*University of Leicester, Department of Health Sciences, Leicester, United Kingdom*
[2]*Women's and Children's CBU, University Hospitals of Leicester, Leicester, United Kingdom*
[3]*Gynaecology, Rosie Hospital, Cambridge University Hospitals NHS Trust, Cambridge, United Kingdom*

Objective: To evaluate potential predictors of non-response to treatment with 200U onabotulinum toxin A (onaBoNTA) in women with refractory detrusor overactivity (DO). **Subjects and Methods:** A secondary analysis of a randomized trial of 200U onaBoNTA versus placebo in women with refractory DO analyzed baseline and 6 week follow-up data. Univariate and multivariate logistic regression were used to assess demographic factors and baseline clinical parameters on non-response to treatment defined as 20% or less improvement in urinary urgency and leakage episodes, 10% or less in voiding frequency, not achieving continence, and "no change" or worse on PGI-I score at 6 weeks. **Results:** One Hundred and twenty-two women were included. Twenty-nine (23.8%), 24 (19.7%), and 19 (15.6%) were non-responders to treatment for urgency, voiding, and leakage episodes, respectively. Fifty-nine (48.4%) failed to achieve continence, and 28 (23%) were non-responders on the PGI-I scale. Smoking status (OR: 2.89 95%CI 1.08, 7.73, $P = 0.034$) predicted non-response in urgency episodes, and higher baseline leakage episodes (OR: 1.17 95%CI 1.04, 1.31, $P = 0.007$) predicted non-response in achieving continence. Increasing age (OR 1.04, 95%CI 1.0, 1.09, $P = 0.063$) and body mass index (BMI) (OR 1.07, 95%CI 1.0, 1.16, $P = 0.065$) showed marginal associations with non-response on the PGI-I scale. **Conclusion:** onaBoNTA is an effective treatment for refractory DO, but some fail to respond. For identification of women at risk, our data indicate smokers should be advised of a lesser chance of successful treatment. Older women, those with high BMI and with more severe leakage also have a higher risk of failure. *Neurourol. Urodynam.* © 2016 Wiley Periodicals, Inc.

Key words: detrusor overactivity; efficacy; onabotulinum toxin; overactive bladder; treatment

INTRODUCTION

Detrusor overactivity (DO) is characterized by spontaneous contractions of the detrusor muscle during bladder filling, causing symptoms of urgency, frequency, nocturia, and incontinence, which together are symptoms of the overactive bladder syndrome (OAB).[1] Initial treatments include behavioral therapy[2] and antimuscarinic drugs which have moderate efficacy but troublesome side effects leading to frequent discontinuations.[3–6] In the last 5 years, botulinum toxin A (BoNTA) has been rapidly adopted as a treatment for DO and OAB, based upon data from several case series and more recently several randomized trials which enrolled women with both urodynamically proven DO and those with only OAB symptoms.[7–13] The data from these trials consistently show profound and long lasting effects on all symptoms of OAB, typically with reductions in excess of 50% of baseline and duration of between 3 and 9 months. However, not all patients respond to treatment with BoNTA, but very little data have been published exploring the reasons for this. Here we present a secondary analysis of data from a large single dose RCT comparing a dose of 200 units of onabotulinum toxin A (onaBoNTA) (BOTOX®, Allergan, Dublin, Ireland) with placebo in women[11] to examine if any patient factors can predict the likelihood of treatment failure at 6 weeks after treatment. At the time the trial was conducted, 200 units of onaBoNTA was the accepted dose for patients with idiopathic DO, with the licences for 100 units only being granted after 2010.

MATERIALS AND METHODS

The RELAX trial[11] was registered on Current Controlled Trials (ISRCTN26091555) on 26th May 2005 and recruited between July 2006 and November 2009 from eight UK hospitals. The trial received ethical approval from the Scottish Multicenter Research Ethics Committee (ref: 04/MRE10/67). Briefly, women with proven DO on urodynamics[1] within 2 years of recruitment, at least eight voids and at least two "moderate" or "severe" urgency episodes per 24 hr,[14] refractory to standard treatment were randomized 1:1 to receive 200 units of onaBoNTA (BOTOX®, Allergan) or placebo injected in 20 sites, sparing the trigone. This was the accepted dosage recommended for idiopathic DO worldwide during the lifetime of the trial. Blinded outcome data were collected at baseline, 6 weeks, 3 months (by post), and 6 months. Urinary voiding frequency, leakage episode frequency, urgency episode frequency (moderate or severe on Indevus Urgency Severity Scale [IUSS])[14] were recorded in 3-day voiding diaries; the International Consultation on Incontinence Questionnaire short form (ICIQ-SF),[15] Incontinence Quality of Life (IQOL) questionnaire,[16] and Patient Global Impression of Improvement (PGI-I)[17,18] scale were completed at each time point. Details of the trial design and

Dr. Hashim Hashim led the peer-review process as the associate editor responsible for the paper.
Potential conflicts of interest: DGT, CJM, and MCS received a grant from Wellbeing of Women during the conduct of the study. DGT reports personal fees for consultancy with Allergan for teaching and for advisory board participation with Allergan outside the submitted work. KRA reports personal fees for consultancy work for Allergan and Astellas outside the submitted work. RKO reports consultancy payments from Allergan outside the submitted work.
*Correspondence to: Douglas G. Tincello, B.Sc., M.D., F.R.C.O.G., Department of Health Sciences, University of Leicester, 22–28 Princess Road West, Leicester LE1 6TP, United Kingdom. E-mail: dgt4@le.ac.uk
Received 22 September 2015; Accepted 30 November 2015
Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/nau.22948

primary study outcome data have been reported elsewhere.[11] The analysis presented here focused on baseline and 6-week follow up data from those women who received active drug. After 6 weeks follow up, the original trial protocol allowed women with little response to resume antimuscarinic medication, so this time point was the only opportunity to analyze the effects of onaBoNTA alone.

Statistical Analysis

There is no agreed definition of "non-response" to any treatment for DO or OAB. To search for factors influencing the effect of onaBoNT-A in women receiving active drug we have defined non-response for each of the outcome measures used in the trial: 20% or less improvement in urgency episode frequency; 10% or less improvement in voiding episode frequency; 20% or less improvement in leakage episode frequency; those women not achieving continence, and those women reporting a response of "no change" or worse on the PGI-I scale.[17] These definitions represented a reduction of approximately one episode per day for the median of each variable at baseline: urgency (1.6), leakage (1.2), and voiding frequency (1.0) episodes.[11] We explored both primary and all secondary outcomes from the blinded trial in this analysis, due to the lack of a standardized definition of non-response. A sensitivity analysis was included to examine different cut-off thresholds (at 10% and 50%) in view of the absence of an agreed definition. All outcomes were studied at 6 weeks (first follow-up) to avoid confounding of the use of additional treatments (usually antimuscarinic drugs) which were allowed during the remainder of follow-up between 6 weeks and 6 months, and because the greatest treatment effect was seen at this time.[11]

Relevant demographic factors and baseline clinical parameters were analyzed in a complete case analysis using univariate logistic regression. Factors found to be significant at the 10% level in the univariate analysis were entered into stepwise forward multiple regression. Variables with *P*-values <0.05 were considered significant. The following potential factors were examined: age at treatment; body mass index (BMI); ethnicity; parity; smoking status; previous continence surgery; baseline voiding frequency; baseline leakage episodes; baseline urgency episodes; baseline Urgency Severity Score; baseline maximum voided volume from the urinary diary; baseline mean voided volume from the urinary diary; baseline ICIQ score; baseline IQOL score; and baseline urodynamic data (including volume at first desire, volume at capacity, maximum detrusor pressure, maximum voiding detrusor pressure, maximum free flow rate, detrusor pressure at maximum flow). We also included the occurrence of urinary retention and urinary tract infection identified at the 6 week visit as additional, post-treatment factors which might influence perceived efficacy. All statistical analyses were performed using STATA, version 13.0 (StataCorp, College Station, TX).

RESULTS

Four hundred fifteen women were screened and 240 were enrolled and treated in the randomized study.[11] One Hundred and twenty-two women were randomized to BoNTA and 118 women to placebo. All outcome data were comparable at baseline (data not shown).[11] Of the 122 women receiving active treatment, 29 (23.8%) had a 20% or less change in urgency episodes at 6 weeks, 24 (19.7%) had a 10% or less change in voiding episodes at 6 weeks, 19 (15.6%) had a 20% or less change in leakage episodes at 6 weeks, 59 (48.4%) failed to achieve continence, and 28 (23%) reported a response of "no change" or worse on the PGI-I scale. Baseline characteristics studied for these women are presented in Tables I–IV.

Univariate analysis identified smoking status as a potential predictor of non-response in urgency episodes, with smokers having nearly three times increased odds of non-response (OR:2.89 95%CI 1.08, 7.73, $P=0.034$) compared to non-smokers.

TABLE I. Baseline Clinical Characteristics of the Patients Receiving Botulinum Toxin by Change in Urinary Diary Symptoms

	Change in urgency episodes <= 20%			Change in voiding frequency <= 10%			Change in leakage episodes <= 20%		
	Non-response	Response	Missing	Non-response	Response	Missing	Non-response	Response	Missing
Characteristics									
n (122)	29 (23.8%)	78 (63.9%)	15 (12.3%)	24 (19.7%)	86 (70.5%)	12 (9.8%)	19 (15.6%)	91 (74.6%)	12 (9.8%)
Age	59.8 (9.7)	59.3 (11.8)	60.2 (14.7)	63.2 (10.8)	58.8 (11.5)	57.3 (14.2)	62.2 (11.1)	59.3 (11.5)	57.3 (14.1)
Ethnicity									
White (n = 117)	27 (93.1 %)	75 (96.2%)	15 (100%)	24 (100%)	81 (94.2%)	12 (100%)	18 (94.7%)	87 (95.6%)	12 (100%)
Other (n = 5)	2 (6.9%)	3 (3.6%)	0 (0%)	0 (0%)	5 (5.8%)	0 (0%)	1 (5.3%)	4 (4.4%)	0 (0%)
Parity									
0 (n = 10)	2 (6.9%)	6 (7.7%)	2 (13.3%)	3 (12.5%)	5 (5.8%)	2 (16.7%)	1 (5.3%)	7 (7.7%)	2 (16.7%)
1+ (n = 122)	27 (93.1%)	72 (92.3%)	13 (86.7%)	21 (87.5%)	81 (94.2%)	10 (83.3%)	18 (94.7%)	84 (92.3%)	10 (83.3%)
Previous surgery									
Yes	12 (41.4%)	28 (35.9%)	4 (26.7%)	11 (45.8%)	30 (34.9%)	3 (25%)	9 (47.4%)	32 (35.2%)	3 (25%)
No	17 (58.6%)	50 (64.1%)	11 (73.3%)	13 (54.2%)	56 (65.1%)	9 (75%)	10 (52.6%)	59 (64.8%)	9 (75%)
BMI	30.4 (6.5)	28.5 (6.0)	29.0 (7.3)	30.3 (7.1)	28.6 (5.8)	29.5 (7.8)	31 (6.5)	28.5 (6.0)	29.5 (7.8)
Baseline leakage	6.5 (5.3)	6.4 (4.0)	6.8 (4.2)	6.0 (3.8)	6.5 (4.4)	7.4 (4.3)	7.5 (4.4)	6.1 (4.3)	7.4 (4.3)
Baseline urgency	7.4 (3.5)	8.2 (3.1)	8.6 (3.4)	7.1 (3.3)	8.2 (3.2)	9.2 (3.2)	7.8 (3.1)	8 (3.3)	9.2 (3.2)
Baseline voiding	13.9 (10.2)	11.7 (3.3)	11.0 (2. 7)	10.6 (2.0)	12.7 (6.6)	11.4 (2.9)	11.9 (3.3)	12.3 (6.4)	11.3 (2.9)
Smoking status									
Smoker	10 (34.5%)	12 (15.4%)	8 (53.3%)	5 (20.8%)	18 (20.9%)	7 (58.3%)	6 (31.6%)	17 (18.7%)	7 (58.3%)
Non-smoker	19 (65.5%)	66 (84.6%)	7 (46.7%)	19 (79.2%)	68 (79.1%)	5 (41.7%)	13 (68.4%)	74 (81.3%)	5 (41.7%)
Max voided vol	375.1 (184.9)	387.9 (134.9)	316.5 (136.5)	376.5 (152.7)	385.3 (147.8)	309.2 (146.2)	384.2 (137.7)	383.2 (151.1)	309.2 (146.0)
Av. voided vol	166.5 (76.8)	171.6 (61.3)	169.2 (87.8)	173.0 (47.6)	168.7 (47.6)	173.4 (97.3)	170 (50.3)	169.5 (67.5)	173.4 (97.3)
IUSS score	2.0 (0.6)	2.0 (0.5)	2.1 (0.5)	2.0 (0.6)	2.0 (0.5)	2.2 (0.5)	2.1 (0.5)	2.0 (0.5)	2.2 (0.5)
ICIQ score	15.3 (4.5)	14.1 (9.9)	16.4 (4.6)	15.0 (4.8)	14.4 (9.5)	16.3 (5.1)	16.7 (2.7)	14 (9.4)	16.3 (5.1)
IQOL score	24 (18.9)	21.3 (46.8)	24.9 (19.1)	33.3 (20.7)	18.4 (44.1)	28.8 (19.4)	29.8 (19.3)	19.6 (43.6)	28.8 (19.3)

