Technology Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence

Title: Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): Clinical effectiveness systematic review and economic model

Produced by ScHARR, The University of Sheffield

Authors

Rachel Archer, Research Fellow, ScHARR
Paul Tappenden, Reader in Health Economic Modelling, ScHARR
Shijie Ren, Research Associate, ScHARR
Marrissa Martyn-St James, Research Associate, ScHARR
Rebecca Harvey, Research Associate, ScHARR
Hasan Basarir, Research Fellow, ScHARR
John Stevens, Reader in Decision Science, ScHARR
Christopher Carroll, Reader in Health Technology Assessment, ScHARR
Anna Cantrell, Information Specialist, ScHARR
Alan Lobo, Consultant Gastroenterologist, Sheffield Teaching Hospitals NHS Foundation Trust
Sami Hoque, Consultant Gastroenterologist, Barts Health NHS Trust

Correspondence to

Dr Rachel Archer
Research Fellow
Health Economics and Decision Science
School of Health and Related Research
University of Sheffield
Regent Court, 30 Regent Street, Sheffield S1 4DA
Email: r.archer@sheffield.ac.uk

Date completed 24th June 2014

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/51/01

Declared competing interests of the authors

None

Declared competing interests of the clinical advisors/peer reviewers

None

Acknowledgements

We would like to thank Dr Anthony Akobeng (Central Manchester University Hospitals NHS Foundation Trust) and Mr Steven Brown (Sheffield Teaching Hospitals NHS Foundation Trust) for providing clinical advice over the course of this assessment. Dr Stuart Bloom (University College London Hospitals NHS Foundation Trust) also provided clinical advice during the early stages of this assessment but had no subsequent involvement due to the identification of a potential conflict of interest. We would also like to thank Dr Marcus Harbord (Chelsea and Westminster Hospital NHS Foundation Trust), Mrs Sarah Davis (ScHARR, University of Sheffield) and Dr Eva Kaltenthaler (ScHARR, University of Sheffield) for commenting on the draft assessment report. Thanks also to Gill Rooney (ScHARR, University of Sheffield) for providing project administration support. We would also like to thank MSD and AbbVie for providing unpublished data to inform the network meta-analyses undertaken as part of this assessment.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Archer R, Tappenden P, Ren S, Martyn-St James M, Harvey R, Basirir H, Stevens J, Carroll C, Cantrell A, Lobo A, Hoque S. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. *Technology Assessment Report: Final report to the National Institute for Health and Care Excellence*, 2014.

Contributions of authors

Rachel Archer acted as Principal Investigator for this assessment. Anna Cantrell carried out the electronic searches. Rachel Archer, Marrissa Martyn-St James and Christopher Carroll undertook the clinical effectiveness systematic review. Paul Tappenden and Hasan Basarir undertook the health economic review and developed the Assessment Group model. Shijie Ren, Rebecca Harvey and John Stevens conducted the network meta-analyses. Professor Alan Lobo (Sheffield Teaching Hospitals NHS Foundation Trust) and Dr Sami Hoque (Barts Health NHS Trust) provided clinical advice and commented on the draft assessment report.

About ScHARR

The School of Health and Related Research (ScHARR) is one of the nine departments that comprise the Faculty of Medicine, Dentistry and Health at the University of Sheffield. ScHARR specialises in health services and public health research, and the application of health economics and decision science to the development of health services and the improvement of public health.

The ScHARR Technology Assessment Group (ScHARR-TAG) synthesises research on the clinical effectiveness and cost-effectiveness of healthcare interventions for the NIHR Health Technology Assessment Programme on behalf of a range of policy makers, including the National Institute for Health and Care Excellence (NICE). ScHARR-TAG is part of a wider collaboration of a number of units from other regions including Health Economics Research Unit and Health Services Research Unit, University of Aberdeen; Southampton Health Technology Assessment Centre (SHTAC), University of Southampton; Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; Peninsular Technology Assessment Group (PenTAG), University of Exeter; the NHS Centre for Reviews and Dissemination, University of York; Warwick Evidence, University of Warwick; the BMJ Technology Assessment Group (BMJ-TAG), BMJ Evidence Centre and Kleijnen Systematic Reviews Ltd.

Word count: 70,000 words approx.