TABLE II. Baseline Clinical Characteristics of the Patients Receiving Botulinum Toxin by Continence Status and Patient Global Impression of Improvement

	Incontinent at follow up			Patient Global Impression of Improvement		
	Non-response	Response	Missing	Non-response	Response	Missing
Characteristics						
n (122)	59 (48.4%)	45 (36.9%)	18 (14.8%)	28 (23.0%)	80 (65.6%)	14 (11.5%)
Age	61.3 (11.0)	58.4 (11.8)	56.6 (13.2)	63.7 (11.1)	59.0 (11.0)	54.5 (14.4)
Ethnicity						
White (n = 117)	55 (93.2%)	44 (97.8%)	18 (100%)	26 (92.9%)	78 (97.5%)	13 (92.9%)
Other (n = 5)	4 (6.8%)	1 (2.2%)	0 (0%)	2 (7.1%)	2 (2.5%)	1 (7.1%)
Parity						
0 (n = 10)	2 (3.4%)	5 (11.1%)	3 (16.7%)	1 (3.6%)	6 (7.5%)	3 (21.4%)
1+ (n = 122)	57 (96.6%)	40 (88.9%)	15 (83.3%)	27 (96.4%)	74 (92.5%)	11 (78.6%)
Previous surgery						
Yes	25 (42.4%)	14 (31.1%)	25 (42.4%)	14 (50%)	27 (33.8%)	3 (21.4%)
No	34 (57.6%)	31 (68.9%)	34 (57.6%)	14 (50%)	53 (66.3%)	11 (78.6%)
BMI	29.8 (6.5)	28.4 (5.4)	27.9 (7.6)	30.7 (6.1)	28.2 (5.7)	30.2 (9.0)
Baseline leakage	7.7 (4.4)	5.4 (3.4)	4.9 (5.0)	6.7 (5.6)	6.6 (4.0)	5.2 (2.7)
Baseline urgency	7.9 (3.4)	7.9 (3.3)	9.2 (3.1)	7.6 (3.5)	8.3 (3.4)	7.8 (2.2)
Baseline voiding	12.7 (7.5)	11.74 (3.6)	11.5 (2.5)	14.3 (10.3)	11.7 (3.3)	10.5 (1.9)
Smoking status						
Smoker	11 (18.6%)	12 (26.7%)	7 (38.9%)	7 (25%)	19 (23.8%)	4 (28.6%)
Non-smoker	48 (81.4%)	33 (73.3%)	11 (61.1%)	21 (75%)	61 (76.3%)	10 (71.4%)
Max voided vol	379.3 (141.7)	398.3 (161.6)	309.6 (126.5)	356.8 (180.5)	388.4 (136.0)	344.4 (154.1)
Av. voided vol	166.9 (72.3)	172.2 (56.2)	174.6 (84.9)	176.2 (99.9)	168.0 (56.2)	169.3 (56.0)
IUSS score	2.0 (0.5)	2.0 (0.5)	2.1 (0.5)	2.0 (0.6)	2.1 (0.5)	2.1 (0.4)
ICIQ score	16.2 (5.8)	13.8 (10.3)	11.9 (9.3)	15.7 (3.8)	14.5 (9.8)	13.9 (6.1)
IQOL score	20.7 (33.9)	19.9 (49.0)	33.8 (22.6)	26.8 (21.1)	19.1 (45.4)	32.1 (22.2)

For every additional increase in baseline leakage episodes, patients had an 17% increased odds of failing to achieve continence (OR: 1.17 95%CI 1.04, 1.31, $P = 0.007$). At the 10% significance threshold, age, baseline voiding frequency, and IQOL score were individually identified as potential predictors of non-response in voiding frequency. Age and BMI were associated with non-response on the PGI-I scale (Table V).

Multiple characteristics were found to be associated with non-response in voiding frequency, and non-response on the PGI-I scale, and thus these factors were entered in to a multivariate analysis. Having accounted for all associated factors in a multiple regression, both increasing age (OR 1.04, 95%CI 1.0, 1.09, $P = 0.063$) and BMI (OR 1.07, 95%CI 1.0, 1.16, $P = 0.065$) showed a marginal association with non-response on

TABLE III. Baseline Urodynamic Characteristics of the Patients Receiving Botulinum Toxin by Change in Urinary Diary Symptoms

	Change in urgency episodes <= 20%			Change in voiding frequency <= 10%			Change in leakage episodes <= 20%		
	Non-response	Response	Missing	Non-response	Response	Missing	Non-response	Response	Missing
n (122)									
Cystometry data									
Volume at first sensation									
n (%miss)	15 (48.3%)	24 (69.2%)	9 (40%)	12 (50%)	27 (68.6%)	9 (25%)	7 (63.2%)	32 (64.8%)	9 (25%)
Mean (SD)	116.1 (96.3)	128.5 (84.5)	74.7 (50.3)	119.35 (100.2)	125.7 (84.3)	74.7 (50.3)	141.6 (112.4)	119.8 (83.6)	74.7 (50.3)
Bladder capacity at cystometry									
n (%miss)	15 (48.3%)	26 (66.7%)	11 (26.7%)	12 (50%)	30 (65.1%)	10 (16.7%)	7 (63.2%)	35 (61.5%)	10 (16.7%)
Mean (SD)	349.7 (125)	376.3 (126.8)	291.1 (114.2)	384.2 (96.8)	360.0 (133.9)	282.3 (116.4)	427.1 (107.2)	354.8 (124.7)	282.3 (116.4)
Maximum filling detrusor pressure									
n (%miss)	14 (51.7%)	26 (66.7%)	11 (26.7%)	12 (50%)	29 (66.3%)	10 (16.7%)	7 (63.2 %)	34 (62.6%)	10 (16.7%)
Mean (SD)	27 (28.2)	21.7 (16.6)	39.9 (26.3)	20.4 (24.1)	25.4 (19.7)	41.3 (27.3)	22.4 (30.1)	24.2 (19.1)	41.3 (27.3)
Maximum voiding detrusor pressure									
n (%miss)	10 (65.5%)	23 (70.5%)	8 (46.7%)	8 (66.7%)	26 (69.8%)	7 (41.7%)	5 (73.7%)	29 (68.1%)	7 (41.7%)
Mean (SD)	50.9 (23.2)	41.7 (16.3)	56 (25.4)	42.8 (21.3)	44.5 (18.1)	59.3 (25.6)	47.4 (21.9)	43.6 (18.3)	59.3 (25.6)
Detrusor pressure at max flow rate									
n (%miss)	10 (65.5%)	22 (71.8%)	6 (60%)	8 (66.7%)	24 (72.1%)	6 (50%)	5 (73.7%)	27 (70.3%)	6 (50%)
Mean (SD)	39.2 (20.5)	34 (14.4)	36.5 (22.7)	33.6 (18.2)	36.3 (16.1)	36.5 (22.7)	34.6 (16.2)	35.9 (16.7)	36.5 (22.7)
Volume at time of first contraction									
n (%miss)	12 (58.6%)	24 (69.2%)	11 (26.7%)	10 (58.3%)	27 (68.6%)	10 (16.7%)	5 (73.7%)	32 (64.8%)	10 (16.7%)
Mean (SD)	235.6 (176.1)	243.1 (162.8)	139.3 (117.0)	236.5 (163.8)	235.6 (168.7)	146.7 (120.5)	262.6 (218.2)	231.7 (159.3)	146.7 (120.5)
Amplitude of first contraction									
n (%miss)	12 (58.6%)	23 (70.5%)	11 (26.7%)	11 (54.2%)	25 (70.9%)	10 (16.7%)	6 (68.4%)	30 (67%)	10 (16.7%)
Mean (SD)	21.8 (8.2)	20.8 (14.6)	22.6 (16.3)	17.7 (7.4)	22 (14.5)	24.5 (15.9)	19.8 (7.4)	20.9 (13.7)	24.5 (15.9)

TABLE IV. Baseline Urodynamic Characteristics of the Patients Receiving Botulinum Toxin by Continence Status and Patient Global Impression of Improvement

	Incontinent at follow up			Patient Global Impression of Improvement		
	Non-response	Response	Missing	Non-response	Response	Missing
n (122)						
Cystometry data						
Volume at first sensation (mls)						
n (%miss)	24 (59.3%)	13 (71.1%)	11 (38.9%)	9 (67.9%)	27 (66.3%)	12 (14.3%)
Mean (SD)	126 (94.3)	123 (86.4)	79.5 (46.3)	81.8 (60.3)	131.6 (94.7)	100.8 (68.1)
Bladder capacity at cystometry (mls)						
n (%miss)	25 (57.6%)	15 (66.7%)	12 (33.3%)	11 (60.7%)	29 (63.8%)	12 (14.3%)
Mean (SD)	376.4 (138.7)	347.3 (105.8)	301.1 (114.0)	360.1 (133.9)	367.1 (123.0)	302.2 (123.3)
Maximum filling detrusor pressure (cm H_20)						
n (%miss)	24 (59.3%)	15 (66.7%)	12 (33.3%)	11 (60.7%)	28 (65%)	12 (14.3%)
Mean (SD)	24.4 (25.2)	24.4 (14.0)	36.9 (26.8)	31.7 (29.7)	23.1 (17.3)	33.1 (28.1)
Maximum voiding detrusor pressure (cm H_20)						
n (%miss)	21 (64.4%)	12 (73.3%)	8 (55.6%)	9 (67.9%)	22 (72.5%)	10 (28.7%)
Mean (SD)	42.5 (21.7)	45.8 (12.4)	59 (23.7)	52.6 (23.5)	42.1 (15.3)	51.5 (26.5)
Detrusor pressure at max flow rate (cm H_20)						
n (%miss)	21 (64.4%)	10 (77.8%)	7 (61.1%)	8 (71.4%)	21 (73.8%)	9 (35.7%)
Mean (SD)	34.3 (17.7)	36.4 (13.3)	39.4 (22.1)	36.6 (22.0)	36.2 (13.3)	34.1 (22.1)
Volume at time of first contraction (mls)						
n (%miss)	20 (66.1%)	15 (66.7%)	12 (33.3%)	10 (64.3%)	25 (68.8%)	12 (14.3%)
Mean (SD)	255.5 (184.4)	202.9 (140.1)	170 (133.3)	191 (155.5)	239.5 (163.0)	191.4 (163.0)
Amplitude of first contraction (cm H_20)						
n (%miss)	19 (67.8%)	15 (66.7%)	12 (33.3%)	10 (64.3%)	24 (70%)	12 (14.3%)
Mean (SD)	21.5 (14.7)	20.6 (11.2)	22.7 (15.0)	18 (11.9)	20.7 (10.2)	26.2 (19.3)

Considerable data were missing for some patients; this is detailed in the "% miss" figure in brackets for each item.

the PGI-I scale. All other factors had a non-significant association with non-response in the presence of all other associated variables (Table VI).

The occurrence of voiding dysfunction or urinary tract infection had no effect upon the analyses above. In our sensitivity analysis, the cut-off of 10% or 50% did not alter the variables found to be significant in multivariate analysis.

DISCUSSION

In this study we assessed potential patient factors that predicted non-response at 6 weeks after active treatment. Smoking status and baseline leakage episodes were strongly associated factors with non-response to BoNTA. It is not surprising that more severe incontinence is less likely to be

TABLE V. Univariate Analysis of Patient Factors Associated With Lack of Benefit Against Each Outcome

		Change in urgency episodes <= 20%			
	n	Odds ratio	Standard error	95%CI	*P*-value
Characteristics					
Smoking status	107				
Non-smoker		Ref	Ref	Ref	Ref
Smoker		2.89	1.45	(1.08, 7.73)	0.034
constant		0.29	0.07	(0.17, 0.48)	<0.001
		Change in voiding frequency <= 10%			
Characteristics					
Age (centred)	110	1.04	0.02	(0.99, 1.08)	0.104
constant		0.26	0.06	(0.17, 0.42)	<0.001
Baseline voids	110	0.83	0.08	(0.68, 1.00)	0.056
constant		2.36	2.58	(0.28, 20.11)	0.432
IQOL Score	110	1.02	0.01	(1.00, 1.05)	0.048
constant		0.24	0.06	(0.15, 0.41)	<0.001
		Incontinent at follow up			
Baseline leakage	104	1.17	0.07	(1.04, 1.31)	0.007
constant		0.47	0.2	(0.21, 1.07)	0.073
		Patient global impression of improvement			
Age (centred)	108	1.04	0.02	(1.0, 1.09)	0.055
constant		0.32	0.08	(0.21, 0.51)	<0.001
BMI	108	1.07	0.04	(1.0, 1.16)	0.056
constant		0.34	0.08	(0.22, 0.54)	<0.001

Only significant associations are included. 95%CI, 95% confidence interval.

TABLE VI. Multivariate Analysis of Patient Factors Associated With Lack of Benefit Against Each Outcome

		Change in voiding frequency <= 10%			
	N	**Odds ratio**	**Standard error**	**95%CI**	***P*-value**
Characteristics					
Age (centred)	110	1.03	0.02	(0.99, 1.08)	0.172
Baseline voids		0.85	0.09	(0.69, 1.04)	0.107
IQol Score		1.02	0.01	(1.00, 1.04)	0.117
constant		1.61	1.88	(0.16, 15.88)	0.685
		Patient global impression of improvement			
Age (centred)	108	1.04	0.02	(1.0, 1.09)	0.063
BMI		1.07	0.04	(1.0, 1.16)	0.065
constant		0.32	0.08	(0.2, 0.51)	0.065

resolved after treatment, and it seems reasonable to advise the severely incontinent patients that they may have some residual leakage. The way in which smoking may be acting is not clear. It is interesting to note that smoking does not appear to reduce the efficacy of antimuscarinic drugs,[19] although it does increase the likelihood of discontinuation of oral medication.[20] It seems unlikely that an effect is acting via the nicotinic receptors in the parasympathetic ganglia, since onaBoNTA will be preventing the downstream release of neurotransmitter within the bladder. Smoking may be an indirect marker of cardiovascular changes leading to increased hypoxia in the bladder wall; episodes of hypoxia have been shown to increase detrusor contractility in vitro,[21] and in the bladders of atherosclerotic rabbits, with related ischaemia.[22,23]

Several authors have also examined clinical factors predictive of cure, from studies using a range of doses. As mentioned above, during the conduct of the trial from which our results are taken, 200 units of onaBoNTA was the accepted dose. Sahai et al.[24] analyzed data from 33 patients enrolled in their randomized study of 200 units of onaBoNT-A or placebo,[7] five of whom were deemed non-responders (based on patient reports and diary data, but not defined). Non-responsive patients had a significantly higher maximum detrusor pressure before treatment compared to responsive patients, but all other urodynamic factors were similar. Cohen et al.[25] analyzed data from 35 patients with overactive bladder who received intradetrusor onaBoNTA injections (100 and 150 U). Their definition of response was a 40% or more improvement from baseline, compared to our definitions of 10% and 20% (equivalent to one episode per day), but they did not identify any predictive factors. Schmid et al.[26] analyzed 100 patients with idiopathic OAB receiving 100 units of BoNTA and found that low bladder compliance, on pre-operative urodynamic assessment, and a maximal cystometric capacity less than 100 ml (confirmed on biopsy to be fibrosis), was seen in the eight patients who did not respond. Although our data did not confirm the findings from these two papers, it should be emphasized that these authors only conducted univariate analyses, so did not control for potential confounding between variables and did not explore the possibility of potential interactions. Our logistic regression accounts for this and so the data suggest that it is not possible to reliably identify patients who may fail to achieve benefit from onaBoNTA treatment.

It is interesting to note conflicting evidence on success rates related to antimuscarinic drug history. Makovey et al. reported that following 150–200 units of onaBoNTA success rates were lower in patients reporting lack of efficacy of antimuscarinic drugs (34 of 57, 68%), compared to those who stopped because of side effects (24 of 28, 86%).[27] In contrast, a pooled analysis of two trials of 100 units onaBoNTA showed no difference in treatment effect irrespective of the number of antimuscarinic preparations tried, or whether oral medication was stopped due to side effects or lack of efficacy.[28] It may be a possibility that some patients have a more resistant disease state, but whether this is a motor or sensory phenomenon is unclear. We are not able to comment on whether predictors of success vary by treatment dose, but there is no physiological or pharmacological reason why such a difference would exist.

There are limitations to this study. As a result of the efficacy of active treatment, there were few non-responders in the study population. Increasing the number of predictive variables in the model can therefore be problematic and reduce the power of the logistic regression model.[29] There is no agreed definition of what constitutes non-response after onaBoNTA treatment, as can be seen by the different definitions of non-response in the papers above.[24–26] In order to address this perceived limitation, we explored varying definitions of non-response in a sensitivity analysis. We tested non-response set at 10% or less, and 50% or less improvement; however, this had very little impact on the overall conclusions made, and the results mirrored that of the 20% analysis.

Some authors consider voiding dysfunction to be a significant factor in poor outcome after treatment, based on the findings of Brubaker et al.[8] and others.[12] We explored the influence of urinary retention and infection on non-response to treatment at 6 weeks, and found neither were independently associated. This is somewhat counter-intuitive so it may be that an assessment of efficacy at a longer interval would identify these as factors, but in the short term this does not appear to be the case. Due to the additional medication allowed after 6 weeks in our protocol, we are unable to assess this reliably.

A further limitation is the considerably large number of women with missing urodynamic data. Despite confirming the presence of detrusor contractions on the cystometry traces, 91 of the 122 (74.6%) had missing values for at least one of the baseline urodynamic factors. Where data were missing, the patient was excluded from the analyses involving that variable. To ameliorate the effect that this had on the complete case analysis we used multiple imputation techniques. We used multivariate normal regression to impute 10 datasets based on the data collected from the RELAX trial (inclusive of placebo patients); however, due to the limited urodynamic dataset in the original trial, there was considerable uncertainty in the imputed estimates and multiple imputation had very little benefit.

In conclusion, onaBoNTA is well established as a second-line treatment for patients with OAB and DO. However, not all women respond to treatment and being able to predict a

patient's likelihood of response would be clinically and economically advantageous. Previous studies have suggested that maximum detrusor pressure, poor compliance, and low maximum cystometric capacity were predictors of non-response, and identified urinary retention and infection as potential factors. Our regression analysis did not confirm these observations but identified baseline leakage episodes and smoking status as further predictors of non-response to BoNTA in patients with refractory DO.

REFERENCES

1. Abrams P, Cardozo L, Fall M, et al. The standardisation of terminology of lower urinary tract function: Report from the standardisation sub-committee of the international continence society. Neurourol Urodynam 2002;21:167–78.
2. Burgio KL, Locher JL, Goode PS, et al. Behavioral versus drug treatment for urge urinary incontinence in older women. JAMA 1998;280:1995–2000.
3. Kirkali Z, Whitaker RH. The use of oxybutynin in urological practice. Int Urol Nephrol 1987;19:385–91.
4. Abrams P, Freeman RM, Anderstrom C, et al. Tolterodine, a new antimuscarinic agent: As effective but better tolerated than oxybutynin in patients with an overactive bladder. Br J Urol 1999;81:801–10.
5. Chapple CR, Rechberger T, Al-Shukri S, et al. Randomized, double-blind placebo- and tolterodine-controlled trial of the once-daily antimuscarinic agent solifenacin in patients with symptomatic overactive bladder. BJU Int 2004;93:303–10.
6. Kelleher CJ, Cardozo LD, Khullar V, et al. A medium-term analysis of the subjective efficacy of treatment for women with detrusor instability and low bladder compliance. Br J Obstet Gynaecol 1997;104:988–93.
7. Sahai A, Khan MS, Dasgupta P. Efficacy of botulinum toxin-A for treating idiopathic detrusor overactivity: Results from a single center, randomized, double-blind, placebo controlled trial. J Urol 2007;177:2231–6.
8. Brubaker L, Richter HE, Visco A, et al. Refractory idiopathic urge urinary incontinence and botulinum A injection. J Urol 2008;180:217–22.
9. Flynn MK, Amundsen CL, Perevich M, et al. Outcome of a randomized, double-blind, placebo controlled trial of botulinum A toxin for refractory overactive bladder. J Urol 2009;181:2608–15.
10. Dmochowski R, Chapple C, Nitti VW, et al. Efficacy and safety of onabotulinumtoxinA for idiopathic overactive bladder: A double-blind, placebo controlled, randomized, dose ranging trial. J Urol 2010;184:2416–22.
11. Tincello DG, Kenyon S, Abrams KR, et al. Botulinum toxin A versus placebo for refractory detrusor overactivity in women: A randomized blinded placebo-controlled trial of 240 women (the RELAX study). Eur Urol 2012;62:507–14.
12. Denys P, Le NL, Ghout I, et al. Efficacy and safety of low doses of onabotulinumtoxinA for the treatment of refractory idiopathic overactive bladder: A multicentre, double-blind, randomized, placebo-controlled dose-ranging study. Eur Urol 2012;61:520–9.
13. Nitti VW, Dmochowski R, Herschorn S, et al. OnabotulinumtoxinA for the treatment of patients with overactive bladder and urinary incontinence: Results of a phase 3, randomized, placebo controlled trial. J Urol 2013;189:2186–93.
14. Nixon A, Colman S, Sabounjian L, et al. A validated patient reported measure of urinary urgency severity in overactive bladder for use in clinical trials. J Urol 2005;174:604–7.
15. Avery K, Donovan J, Peters TJ, et al. ICIQ: A brief and robust measure for evaluating the symptoms and impact of urinary incontinence. Neurourol Urodynam 2004;23:322–30.
16. Patrick DL, Martin ML, Bushnell DM, et al. Quality of life of women with urinary incontinence: Further development of the incontinence quality of life instrument (I-QOL). Urology 1999;53:71–6.
17. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. Am J Obstet Gynecol 2003;189:98–101.
18. Tincello DG, Owen RK, Slack MC, et al. Validation of the patient global impression scales for use in detrusor overactivity: Secondary analysis of the RELAX study. BJOG 2013;120:212–6.
19. Schneider T, Marschall-Kehrel D, Hanisch JU, et al. Do gender, age or lifestyle factors affect responses to antimuscarinic treatment in overactive bladder patients? Int J Clin Pract 2010;64:1287–93.
20. Brubaker L, Fanning K, Goldberg EL, et al. Predictors of discontinuing overactive bladder medications. BJU Int 2010;105:1283–90.
21. Agartan CA, Leggett RE, Kogan BA, et al. Effect of age on the response to in vitro ischemia and reperfusion of the rabbit bladder. Urol Int 2007;78:155–9.
22. Azadzoi KM, Tarcan T, Kozlowski R, et al. Overactivity and structural changes in the chronically ischaemic bladder. J Urol 1999;162:1768–78.
23. Azadzoi KM, Tarcan T, Siroky MB, et al. Atherosclerosis-induced chronic ischaemia causes bladder fibrosis and non-compliance in the rabbit. J Urol 1999;161:1635.
24. Sahai A, Khan MS, Le GN, et al. Urodynamic assessment of poor responders after botulinum toxin-A treatment for overactive bladder. Urology 2008;71:455–9.
25. Cohen BL, Caruso DJ, Kanagarajah P, et al. Predictors of response to intradetrusor botulinum toxin-A injections in patients with idiopathic overactive bladder. Adv Urol 2009 (2009), Article ID 328364. DOI: 10.1155/2009/328364
26. Schmid DM, Sauermann P, Werner M, et al. Experience with 100 cases treated with botulinum-A toxin injections in the detrusor muscle for idiopathic overactive bladder syndrome refractory to anticholinergics. J Urol 2006;176:177–85.
27. Makovey I, Davis T, Guralnick ML, et al. Botulinum toxin outcomes for idiopathic overactive bladder stratified by indication: Lack of anticholinergic efficacy versus intolerability. Neurourol Urodynam 2011;30:1538–40.
28. Sievert KD, Chapple C, Herschorn S, et al. OnabotulinumtoxinA 100U provides significant improvements in overactive bladder symptoms in patients with urinary incontinence regardless of the number of anticholinergic therapies used or reason for inadequate management of overactive bladder. Int J Clin Pract 2014;68:1246–56.
29. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.