TABLE OF CONTENTS

1	DEFINITION OF TERMS AND LIST OF ABBREVIATIONS	1
2	EXECUTIVE SUMMARY	3
2.1	Background	3
2.2	Objectives	3
2.3	Methods	3
2.4	Results	3
2.5	Discussion	7
3	BACKGROUND	9
3.1	Description of health problem	9
3.2	Impact of health problem	10
3.3	Current service provision	13
3.4	Description of technology under assessment	16
4	DEFINITION OF THE DECISION PROBLEM	20
4.1	Decision problem	20
4.2	Overall aims and objectives of assessment	22
5.	ASSESSMENT OF CLINICAL EFFECTIVENESS	23
5.1	Methods for reviewing clinical effectiveness	23
5.2	Results	29
5.3	Discussion	154
6	ASSESSMENT OF COST-EFFECTIVENESS	157
6.1	Systematic review of existing cost-effectiveness evidence	157
6.2	*De novo* Assessment Group model	213
6.3	Budget impact analysis	243
6.4	Discussion	244
7	ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES	247
7.1	Surgery and patient choice	247
7.2	Administration route – i.v. infusions versus subcutaneous injection	247
7.3	Training for subcutaneous injections	247
7.4	Screening for TB and other infectious diseases (e.g. hepatitis B)	247
7.5	Off-license use in children for golimumab and adalimumab	247
8	DISCUSSION	248

8.1	Statement of principal findings	248
8.2	Strengths and limitations of the assessment	251
8.3	Uncertainties	252
9	CONCLUSIONS	253
9.1	Implications for service provision	254
9.2	Suggested research priorities	254
10	APPENDICES	255
11	REFERENCES	411

LIST OF TABLES

Table 1	Features of the Truelove and Witts' Severity Index	11
Table 2	Features of the Mayo Score	12
Table 3	Features of the Paediatric Ulcerative Colitis Activity Index (PUCAI)	12
Table 4	Acquisition costs associated with infliximab, adalimumab and golimumab	19
Table 5	Licenced dose comparisons and assessment time point for included RCTs	31
Table 6	Trial design characteristics of included clinical effectiveness studies in adults	36
Table 7	Trial design characteristics of included paediatric population clinical effectiveness studies	48
Table 8	Population characteristics of included clinical effectiveness studies in adults	53
Table 9	Population characteristics of included paediatric population clinical effectiveness studies	61
Table 10	Definitions of clinical response and remission in adult population RCTs included in the clinical effectiveness systematic review	63
Table 11	Summarised clinical response and remission data from RCTs in adults	67
Table 12	Definitions of clinical response and remission in included paediatric population RCT	87
Table 13	Base case – probabilities of being in each category for the induction phase	96
Table 14	Base case – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response	99
Table 15	Base case – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission	102

Table 16	Base case – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response	105
Table 17	Base case – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission	108
Table 18	Sensitivity analysis 1 – probabilities of being in each category for the induction phase	111
Table 19	Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response	114
Table 20	Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission	117
Table 21	Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response	120
Table 22	Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission	123
Table 23	Sensitivity analysis 2 – probabilities of being in each category for the induction phase	126
Table 24	Sensitivity analysis 2 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response	128
Table 25	Sensitivity analysis 2 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission	131
Table 26	Sensitivity analysis 2 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response	134
Table 27	Sensitivity Analysis 2 – Probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission	137
Table 28	Sensitivity Analysis 3 – Probabilities of being in each category for the induction phase	140
Table 29	Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response	142
Table 30	Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission	145
Table 31	Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response	148
Table 32	Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission	151
Table 33	ACR20/50/70 responders at week 30 in PLANET-RA (all randomised population)	153
Table 34	Summary of search results for existing economic evaluations	159
Table 35	Summary table of included published studies	160
Table 36	Headline cost-effectiveness results presented by Park *et al*	163
Table 37	Headline cost-effectiveness results presented by Tsai *et al*	167
Table 38	Headline cost-effectiveness results reported by Xie *et al*	171
Table 39	Summary table of evidence sources used to inform the MSD model parameters	176