F.3 Hierarchical network meta-analysis: development of a three-level hierarchical modelling approach incorporating dose-related constraints

In this paper, I contributed to study design, undertook the full systematic review and data extraction, performed all of the statistical analyses, wrote the first draft of the paper and critically revised the final manuscript. My total contribution was approximately 70% of the research paper.

VALUE IN HEALTH 18 (2015) 116–126

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Network Meta-Analysis: Development of a Three-Level Hierarchical Modeling Approach Incorporating Dose-Related Constraints

Rhiannon K. Owen, BSc, MSc[1,*], *Douglas G. Tincello, BSc, MBChB, MD, FRCOG*[2], *Keith R. Abrams PhD, CStat*[1]

[1]*Department of Health Sciences, University of Leicester, Leicester, UK;* [2]*Reproductive Science Section, Department of Cancer Studies and Molecular Medicine, University of Leicester, Leicester, UK*

ABSTRACT

Background: Network meta-analysis (NMA) is commonly used in evidence synthesis; however, in situations in which there are a large number of treatment options, which may be subdivided into classes, and relatively few trials, NMAs produce considerable uncertainty in the estimated treatment effects, and consequently, identification of the most beneficial intervention remains inconclusive. **Objective:** To develop and demonstrate the use of evidence synthesis methods to evaluate extensive treatment networks with a limited number of trials, making use of classes. **Methods:** Using Bayesian Markov chain Monte Carlo methods, we build on the existing work of a random effects NMA to develop a three-level hierarchical NMA model that accounts for the exchangeability between treatments within the same class as well as for the residual between-study heterogeneity. We demonstrate the application of these methods to a continuous and binary outcome, using a motivating example of overactive bladder. We illustrate methods for incorporating ordering constraints in increasing doses, model selection, and assessing inconsistency between the direct and indirect evidence. **Results:** The methods were applied to a data set obtained from a systematic literature review of trials for overactive bladder, evaluating the mean reduction in incontinence episodes from baseline and the number of patients reporting one or more adverse events. The data set involved 72 trials comparing 34 interventions that were categorized into nine classes of interventions, including placebo. **Conclusions:** Bayesian three-level hierarchical NMAs have the potential to increase the precision in the effect estimates while maintaining the interpretability of the individual interventions for decision making.

Keywords: network meta-analysis, statistical methods, mixed treatment comparisons, overactive bladder.

Introduction

Network meta-analyses (NMA) are widely used in an evidence synthesis setting due to the attractive nature of utilizing all relevant information from both direct and indirect evidence [1–4]. Nevertheless, in situations in which there are a large number of interventions of interest and relatively few trials, there is a potential issue with the sparsity of data in the treatment networks, which can lead to parameter uncertainty. Collapsing the intervention arms into their respective treatment classes increases the evidence base and precision in the effect estimates, but with such a class-based approach, the direct interpretation of individual intervention effects is lost, which makes decision making difficult. To overcome this issue, a three-level hierarchical NMA can be applied [5–7]. This approach incorporates the exchangeability between interventions of the same class to predict an effect estimate for each of the interventions individually [8]. Thus, this approach allows strength to be borrowed within the classes of interventions, strengthening inferences and potentially reducing the uncertainty around the individual intervention effects, and consequently increasing the ability to rank these and inform decision-making frameworks. To further increase the precision in the effect estimates, constraints can be applied on increasing doses of the same intervention, making the assumption that higher doses have an effect greater or equal to that of lower doses [9,10].

To illustrate the use of the hierarchical framework, we applied the proposed methods to a real clinical question in overactive

Conflict of interest: Douglas G. Tincello has received consultancy fees from Ethicon and Allergan in the last 3 years. Keith Abrams has acted as a paid consultant providing methodological advice to the health care industry generally, including to both Allergan and Astellas.

* *Address correspondence to*: Rhiannon K. Owen, Department of Health Sciences, University of Leicester, Room 212, Adrian Building, University Road, Leicester LE1 7RH, UK.

E-mail: rhiannon.owen@leicester.ac.uk.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2014.10.006

bladder (OAB) syndrome. To manage the OAB syndrome, the National Institute for Health and Care Excellence in the United Kingdom [11] currently recommends a course of supervised pelvic floor muscle training, behavioral therapy, anticholinergic medication, sacral nerve stimulation, and more recently, botulinum toxin type A (BoNTA). Given the availability of numerous interventions and emerging alternative treatments such as BoNTA, there is an increasing need to identify the most beneficial intervention. However, given the large number of interventions and the limited evidence base, in terms of both the number of trials and the number of direct comparisons between active interventions, the estimated intervention effects from a standard NMA will have a considerable level of uncertainty associated with them. In situations in which there are a limited number of trials in a meta-analysis, estimating the heterogeneity between trials may also be problematic. One approach to overcome this issue, and increase precision in the treatment effects, involves incorporating external information from similar studies relevant to the treatment of interest [12]. In an NMA that includes all available trials in a specific field, however, such external information may be limited. The aim of this article was to develop and apply hierarchical NMAs to evaluate the clinical effectiveness of interventions for the OAB syndrome by borrowing strength between interventions of the same class and applying ordering constraints on increasing doses of BoNTA, thus increasing the precision that we have in our effect estimates but maintaining the interpretability of results at the individual intervention level. For illustration purposes, we focus on two outcomes associated with intervention effectiveness (mean change in incontinence episodes from baseline) and treatment tolerability (number of patients reporting one or more adverse events).

In this article, we demonstrate the individual treatment, class-based, and three-level hierarchical random effects model approaches, and where applicable we demonstrate the use of extending hierarchical NMAs to incorporate ordering constraints. We apply these models to a motivating clinical example in the OAB syndrome. Furthermore, we demonstrate a comprehensive technique to assess inconsistency between the direct and indirect estimates of an extensive network using the method of node-splitting [13] and assess model fit using residual deviance [14] and the deviance information criterion (DIC) [15].

Methods

Illustrative Data Set

Almost all published articles reporting data on interventions for the OAB syndrome compare the intervention against placebo, which makes comparison across active interventions difficult without using indirect comparisons or NMA. This is particularly evident for trials evaluating anticholinergic drugs. Only three meta-analyses have been undertaken in the field of the OAB syndrome [16–18]. The interventions were assessed on a head-to-head basis, where studies comparing the interventions directly were pooled in a pairwise meta-analysis. Chapple et al. [16] focused on the evaluation of the clinical effectiveness of anticholinergic drugs compared with placebo, while Novara et al. [17] compared the efficacy of increased doses of each anticholinergic drug with that of their respective lower dose. Anger et al. [18] evaluated the effect of BoNTA against that of a placebo intervention. In the current literature, there is no coherent comparison between all the available interventions, and consequently, there is little information of a superior treatment for the OAB syndrome.

Figure 1A,B illustrates the network diagrams of direct comparisons for the individual intervention and classes of

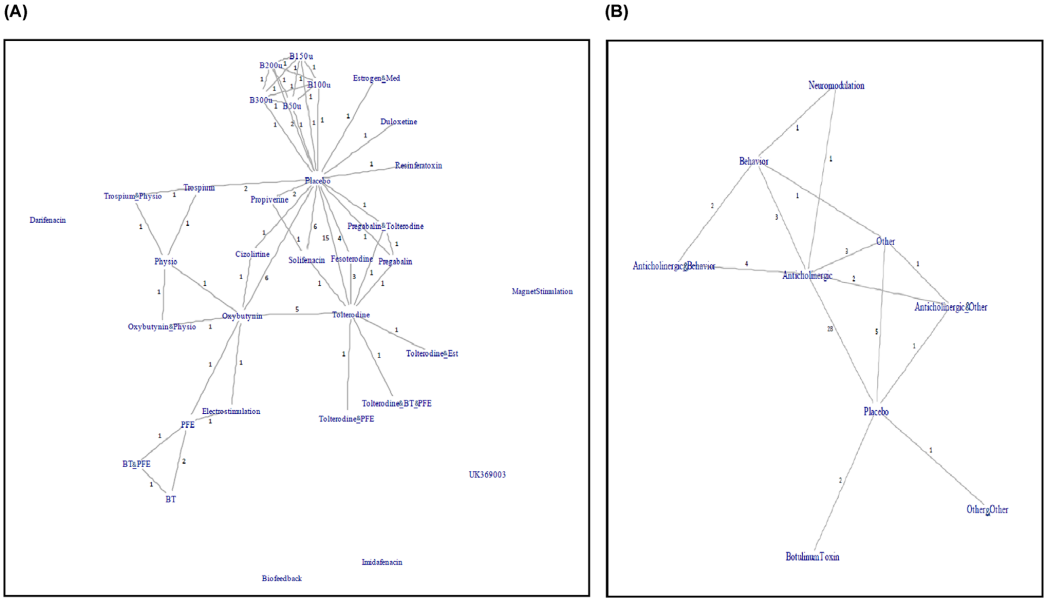


Fig. 1 – Network diagrams for urinary incontinence. (A) Individual and hierarchical network diagram. (B) Classified network diagram. B50u, Botulinum toxin type A 50 units; B100u, Botulinum toxin type A 100 units; B150u, Botulinum toxin type A 150 units; B200u, Botulinum toxin type A 200 units; B300u, Botulinum toxin type A 300 units; BT, Bladder training; Est, Estrogen; Med, Medroxyprogesterone; PFE, Pelvic floor exercises; Physio, Physiotherapy.

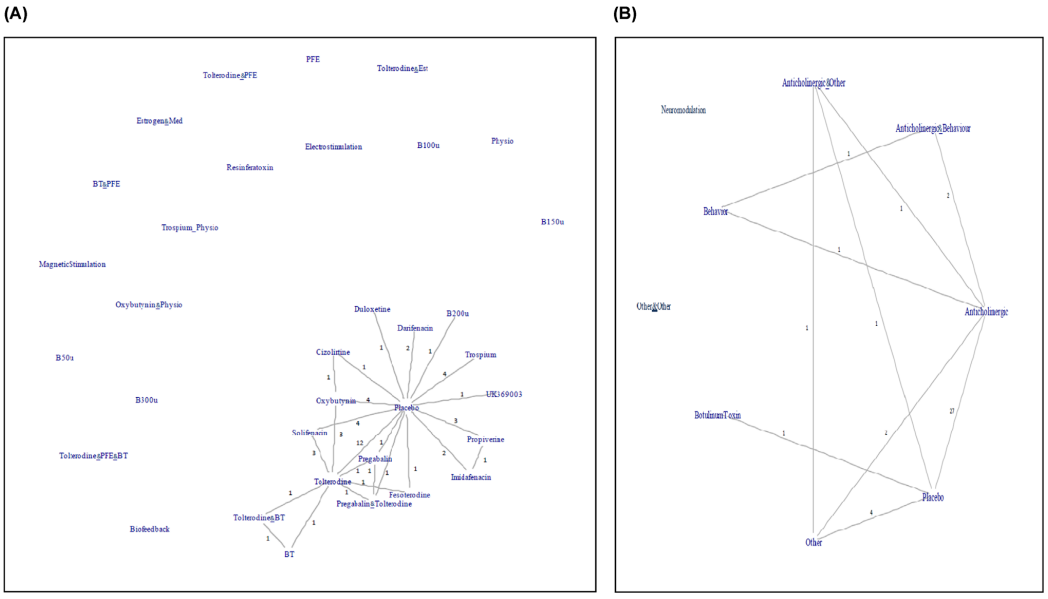


Fig. 2 – Network diagrams for adverse events. (A) Individual and hierarchical network diagram. (B) Class-based network diagram. B50u, Botulinum toxin type A 50 units; B100u, Botulinum toxin type A 100 units; B150u, Botulinum toxin type A 150 units; B200u, Botulinum toxin type A 200 units; B300u, Botulinum toxin type A 300 units; BT, Bladder training; Est, Estrogen; Med, Medroxyprogesterone; PFE, Pelvic floor exercises; Physio, Physiotherapy.

interventions that evaluate the mean reduction in incontinence episodes, respectively. Similarly, Figure 2A,B demonstrates the network diagrams of the individual intervention and classes of interventions that evaluate the number of patients reporting one or more adverse events, respectively. The nodes represent either the individual intervention or classes of interventions. The interconnecting lines demonstrate a direct comparison, and the corresponding values represent the number of trials that directly compare those interventions. For interventions identified in the systematic literature review but disconnected from the network (i.e., fail to report outcome of interest), we were unable to obtain effect estimates and thus these were excluded from the analysis.

For analyses associated with the class of interventions, treatments were grouped, according to clinical opinion (D.G.T.), into anticholinergic drug therapy, botulinum toxin, neuromodulation, behavior therapy, other drugs, anticholinergic drugs in combination with behavior therapy, anticholinergic drugs in combination with other drugs, and a combination of other drugs. Figure 3 demonstrates the classification of each of the individual interventions, where the central node represents the class of treatments and the linked arms represent each of the individual interventions within that class. In this example, anticholinergic drug therapy consisted of all the members of the anticholinergic class of drugs. The botulinum toxin group contained all BoNTA interventions regardless of the site of administration or dose. The neuromodulation classification included all interventions involved in nerve stimulation or electrostimulation. *Behavior therapy* was defined as interventions that focused on attaining change in behavioral factors relevant to symptoms of the OAB syndrome, including physiotherapy, bladder retraining, and biofeedback. *Other drugs* were defined as all other pharmacotherapy interventions that were not classified as anticholinergic or BoNTA therapies.

Individual Treatment and Class-Based NMA Models

Equations 1 and 2 illustrate the general model described by Welton et al. [19] for the continuous and binary outcome case, respectively. It is these models that form the foundation for both the individual treatment and class-based NMAs.

For an intervention j, in study i, the continuous outcome can be interpreted as the mean change in 3-day diary data for the number of incontinence episodes from baseline y_{ij} and assumed to follow a normal distribution with mean equal to the underlying intervention effect $\theta_{i,j}$ and observed standard error $SE_{i,j}$. Let μ_i represent the baseline mean change in the number of incontinence episodes corresponding to the b_i intervention arm in the ith study, and let $\delta_{i,j}$ represent the mean difference in change in the number of incontinent episodes of intervention j relative to the b_i. intervention arm. $\delta_{i,j}$ is obtained from a normal distribution with the mean equal to the mean differences $(d_j - d_{b_i})$ and between-study variance τ^2, where d_j and d_{b_i} represent the effect estimate of intervention j and study-specific baseline intervention b_i, respectively. Notably, when the between-study variance is zero, that is, $\tau^2 = 0$, we obtain a fixed effects model. Thus, the overall model is based on a linear regression model on a natural additive scale:

$$y_{i,j} \sim \text{Normal}(\theta_{i,j}, SE_{i,j}^2)$$

where

$$\theta_{i,j} = \begin{cases} \mu_i & \text{Intervention } b_i \\ \mu_i + \delta_{i,j} & \text{Intervention } j \end{cases}$$

and

$$\delta(i,j) \sim \text{Normal}((d_j - d_{b_i}), \tau^2) \quad (1)$$

The number of reported adverse events is considered a binomial count $r_{i,j}$ from a sample number at risk $n_{i,j}$ for an intervention j

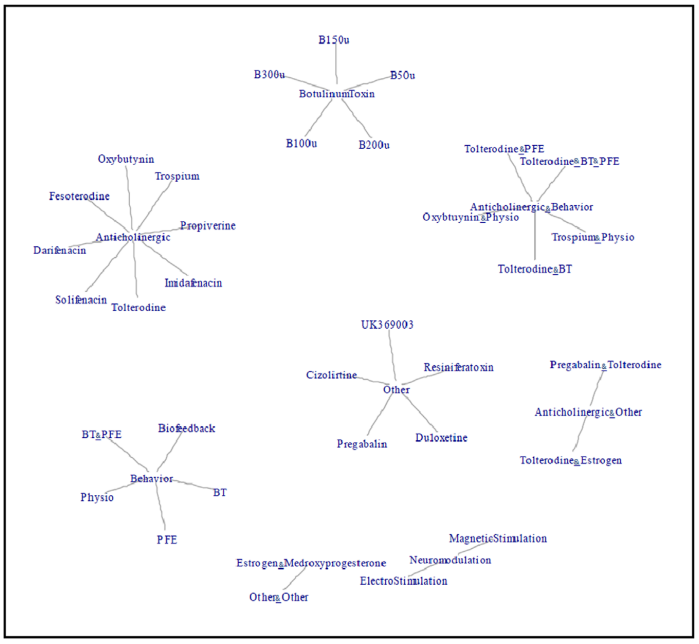


Fig. 3 – Treatment classification. B50u, Botulinum toxin type A 50 units; B100u, Botulinum toxin type A 100 units; B150u, Botulinum toxin type A 150 units; B200u, Botulinum toxin type A 200 units; B300u, Botulinum toxin type A 300 units; BT, Bladder training; PFE, Pelvic floor exercises; Physio, Physiotherapy.

within the ith study. This information allows estimation of the probability $p_{i,j}$, which is associated with the risk of adverse events. We assume a logistic regression model for the binary outcome. Thus, $\varphi_{i,j}$ represents the log-odds of treatment j relative to a baseline treatment b_i with between-study variance τ^2. In this case, d_j and d_{b_i} represent the estimated log-odds of intervention j and baseline intervention b_i, respectively. Therefore, the overall model is given by

$$r_{i,j} \sim \text{Binomial}(p_{i,j}, n_{i,j})$$

where

$$\text{logit}(p_{i,j}) = \begin{cases} \mu_i & \text{Intervention } b_i \\ \mu_i + \delta_{i,j} & \text{Intervention } j \end{cases}$$

and

$$\varphi(i,j) \sim \text{Normal}((d_j - d_{b_i}), \tau^2) \quad (2)$$

The baseline intervention means μ_i are assumed to have a normal (0, 1000) prior distribution. Therefore, for the continuous case, the mean reduction in incontinence episodes from baseline for the reference intervention could plausibly be in the range of 0 ± 62 incontinence episodes. For the binary case, the log-odds of an adverse event could plausibly be in the range of 0 ± 62. The between-study standard deviation values of τ are assumed to have a uniform (0, 5) prior distribution, suggesting that the between-study SD can take any value between, but not including, 0 and 5, and small values of τ are equally likely as large values [1]. A value of 5, for example, would indicate that for a random pair of studies, the difference in the mean reduction in incontinence episodes from baseline could be as large as 5.5 while the ratio of odds ratios could be as large as 232.8. Sensitivity analyses considering two other variance-component priors were considered: 1) gamma (0.001, 0.001) on the precision scale, that is, 1/variance, and 2) half-normal (0, 1) on the SD scale [1].

Hierarchical NMA

A random effects model was used to estimate the effect of each individual intervention for both continuous (Equation 3) and binary (Equation 4) outcomes. To account for the exchangeability between the treatments within each class, the treatments within each class were assumed to follow a normal distribution with a class-specific mean and variance (Equation 5).

$$y_{i,j} \sim \text{Normal}(\theta_{i,j}, \text{SE}_{i,j}^2)$$

$$\theta_{i,j} = \begin{cases} \mu_i & \text{Intervention } b_i \\ \mu_i + \delta^*_{i,j,k} & \text{Intervention } j \end{cases} \quad (3)$$

$$r_{i,j} \sim \text{Binomial}(p_{i,j}, n_{i,j})$$

$$\text{logit}(p_{i,j}) = \begin{cases} \mu_i & \text{Intervention } b_i \\ \mu_i + \varphi^*_{i,j,k} & \text{Intervention } j \end{cases} \quad (4)$$

Following Dakin et al. [5] and Warren et al. [6], the effect estimate for a specific intervention class combination $d_{j,k}$ is described as

$$d_{j,k} \sim \text{Normal}(\mu_k, \sigma^2) \quad (5)$$

where μ_k denotes the pooled effect estimate for the kth class of interventions, with a common between-intervention variance σ^2. Class-specific between-intervention variances, σ_k^2, were also considered in exploratory analyses and assessed through model fit statistics.

At the individual intervention level, the effect estimate compared with that of a baseline treatment $\delta^*_{i,j,k}$ and $\varphi^*_{i,j,k}$ for the continuous case and the binary case, respectively, is expressed in terms of the effect estimate for a specific individual intervention class $d_{j,k}$, compared with a class-specific baseline treatment $d_{b_{i,k}}$, given by

$$\delta^*_{i,j,k} \sim \text{Normal}((d_{j,k} - d_{b_{i,k}}), \tau^2) \quad \text{for the continuous case}$$

and

$$\varphi^*_{i,j,k} \sim \text{Normal}((d_{j,k} - d_{b_{i,k}}), \tau^2) \quad \text{for the binary case} \tag{6}$$

where $\delta^*_{i,j,k}$ represents the estimated mean difference of treatment j compared with a baseline treatment b_i and $\varphi^*_{i,j,k}$ represents the estimated log-odds of treatment j relative to a baseline treatment b_i. At level 3 in the hierarchical model, the between-study variance τ^2 was given a uniform (0, 5) prior distribution as in the random effect NMA model (see Equations 1 and 2). The pooled class-effect estimates μ_k were assumed to have a vague normal (0, 1000) prior distribution, with a common between-intervention variance σ^2 assumed to follow a uniform (0, 5) distribution on the SD scale. Thus, the pooled class-effect estimates could plausibly take values in the range of 0 ± 62 change in incontinent episodes from baseline for the continuous case and 0 ± 62 log-odds of an adverse event for the binary case. The common between-intervention variance could plausibly take values between, but not including, 0 and 5. Two other prior distributions for the variance for σ^2 were considered in the form of a sensitivity analysis: 1) gamma (0.001, 0.001) on the precision scale, and 2) half-normal (0, 1) on the SD scale [1].

For the continuous outcome, treatments were ranked on the basis of posterior distributions of the relative effect estimate δ and δ^*, where treatments with the largest relative reduction in mean incontinence episodes were ranked first for each Markov chain Monte Carlo (MCMC) iteration from the individual treatment and the hierarchical model, respectively. The estimated rankings overall were then calculated from a summary of these individual ranks at each iteration. Therefore, a higher rank indicates a more efficacious intervention overall. Similarly, for the binary outcome, the treatments were individually ranked on their posterior summaries φ and φ^*, where treatments with the highest log-odds of an adverse event were ranked in the first place for the individual treatment and the hierarchical model, respectively, where a higher rank indicates a larger prevalence. Thus, interventions ranked first are regarded as the "worst" treatments associated with adverse events. The corresponding probabilities were calculated by monitoring the number of MCMC iterations for which each of the treatments was ranked in the first place.

Incorporating Constraints on Increasing Doses

Ordering constraints were placed on multiple doses of BoNTA interventions, with the assumption that larger doses would have a greater or equal treatment effect compared with its respective lower dose (e.g., $d_1 \le d_2 \le \cdots \le d_m$). We applied these constraints by assigning an indicator function γ, equal to 1, given by

$$\gamma = \prod_{l=1}^{m-1} I(d_{l+1} - d_l) \tag{7}$$

where $I(x) = 1$, if $x \ge 0$, and $I(x) = 0$, otherwise. This forces $(d_{l+1} - d_l) \ge 0$ and consequently imposes ordering constraints on the treatment effects of increasing doses (i.e., $d_l \le d_{l+1}$) [7]. Ordering constraints can be placed in either direction depending on the outcome of interest [6].