Table 40	Manufacturer's NMA results – induction treatment	178
Table 41	Manufacturer's NMA results – maintenance treatment	178
Table 42	Transition matrix for adalimumab	179
Table 43	Health utility values used in the MSD models	181
Table 44	Background therapies resource use and costs used in MSD golimumab model	182
Table 45	Background therapies resource use and costs used in MSD infliximab model	183
Table 46	Biologic treatment resource use and costs used in MSD models	184
Table 47	Other health state costs per 2-month cycle – biologic treatments	185
Table 48	Other health state costs 2-month cycle – colectomy group	186
Table 49	Model results from golimumab submission (including PAS)	187
Table 50	Model results from infliximab submission (including PAS)	187
Table 51	Adherence of the MSD model to the NICE Reference Case	190
Table 52	Model results from golimumab submission (excluding PAS)	196
Table 53	Model results from infliximab submission (excluding PAS)	196
Table 54	Summary table of evidence sources used to inform the AbbVie model parameters	200
Table 55	Adalimumab group - 2-week transition probabilities for pre-surgery states by time interval	202
Table 56	Conventional management group - 2-week transition probabilities for pre-surgery states by time interval	203
Table 57	Transition matrix for surgery and post-surgical states (all time periods, both treatment groups)	203
Table 58	Health utilities assumed in the AbbVie model	204
Table 59	Drug resource cost parameters used in the AbbVie model	204
Table 60	Other health state costs used in the AbbVie model	205
Table 61	Model results obtained from the AbbVie model – base case analysis	206
Table 62	Model results obtained from the AbbVie model – anti-TNF-α naïve subgroup	206
Table 63	Adherence of the AbbVie model to the NICE Reference Case	209
Table 64	Comparison of observed and predicted induction outcomes	212
Table 65	Summary of evidence sources used to inform the model's parameter values	219
Table 66	Transition probabilities applied in the Assessment Group model	220
Table 67	Summary of studies reporting on long-term colectomy rates in UC population	221
Table 68	Surgery-related complication frequency and treatment approach	223
Table 69	EQ-5D utilities search results	224
Table 70	Studies included in the systematic review of utility values	226

Table 71	Utility vectors for all states and treatment options	227
Table 72	Dosing regimens and costs for biologic options	228
Table 73	Conventional drug regimen costs per induction/maintenance cycle	230
Table 74	UC resource use per year	231
Table 75	Relative risks of hospitalisation for infliximab, adalimumab and golimumab	231
Table 76	Unit costs for UC health state resource use, surgery and complications	232
Table 77	Probabilistic cost-effectiveness results, base case analysis, adult patients in whom colectomy is an option (medical and surgical treatments)	235
Table 78	Probabilistic cost-effectiveness results, base case analysis, adult patients in whom colectomy is not an option (medical treatments only)	237
Table 79	Probabilistic cost-effectiveness results, base case analysis, paediatric patients in whom colectomy is an option (medical and surgical treatments)	238
Table 80	Probabilistic cost-effectiveness results, base case analysis, paediatric patients in whom colectomy is not an option (medical treatments only)	239
Table 81	Results of probabilistic NMA sensitivity analyses	240
Table 82	Head-to-head analysis – adult population, infliximab versus conventional management	241
Table 83	Other deterministic sensitivity analyses	242
Table 84	Estimated absolute and net budgetary impact of introducing biologics for the treatment of moderate to severe UC in England and Wales	244
Table 85	Summary of key differences between the Assessment Group model and the manufacturers' models	245

LIST OF FIGURES

Figure 1	Treatment pathway for moderate to severe UC	15
Figure 2	Flow diagram of study inclusion (adapted from PRISMA)	30
Figure 3	Risk of bias summary - Review authors' judgements about each risk of bias item for each included study	50
Figure 4	Risk of bias graph - Review authors' judgements about each risk of bias item presented as percentages across all included studies.	51
Figure 5	Forest plot of comparison: Colectomy - adults, outcome: Colectomy - induction licenced dose	73
Figure 6	Forest plot of comparison: Colectomy - adults, outcome: Colectomy - maintenance licenced dose	73
Figure 7	Forest plot of comparison: Ostomy - adults, outcome: Ostomy - maintenance licenced dose	74
Figure 8	Forest plot of comparison: Serious infections - adults, outcome: Serious infections - induction licenced dose	80

Figure 9	Forest plot of comparison: Serious infections - adults, outcome: Serious infections - maintenance licenced dose	81
Figure 10	Forest plot of comparison: Injection site reactions - adults, outcome: Injection site reactions (ADA and GOL only) - induction licenced dose	83
Figure 11	Forest plot of comparison: Injection site reactions - adults, outcome: Injection site reactions (ADA and GOL only) - maintenance licenced dose	84
Figure 12	Forest plot of comparison: Colectomy - children, outcome: Infliximab maintenance	89
Figure 13	Base case – network of evidence for the induction phase	94
Figure 14	Base case - comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))	95
Figure 15	Base case – ranking probability histograms for the induction phase	96
Figure 16	Base case – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response	97
Figure 17	Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~HN(0,0.32^2))	98
Figure 18	Base case – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response	98
Figure 19	Base case – network of evidence for the maintenance phase at 8-32 weeks for patients starting in remission	100
Figure 20	Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))	101
Figure 21	Base case – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission	101
Figure 22	Base case – network of evidence for the maintenance phase at 32-52 weeks for patients starting in response	103
Figure 23	Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))	104
Figure 24	Base case – Ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response	104
Figure 25	Base case – network of evidence for the maintenance phase at 32-52 weeks for patients starting in remission	106
Figure 26	Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))	107
Figure 27	Base case – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission	107
Figure 28	Sensitivity analysis 1 – network of evidence for the induction phase	109
Figure 29	Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))	110