Assessment of Inconsistency

Consistencies between direct and indirect comparisons were evaluated using the method of "node-splitting" [13]. This approach allows the calculation of two posterior distributions, one of which is derived from trials that directly compare the interventions (e.g., interventions *X* and *Y*), d^{Dir}_{XY}, whereas the other is indirectly derived from the remaining trials d^{Ind}_{XY}. The fundamental model described in Equations (1) and (2) remains the same; however, for direct comparison, the effect estimates δ_{iXY} obtained from splitting the (*X*, *Y*) node are selected from a normal distribution with mean d^{Dir}_{XY} and variance sd^2, that is,

$$\delta iXY \sim N(d^{\text{Dir}}_{XY}, sd^2) \tag{8}$$

Simultaneously, indirect comparisons are obtained using the consistency assumption, which states that for treatment effects d_{XY}, d_{XZ}, and d_{YZ}, relative to treatments X, Y, and Z,

$$\begin{aligned} d_{XY} &= d_{YZ} - d_{XZ} \\ d_{YZ} &= d_{XZ} - d_{XY} \\ d_{XZ} &= d_{YZ} - d_{XY} \end{aligned} \tag{9}$$

To test for consistency between direct and indirect estimates, we simply calculated the difference for each pair of interventions, together with the probability that the direct estimate surpasses that of the indirect estimate. Thus, a Bayesian P value can be calculated using the derived test statistic and comparing it with a standard normal distribution. This method, however, can only be applied to interventions within a closed loop, meaning that there is both direct and indirect evidence available for all pairs of interventions under consideration [13].

Model Fit and Selection

The DIC [15] was used to compare models. It is a measure of the deviance, estimated by the posterior mean of minus twice the log-likelihood plus the effective number of parameters in the model. Thus, it is considered as a Bayesian measure of goodness of fit that can be used as a relative measure of model suitability and easily applied to hierarchical modeling [20]. In parallel, the total residual deviances for each of the models were also compared with the respective number of data points. To illustrate and assess the goodness of fit for each of the models, we plotted the residual deviances for each of the included studies against the respective number of data points for that study [14], that is, 2 for two arm studies, 3 for three arm studies, and so forth.

Model Estimation

All models were estimated using WinBUGS 1.4.3 [21]. The results are based on 60,000 samples, where the first 10,000 samples were discarded from the analyses as a "burn-in." Three individual chains with disparate starting values were analyzed and convergence was assessed using Brooks-Gelman-Rubin plots [22]. Sensitivity analyses were also undertaken to assess the impact of the choice of prior distributions especially for the variance parameters [1,23]. Full codes for continuous and binary models are given in Appendices A and B in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.10.006, respectively. Both node-splitting analyses and network diagrams were implemented in R [24] using the "R2WinBUGs" software package [25] and the GeMTC package [26], respectively.

Results

Model Fit and Selection

Table 1 contains the goodness-of-fit statistics for each of the models individually. Notably, analyses for class-based models were calculated on different data sets—a consequence of the treatment clustering into endonodal treatment classes [27], which resulted in the omission of several studies that compared

Table 1 – Goodness-of-fit statistics for fixed and random effects models.

Outcome	Model		Between-study SD (95% CrI)	Residual deviance (no. of data points)	DIC
Number of incontinence episodes	Individual	FE	–	158.1 (112)	51.68
		RE	0.20 (0.12–0.31)	116.9 (112)	29.05
	Class*	FE	–	118.3 (87)	1.89
		RE	0.15 (0.06–0.25)	90.66 (87)	–11.53
	Hierarchical	FE	–	158.6 (112)	41.33
		RE	0.20 (0.12–0.30)	114.9 (112)	17.02
	Hierarchical with constraints	FE	–	158.8 (112)	41.37
		RE	0.20 (0.12–0.30)	114.5 (112)	15.81
Number of patients reporting one or more adverse events	Individual	FE	–	134.5 (79)	591.97
		RE	0.32 (0.21–0.49)	74.42 (79)	547.74
	Class*	FE	–	155.6 (68)	554.81
		RE	0.36 (0.25–0.52)	65.74 (68)	484.85
	Hierarchical	FE	–	135 (79)	592.66
		RE	0.30 (0.20–0.46)	75.43 (79)	548.68

CrI, credible interval; DIC, deviance information criterion; FE, fixed effects; RE, random effects.
* Class-based analyses conducted on a different population and are not directly comparable.

two or more interventions from the same class. Thus, model fit statistics for class-based NMAs are solely presented for completeness and cannot be directly compared with the remaining models. In relation to individual analyses, hierarchical models appeared to have a slightly better fit to the data for both outcome measures as illustrated through the reduced residual deviance for continuous (Fig. 4A) and binary (Fig. 4B) outcomes. For the continuous outcome, incorporating ordering constraints slightly improved model fit further with respect to both the DIC and the total residual deviance. The hierarchical random effects model with ordering constraints had a lower DIC (15.81) than did the individual random effects model (29.05). Similarly, for the hierarchical random effects model with constraints, the total residual deviance of 114.5 was closer to the number of data points (112) compared with the individual random effects model, which had a total residual deviance of 116.9. The random effects models were of a better fit to the data in comparison to the fixed effects models for all sets of analyses and thus, for illustration purposes, the results presented are based on estimates derived from the random effects NMAs.

Number of Incontinence Episodes

Table 2 illustrates the interventions ranked in order of their estimated efficacy for reducing incontinence episodes. Effect estimates derived from hierarchical models correspond with estimates derived from the individual analysis; however, there was a substantial increase in the precision surrounding effect

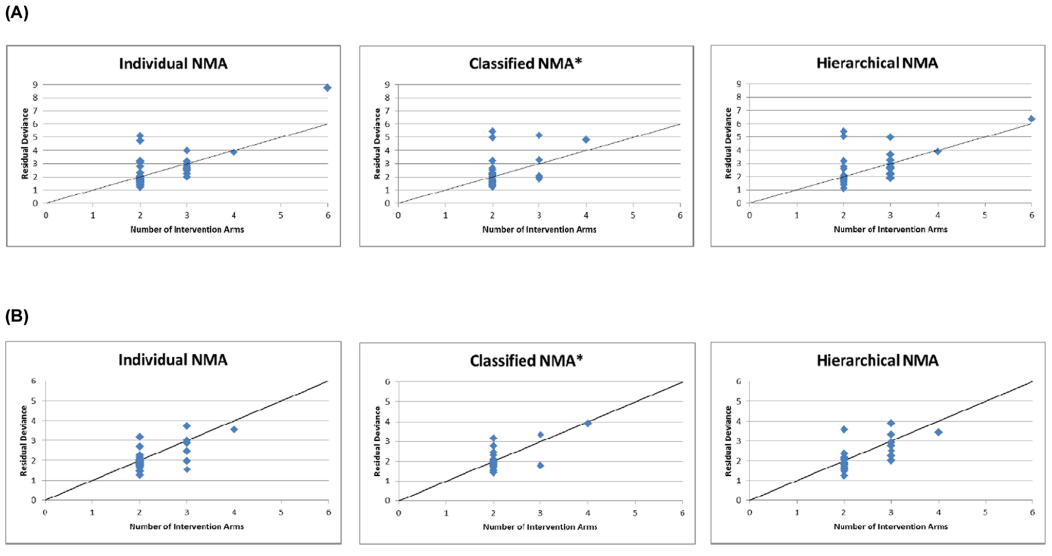


Fig. 4 – Residual deviance plots. (A) Incontinence episodes. (B) Number of patients reporting adverse events. NMA, network meta-analysis. *Classified analyses conducted on a different population and are not directly comparable.

Table 2 – Treatments placed in ranked order of their relative difference compared with placebo (δ and δ*) and corresponding probabilities for the mean reduction in incontinence episodes from baseline.

S. no.	Individual NMA				Hierarchical NMA				Hierarchical NMA with constraints			
	Treatment	Rank (95% Crl)	*p* (best) (%)	δ (95% Crl)	Treatment	Rank (95% Crl)	*p* (best) (%)	δ* (95% Crl)	Treatment	Rank (95% Crl)	*p* (best) (%)	δ* (95% Crl)
1	Trospium and physiotherapy	2 (1–10)	47.66	−1.89 (−2.97 to −0.82)	BoNTA 200 U	3 (1–13)	20.58	−1.33 (−2.13 to −0.61)	BoNTA 300 U	1 (1–7)	83.18	−1.44 (−2.09 to −0.78)
2	BoNTA 150 U	4 (1–22)	17.67	−1.45 (−2.73 to −0.17)	BoNTA 150 U	3 (1–16)	19.32	−1.31 (−2.15 to −0.55)	BoNTA 200 U	2 (2–10)	0	−1.38 (−2 to −0.73)
3	BoNTA 200 U	4 (1–12)	9.46	−1.46 (−2.28 to −0.65)	BoNTA 50 U	4 (1–16)	14.97	−1.29 (−2.13 to −0.54)	BoNTA 150 U	3 (3–12)	0	−1.32 (−1.94 to −0.68)
4	BoNTA 50 U	6 (1–23)	4.35	−1.17 (−2.26 to −0.09)	BoNTA 300 U	4 (1–17)	12.66	−1.27 (−2.12 to −0.51)	BoNTA 100 U	4 (4–14)	0	−1.26 (−1.89 to −0.61)
5	Physiotherapy	6 (2–19)	1.08	−1.15 (−1.93 to −0.38)	BoNTA 100 U	4 (1–18)	11.75	−1.27 (−2.11 to −0.49)	BoNTA 50 U	5 (5–18)	0	−1.19 (−1.84 to −0.51)
6	BoNTA 300 U	8 (2–25)	2.06	−0.97 (−2.10 to 0.17)	BT and PFE	9 (1–25)	5.29	−0.80 (−1.59 to −0.05)	BT and PFE	9 (1–24)	4.54	−0.82 (−1.53 to −0.14)
7	BT and PFE	9 (1–26)	11.16	−0.86 (−2.94 to 1.17)	Trospium and physiotherapy	9 (1–22)	3.648	−0.84 (−1.36 to −0.38)	Trospium and physiotherapy	9 (1–20)	2.88	−0.85 (−1.31 to −0.43)
8	Oxybutynin and physiotherapy	10 (2–25)	1.22	−0.81 (−1.77 to 0.11)	PFE	10 (1–25)	3.67	−0.77 (−1.55 to −0.03)	PFE	10 (1–24)	2.85	−0.79 (−1.47 to −0.13)
9	BoNTA 100 U	10 (2–26)	1.17	−0.85 (−1.99 to 0.31)	Oxybutynin and physiotherapy	10 (2–23)	1.748	−0.78 (−1.26 to −0.33)	Tolterodine, PFE, and BT	10 (2–21)	1.22	−0.79 (−1.19 to −0.39)
10	PFE	11 (1–26)	3.07	−0.72 (−2.61 to 1.15)	Physiotherapy	11 (2–24)	1.724	−0.75 (−1.46 to −0.08)	Oxybutynin and physiotherapy	10 (3–22)	1.14	−0.78 (−1.2 to −0.37)
11	Tolterodine, PFE, and BT	11 (3–24)	0.57	−0.78 (−1.55 to −0.02)	Duloxetine	11 (2–23)	1.216	−0.76 (−1.21 to −0.33)	Physiotherapy	11 (2–24)	1.43	−0.77 (−1.37 to −0.16)
12	Tolterodine and estrogen	11 (4–21)	0.1	−0.78 (−1.27 to −0.29)	Tolterodine and PFE	11 (2–23)	1.11	−0.75 (−1.21 to −0.30)	Duloxetine	11 (3–22)	0.68	−0.76 (−1.15 to −0.37)
13	Solifenacin	11 (6–16)	0	−0.77 (−0.96 to −0.57)	Bladder training	13 (2–26)	0.994	−0.67 (−1.46 to 0.12)	Tolterodine and PFE	11 (3–22)	0.63	−0.75 (−1.14 to −0.34)
14	Duloxetine	14 (5–24)	0.04	−0.60 (−1.22 to 0.02)	Tolterodine and estrogen	13 (5–23)	0.426	−0.66 (−1.02 to −0.28)	Bladder training	13 (3–26)	0.75	−0.68 (−1.35 to 0.04)
15	Trospium	14 (7–22)	0	−0.61 (−0.97 to −0.27)	Solifenacin	13 (6–21)	0.068	−0.66 (−0.86 to −0.46)	Solifenacin	13 (6–21)	0.06	−0.66 (−0.86 to −0.48)
16	Cizolirtine	17 (5–26)	0.04	−0.49 (−1.15 to 0.19)	Tolterodine, PFE, and BT	14 (4–24)	0.536	−0.64 (−1.05 to −0.23)	Tolterodine and estrogen	15 (5–24)	0.36	−0.63 (−1.02 to −0.24)
17	Tolterodine and PFE	17 (5–26)	0.03	−0.49 (−1.17 to 0.20)	Pregabalin and tolterodine	16 (6–25)	0.144	−0.57 (−0.93 to −0.17)	Pregabalin and tolterodine	17 (6–25)	0.15	−0.54 (−0.94 to −0.12)
18	Tolterodine	17 (12–21)	0	−0.48 (−0.62 to −0.35)	Trospium	17 (9–23)	0.004	−0.54 (−0.75 to −0.33)	Trospium	18 (10–24)	0	−0.54 (−0.75 to −0.33)
19	Oxybutynin	17 (11–22)	0	−0.48 (−0.69 to −0.28)	Propiverine	18 (9–24)	0.004	−0.52 (−0.72 to −0.30)	Propiverine	18 (10–24)	0	−0.52 (−0.72 to −0.3)
20	Propiverine	17 (9–23)	0	−0.49 (−0.82 to −0.16)	Oxybutynin	18 (11–24)	0	−0.51(−0.66 to −0.33)	Tolterodine	19 (13–23)	0	−0.5 (−0.63 to −0.38)
21	Fesoterodine	18 (11–23)	0	−0.46 (−0.70 to −0.23)	Fesoterodine	18 (10–24)	0	−0.50 (−0.67 to −0.32)	Oxybutynin	19 (11–24)	0	−0.51 (−0.67 to −0.34)
22	Bladder training	21 (3–28)	0.14	−0.30 (−2.31 to 1.74)	Tolterodine	19 (12–23)	0	−0.50 (−0.63 to −0.37)	Fesoterodine	19 (11–24)	0	−0.5 (−0.67 to −0.32)
23	Pregabalin and tolterodine	21 (9–26)	0	−0.28 (−0.80 to 0.24)	Resiniferatoxin	24 (12–27)	0.032	−0.19 (−0.70 to 0.39)	Cizolirtine	24 (12–27)	0.01	−0.24 (−0.71 to 0.22)
24	Estrogen and medroxyprogesterone	23 (9–27)	0	−0.10 (−0.83 to 0.63)	Cizolirtine	24 (11–27)	0.014	−0.24 (−0.71 to 0.25)	Resiniferatoxin	24 (13–27)	0.01	−0.19 (−0.7 to 0.36)

continued on next page

25	Placebo	24 (21–27)	0		NA	Pregabalin	24 (14– 27)	0.002	–0.18 (–0.64 to 0.31)	Pregabalin	24 (14–27)	0	–0.19 (–0.65 to 0.29)
26	Pregabalin	24 (13–27)	0	0.01 (–0.63 to 0.66)		Estrogen and medroxy-progesterone	26 (9–28)	0.084	–0.10 (–0.79 to 0.58)	Estrogen and medroxy-progesterone	26 (9–28)	0.12	–0.1 (–0.81 to 0.62)
27	Resiniferatoxin	27 (7–28)	0.19	0.60 (–1.09 to 2.28)		Placebo	26 (23–27)	0	NA	Placebo	26 (23–28)	0	NA
28	Electros-timulation	28 (23–28)	0	1.27 (–0.19 to 2.76)		Electro-stimulation	28 (25–28)	0.006	1.19 (–0.13 to 2.69)	Electro stimulation	28 (23–28)	0	1.21 (–0.3 to 2.66)

BoNTA, botulinum toxin type A; BT, bladder training; NA, not available/not applicable; NMA, network meta-analysis; PFE, pelvic floor exercises.

estimates produced by hierarchical analyses. For example, in comparison to placebo, BoNTA 150 U had a similar reduction of –1.45 (95% credible interval [CrI] –2.73 to –0.17) and –1.31 (95% CrI –2.15 to –0.55) leakage episodes per 24 hours for individual and hierarchical NMAs, respectively, though the precision increased by approximately 150% for the effect estimate derived from the hierarchical analysis. Imposing ordering constraints on increasing doses of BoNTA further increased the precision of the effect estimates, given that the estimated mean reduction in incontinence episodes for BoNTA 150 U became –1.32 (95% CrI –1.94 to –0.68), with the corresponding precision increasing by approximately 310% compared with the individual analysis. The reduction in posterior uncertainty is particularly apparent through the narrower CrIs. The synthesis of all available data in a hierarchical analysis suggested that BoNTA 200 U is the most efficacious intervention for reducing incontinent episodes with a corresponding probability of 20.58% of it being the "best" intervention overall. Imposing additional ordering constraints identified BoNTA 300 U as the most efficacious intervention with an increased probability of 83.18% of being the best intervention compared with an estimated 12.66% from the unconstrained model.

In addition, the posterior summaries for the class effects obtained from both class-based and hierarchical analyses also agree with one another (Table 3). These suggest that botulinum toxins, as a class of interventions, are the most efficacious at reducing the number of incontinent episodes per 24 hours, with an estimated mean reduction of –1.40 (95% CrI –2.14 to –0.66), –1.29 (95% CrI –2.12 to –0.56), and –1.32 (95% CrI –1.95 to –0.66) for class-based, hierarchical, and hierarchical with constraints models, respectively. Although categorizing the interventions into their respective classes for the class-based NMA increased the precision of the effect estimates compared with the individual NMA, it restricts the interpretability of the result at an individual intervention level.

Number of Patients Reporting Adverse Events

Table 4 presents the interventions ranked in order of the estimated odds of a patient reporting an adverse event. The estimated odds are comparable in both individual and hierarchical models; however, precisions in estimates derived from hierarchical analyses have substantially increased. In both sets of analyses, pregabalin is identified as the "worst" intervention for causing patients to report adverse events. The estimated odds ratio relative to placebo is 3.25 (95% CrI 1.52–7.03) for the individual NMA and 2.77 (95% CrI 1.55– 5.05) for the hierarchical NMA. Although the effect estimates derived from both models are broadly similar, there is a 67% increase in the posterior precision of the estimate obtained from the hierarchical model. This increase in precision is demonstrated through the consistently narrower CrIs. There is still considerable uncertainty in the estimated odds and consequently in the estimated ranks of the interventions, which is further highlighted by the associated 95% CrIs. Although pregabalin is ranked "worst" overall, it is ranked worst only 30% of the time in both individual and hierarchical NMAs, and thus this result should be interpreted cautiously [13].

The median of posterior distributions for class effects are comparable in both class-based and hierarchical NMAs (Table 5). In addition, both sets of analyses suggest that other drugs have the highest prevalence of patients reporting one or more adverse events, with 70.73% and 60.52% probability of these being the highest ranking intervention for causing one or more adverse events for class-based and hierarchical NMA, respectively.

Sensitivity Analysis

Sensitivity analyses suggested that changing the prior distributions of the variance parameters had very little impact on the estimated treatment effects for both the outcome measures (see Appendices C and D in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.10.006). Sensitivity to both the prior distribution for the between-study variance and the between-intervention class-specific variances showed little evidence of an impact on the overall treatment effect estimates and precision for individual and hierarchical models, respectively, thus suggesting that all sets of analyses are insensitive to the choice of the prior distributions.

Node-Splitting

There was little evidence of an inconsistency between direct and indirect estimates obtained from hierarchical NMAs as assessed by methods of node-splitting (see Appendices E and F in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.10.006). For the individual and hierarchical analysis, tolterodine, oxybutynin, BoNTA 100 U, solifenacin, and trospium demonstrate an inconsistent direct and indirect estimate for the mean reduction in incontinent episodes from baseline, relative to placebo (see Appendix E in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.10.006), although between the active interventions, there was little evidence of an inconsistency. Further investigation of the individual classes of treatments (e.g., anticholinergic drugs alone) showed little evidence of an inconsistency between direct and indirect estimates when compared with placebo. This would suggest that the potential pooling of the placebo interventions *between* classes might not be an appropriate assumption because placebo for one class of interventions could be very different from that for another class of interventions. For the adverse event outcome, however, there was very little evidence of any inconsistency between direct and indirect estimates for any set of pairwise interventions (see Appendix F in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.10.006).

Discussion

In this article, we have described and demonstrated the use of hierarchical modeling for mixed treatment NMAs–a useful methodology that can be used in clinical areas in which the available interventions are particularly extensive and the evidence base is somewhat limited both in terms of the number of trials and the number of direct comparisons [6]. With the development of MCMC methods for fitting these models implemented in WinBUGS, hierarchical NMAs are not only computationally feasible but also widely applicable to other clinical settings.

Characteristically, NMAs performed on large networks with a relatively small evidence base frequently evaluate the interventions using an individual treatment NMA, thereby presenting extremely uncertain effect estimates. Alternatively, and in the case of the OAB syndrome example, the NMAs will focus on analyzing a specific set, or class of interventions. Both methodological approaches, however, can often make it difficult to infer the most efficacious intervention, making health policy decision making difficult. Conducting an individual treatment NMA, with a limited evidence base, produces considerable uncertainty in the effect estimates and thus any inferences regarding treatment effectiveness will remain cautious. Reducing the network, by collapsing arms to compare classes of interventions, will severely hinder the ability to specifically identify the most efficacious treatment overall. For example, the class-based NMA identified botulinum toxin to be the most efficacious class of interventions

Table 3 – Classification of treatments placed in ranked order of their relative difference compared with placebo (δ and δ*) and corresponding probabilities for the mean reduction in incontinence episodes from baseline.

S. no.	Class-based NMA				Hierarchical NMA				Hierarchical NMA with constraints			
	Treatment class	Rank (95% Crl)	*p* (best) %	δ (95% Crl)	Treatment class	Rank (95% Crl)	*p* (best) %	δ* (95% Crl)	Treatment class	Rank (95% Crl)	*p* (best) %	δ* (95% Crl)
1	Botulinum toxin	1 (1–3)	89.38	−1.40 (−2.14 to −0.66)	Botulinum toxin	1 (1–4)	11.75	−1.29 (−2.12 to −0.56)	Botulinum toxin	1 (1–3)	82.6	−1.32 (−1.95 to −0.66)
2	Anticholinergic drugs and behavior therapy	2 (1–5)	7.216	−0.83 (−1.21 to −0.44)	Anticholinergic drugs and behavior therapy	3 (1–6)	3.67	−0.78 (−1.23 to −0.36)	Behavior therapy	3 (1–7)	9.93	−0.76 (−1.41 to −0.12)
3	Anticholinergic and other drugs	3 (2–6)	1.096	−0.65 (−0.95 to −0.35)	Behavior therapy	3 (1–7)	0	−0.74 (−1.5 to −0.03)	Anticholinergic drugs and behavior therapy	3 (1–5)	5.71	−0.79 (−1.15 to −0.43)
4	Behavior therapy	4 (2–7)	2.122	−0.60 (−1.11 to −0.08)	Anticholinergic and other drugs	4 (2–6)	0.536	−0.62 (−0.98 to −0.24)	Anticholinergic and other drugs	4 (2–6)	1.25	−0.59 (−1 to −0.17)
5	Anticholinergic drugs	5 (3–6)	0.004	−0.54 (−0.63 to −0.45)	Anticholinergic drugs	5 (3–6)	0	−0.53 (−0.69 to −0.38)	Anticholinergic drugs	5 (3–6)	0.05	−0.54 (−0.69 to −0.38)
6	Other drugs	6 (4–7)	0.016	−0.39 (−0.64 to −0.15)	Other drugs	6 (3–7)	0.006	−0.20 (−0.66 to 0.30)	Other drugs	6 (4–7)	0.03	−0.21 (−0.67 to 0.26)
7	Other and other drug	7 (3–9)	0.158	−0.09 (−0.77 to 0.59)	Other and other drug	7 (2–8)	0.068	−0.11 (−0.84 to 0.64)	Other and other drug	7 (2–8)	0.43	−0.1 (−0.87 to 0.67)
8	Placebo	8 (7–8)	0	NA	Neuromodulation	8 (7–8)	0	1.19 (−0.16 to 2.72)	Neuromodulation	8 (6–8)	0	1.21 (−0.31 to 2.69)
9	Neuromodulation	9 (8–9)	0.01	1.29 (−0.02 to 2.61)	Placebo	NA	NA	NA	Placebo	NA	NA	NA

NA, not available/not applicable; NMA, network meta-analysis.

Table 4 – Treatments placed in ranked order of their relative odds compared with placebo (exp(φ) and exp(φ^*)) and corresponding probabilities for the number of patients reporting one or more adverse events.

	Individual NMA				Hierarchical NMA			
	Treatment	Rank worst (95% CrI)	p (worst) (%)	exp(φ) (95% CrI)	Treatment	Rank worst (95% CrI)	p (worst) (%)	exp(φ^*) (95% CrI)
1	Pregabalin	2 (1–9)	30.41	3.25 (1.52–7.03)	Pregabalin	3 (1–9)	27.66	2.77 (1.55–5.05)
2	Duloxetine	3 (1–11)	26.06	3.03 (1.32–7.03)	Duloxetine	3 (1–10)	23.27	2.66 (1.45–4.95)
3	Oxybutynin	3 (1–7)	14.09	3.06 (1.97–4.85)	Oxybutynin	3 (1–8)	16.09	2.61 (1.77–3.94)
4	Propiverine	4 (1–9)	6.14	2.61 (1.62–4.22)	Cizolirtine	4 (1–13)	14.29	2.36 (1.16–4.81)
5	Cizolirtine	5 (1–15)	15.88	2.39 (–0.81 to 7.17)	Propiverine	5 (1–10)	7.08	2.29 (1.54–3.53)
6	Darifenacin	5 (1–10)	5.21	2.41 (1.36–4.31)	Darifenacin	6 (1–11)	4.19	2.12 (1.36–3.39)
7	Pregabalin and tolterodine	7 (2–14)	1.08	1.86 (0.90–3.82)	UK-369,003	7 (1–14)	3.51	1.92 (0.96–3.60)
8	Solifenacin	8 (4–11)	0.05	1.77 (1.26–2.53)	Pregabalin and tolterodine	8 (1–15)	3.15	1.77 (0.9–3.49)
9	Imidafenacin	9 (4–14)	0.14	1.60 (0.93–2.75)	Solifenacin	8 (4–12)	0.2	1.77 (1.31–2.44)
10	UK-369,003	11 (3–16)	0.83	1.34 (0.56–3.22)	Imidafenacin	9 (4–14)	0.2	1.62 (1.06–2.48)
11	Fesoterodine	11 (5–15)	0.1	1.31 (0.69–2.46)	Fesoterodine	10 (4–15)	0.2	1.49 (0.92–2.39)
12	Trospium	11 (7–15)	0.001	1.27 (0.88–1.82)	Trospium	11 (7–14)	0.01	1.36 (0.99–1.88)
13	Tolterodine	12 (9–14)	0	1.17 (0.92–1.49)	Tolterodine	13 (10–14)	0	1.21 (0.97–1.52)
14	Tolterodine and BT	14 (7–16)	0.02	1.0 (0.52–2.03)	Tolterodine and BT	14 (6–16)	0.15	1.04 (0.54–2.03)
15	Placebo	14 (12–15)	0	NA	Placebo	14 (13–15)	0	NA
16	BoNTA 200 U	16 (12–16)	0.002	0.50 (0.21–1.19)	BoNTA 200 U	16 (13–16)	0.004	0.50 (0.21–1.16)
17	BT	17 (17–17)	0	0.001 (0.00–0.1)	BT	17 (17–17)	0	0.02 (0.00–0.21)

BT, bladder training; NMA, network meta-analysis.

with a probability of 89.38%, though it is unclear which specific BoNTA intervention, that is, dose, is the most efficacious overall.

In the current literature, use of the term "hierarchical NMA" is intermittently used to describe what is commonly known as a "random effects NMA" with variance components at two levels in the model, one at the within-study level and one at the between-study (within intervention) level. In this article, we demonstrate a third level in the model, accounting for an additional variance component between interventions within a class. Adding an additional level to the model changes the assumption of exchangeability and, consequently, the degree of shrinkage [1]. For this reason, there is a notable change in the estimated mean treatment effects and precision of the hierarchical NMA compared with that of the individual treatment NMA.

In comparison to the above methods, use of a hierarchical model as described in this article has several advantages. Principally, there is a substantial increase in the precision surrounding the effect estimates, and this was particularly apparent for the interventions for which there are few trials and a limited number of direct comparisons between other active interventions [6]. In addition, the hierarchical model maintains the interpretability of the effect estimates at an individual intervention level.

Nevertheless, the hierarchical models make a fundamental assumption that the intervention effects, within treatment classes, are exchangeable, and a judgment of appropriateness of such an assumption has to be made [1]. If this assumption does not hold for every class of interventions, use of a hierarchical model will introduce inappropriate results; thus, it is important for researchers to classify treatments into clinically plausible classes. Of course, the treatments do not have to be classified if there is no reason to do so. A further limitation of the hierarchical model is the subjective classification of the interventions when there is potential treatment overlap. For example, in the OAB syndrome case, trospium and physiotherapy as a combination intervention will overlap with both anticholinergic and behavior therapy classes. Moreover, the combination of interventions individually estimated in the NMA could be modeled as the sum of individual components, with the potential to incorporate a synergistic or subadditive interaction between the interventions [19]. Furthermore, the number of interventions and trials within each class can vary substantially. Therefore, in particular classes in which there are few interventions and a small evidence base, the estimates will remain fairly uncertain. In situations such as these, the impact of the prior distributions on the variance parameters could be substantial, and use of extensive sensitivity analyses would be crucial [1,23].

Extending the hierarchical model to incorporate ordering constraints [9,10] on the BoNTA interventions for the OAB syndrome example resulted in the highest dose, BoNTA 300 U, to consistently be the most effective dose and therefore all other BoNTA interventions have a 0% probability of being the "best" intervention overall. In other examples, including ordering constraints in this way, that is, allowing the treatment effects of higher doses to be greater than or *equal* to those of lower doses, allows lower doses of interventions to have an equal probability of being the best intervention. Introducing these constraints for the mean reduction in incontinence episodes resulted in an estimate of –1.44 (95% CrI –2.09 to –0.78) for BoNTA 300 U compared with an estimate of –1.27 (95% CrI –2.12 to –0.51) for the unconstrained hierarchical model. Thus, assuming that larger doses of an intervention have a greater or equal effect to that of lower doses does not alter the effect estimates to any noticeable extent. This approach does, however, reduce the uncertainty in the effect estimates, the estimated ranking of

the interventions, and the probability that each of the interventions is the most effective, and consequently aids decision making.

The hierarchical model can be further extended to fit a multivariate hierarchical NMA [28] in which all outcomes are measured simultaneously [29]. This method of analysis can help ameliorate the potential effect of outcome reporting bias in trials that fail to report all the outcome measures of interest. This approach estimates a correlation between the outcomes to predict a value for the missing data points conditional on the outcome measures already reported and the model [30].

To further investigate the inconsistency detected between the direct and indirect evidence for the continuous outcome, exploratory analyses could investigate the association of baseline risk and treatment effect. Baseline risk represents the average response of a patient under the control group (e.g., placebo). If inconsistency is a result of pooling placebo interventions, incorporating baseline risk in to the model will explain some of the heterogeneity between studies [31].

In summary, we have shown that the use of hierarchical modeling, in NMAs, can increase the precision of intervention estimates, without hindering the interpretability of individual treatments. As demonstrated by the OAB syndrome example, borrowing strength within the classes of treatments reduced the uncertainty in individual estimates, yet estimated relative effects were still comparable with results obtained from individual analyses.

Source of financial support: This article presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the National Health Service, the NIHR, or the Department of Health. Rhiannon Owen is funded by an NIHR Doctoral Research Fellowship (grant no. NI-DRF-2013-06-062), and Keith Abrams is partially supported as an NIHR Senior Investigator (grant no. NF-SI-0512-10159).

Supplemental Materials

Supplemental material accompanying this article can be found in the online version at http://dx.doi.org/10.1016/j.jval.2014.10.006 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

REFERENCES

[1] Spiegelhalter D, Abrams KR, Myles J. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. England: John Wiley, 2003.
[2] Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. Value Health 2008;11:956–64.
[3] Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005;331:897–900.
[4] Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. Stat Med 2013;32:2935–49.
[5] Dakin HA, Welton NJ, Ades AE, et al. Mixed treatment comparison of repeated measurements of a continuous endpoint: an example using topical treatments for primary open-angle glaucoma and ocular hypertension. Stat Med 2011;30:2511–35.
[6] Warren FC, Abrams KR, Sutton AJ. Hierarchical Bayesian network meta-analysis models to address sparsity of events and differing treatment classifications with regard to adverse outcomes. Stat Med 2014;33:2449–66.
[7] Haas DM, Caldwell DM, Kirkpatrick P, et al. Tocolytic therapy for preterm delivery: systematic review and network meta-analysis. BMJ 2012;345:e6226.
[8] Soares MO, Dumville JC, Ades AE, Welton NJ. Treatment comparisons for decision making: facing the problems of sparse and few data. J R Stat Soc A Stat Soc 2014;177:259–79.
[9] Demiris N, Lunn D, Sharples LD. Survival extrapolation using the poly-Weibull model. Stat Methods Med Res 2011 Nov 21. http://dx.doi.org/10.1177/0962280211419645: [Epub ahead of print].
[10] Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. Stat Med 2000;19:3359–76.
[11] National Institute for Health and Care Excellence. Urinary incontinence: the management of urinary incontinence in women. 2013. Available from: http://publications.nice.org.uk/urinary-incontinence-cg40. [Accessed January 5, 2013].
[12] Higgins JPT, Whitehead A. Borrowing strength from external trials in a meta-analysis. Stat Med 1996;15:2733–49.
[13] Dias S, Welton NJ, Caldwell D. Evidence synthesis for decision making in healthcare. Stat Med 2010;29:932–44.
[14] Welton NJ, Sutton AJ, Cooper NJ, et al. Evidence Synthesis for Decision Making in Healthcare. Chichester, UK: John Wiley & Sons, 2012.
[15] Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc B Stat Methodol 2002;64:583–639.
[16] Chapple CR, Khullar V, Gabriel Z, et al. The effects of antimuscarinic treatments in overactive bladder: an update of a systematic review and meta-analysis. Eur Urol 2008;54:543–62.
[17] Novara G, Galfano A, Secco S, et al. A systematic review and meta-analysis of randomized controlled trials with antimuscarinic drugs for overactive bladder. Eur Urol 2008;54:740–64.
[18] Anger J, Weinberg A, Suttorp M, et al. Outcomes of intravesical botulinum toxin for idiopathic overactive bladder symptoms: a systematic review of the literature. J Urol 2010;183:2258–64.
[19] Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. Am J Epidemiol 2009;169:1158–65.
[20] Ades AE, Mavranezouli I, Dias S, et al. Network meta-analysis with competing risk outcomes. Value Health 2010;13:976–83.
[21] Spiegelhalter DJ, Thomas A, Best N, Lunn D. WinBUGS user manual version 1.4. January 2003. Upgraded to Version 1.4.3. Available from: http://www.mrc-bsu.cam.ac.uk/bugs. [Accessed January 20, 2013].
[22] Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 1998;7:434–55.
[23] Lambert PC, Sutton AJ, Burton PR, et al. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med 2005;24:2401–28.
[24] Ihaka R, Gentleman R. A language for data analysis and graphics. J Comput Graph Stat 1996;5:299–314.
[25] Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. J Stat Softw 2005;12:1–16.
[26] van Valkenhoef G, Lu G, de Brock B, et al. Automating network meta-analysis. Res Synth Methods 2012;3:285–99.
[27] Kanters S, Mills EJ, Thorlund K, et al. Antiretroviral therapy for initial human immunodeficiency virus/AIDS treatment: critical appraisal of the evidence from over 100 randomized trials and 400 systematic reviews and meta-analyses. Clin Microbiol Infect 2014;20:114–22.
[28] Hong H, Carlin BP, Shamliyan TA, et al. Comparing Bayesian and frequentist approaches for multiple outcome mixed treatment comparisons. Med Decis Making 2013;33:702–14.
[29] White I, Barrett J, Jackson D, Higgins J. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. Res Synth Meth 2012;3:111–25.
[30] Nam I, Mengersen K, Garthwaite P. Multivariate meta-analysis. Stat Med 2003;22:2309–33.
[31] Achana FA, Cooper NJ, Dias S, et al. Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. Stat Med 2013;32:752–71.