Figure 30	Sensitivity analysis 1 – ranking probability histograms for the induction phase	110
Figure 31	Sensitivity analysis 1 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response	112
Figure 32	Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32weeks for patients starting in response (SD~HN(0,0.32^2))	113
Figure 33	Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response	113
Figure 34	Sensitivity analysis 1 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in remission	115
Figure 35	Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))	116
Figure 36	Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission	116
Figure 37	Sensitivity analysis 1 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in response	118
Figure 38	Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))	119
Figure 39	Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response	119
Figure 40	Sensitivity analysis 1 – network of evidence for the maintenance phase at 32-52weeks for patients starting in remission	121
Figure 41	Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))	122
Figure 42	Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission	122
Figure 43	Sensitivity analysis 2 – network of evidence for the induction phase	124
Figure 44	Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))	125
Figure 45	Sensitivity analysis 2 – ranking probability histograms for the induction phase	125
Figure 46	Sensitivity analysis 2 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response	126
Figure 47	Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~HN(0,0.32^2))	127
Figure 48	Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response	128
Figure 49	Sensitivity analysis 2 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in remission	129

Figure 50	Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))	130
Figure 51	Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission	130
Figure 52	Sensitivity analysis 2 – network of evidence for the maintenance phase for patients starting in response at 32-52 weeks	132
Figure 53	Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))	133
Figure 54	Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response	133
Figure 55	Sensitivity analysis 2 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in remission	135
Figure 56	Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))	136
Figure 57	Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission	136
Figure 58	Sensitivity analysis 3 – network of evidence for the induction phase	138
Figure 59	Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))	139
Figure 60	Sensitivity analysis 3 – ranking probability histograms for the induction phase	139
Figure 61	Sensitivity analysis 3 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response	140
Figure 62	Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32weeks for patients starting in response (SD~HN(0,0.32^2))	141
Figure 63	Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response	142
Figure 64	Sensitivity analysis 3 – network of evidence for maintenance phase at 8-32 weeks for patients starting in remission	143
Figure 65	Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))	144
Figure 66	Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission	144
Figure 67	Sensitivity analysis 3 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in response	146
Figure 68	Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))	147
Figure 69	Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response	147

Figure 70	Sensitivity analysis 3 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in remission	149
Figure 71	Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))	150
Figure 72	Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission	150
Figure 73	Study selection results for review of economic evaluations	159
Figure 74	Model diagram presented by Park *et al*	161
Figure 75	Model diagram presented by Tsai *et al*	165
Figure 76	Treatment sequences evaluated by Xie *et al*	168
Figure 77	Model diagram presented by Xie *et al*	169
Figure 78	MSD model structure (re-drawn by Assessment Group)	175
Figure 79	MSD model Markov component	175
Figure 80	Evidence network for induction therapy	177
Figure 81	Evidence network for maintenance therapy	177
Figure 82	Cost-effectiveness plane from MSD model	188
Figure 83	Incremental CEACs from MSD model	189
Figure 84	Proportion of patients in post-colectomy states over time (excluding mortality)	192
Figure 85	AbbVie model structure	198
Figure 86	Cost-effectiveness plane reported by AbbVie – base case analysis	207
Figure 87	CEAC reported by AbbVie – base case analysis	208
Figure 88	Assessment Group model structure (induction and maintenance phases)	217
Figure 89	Study selection results	224
Figure 90	CEACs, base case analysis, adult patients in whom colectomy is an option (medical and surgical treatments)	236
Figure 91	CEACs, base case analysis, adult patients in whom colectomy is not an option (medical treatments only)	237
Figure 92	CEACs, base case analysis, paediatric patients in whom colectomy is an option (medical and surgical treatments)	238
Figure 93	CEACs, base case analysis, paediatric patients in whom colectomy is not an option (medical and surgical treatments)	239
Box 1:	Inclusion and exclusion criteria for review of cost-effectiveness studies	158
Box 2:	Main problems and concerns relating to the MSD model	189
Box 3:	Main problems and concerns relating to the AbbVie model	208
Box 4:	Scope of the Assessment Group economic analysis	213
Box 5:	Sensitivity/scenario analyses undertaken using the Assessment Group model	234

1. DEFINITION OF TERMS AND LIST OF ABBREVIATIONS

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context, but a glossary is provided for the non-specialist reader.

ACR	American College of Rheumatology
ADA	Adalimumab
AE	Adverse event
AS	Ankylosing spondylitis
5-ASA	5-aminosalicylates
AZA	Azathioprine
CDAI	Clinical disease activity index
CHMP	Committee for Human Medicinal Products
CI	Confidence interval
CrI	Credible interval
CRP	C-reactive protein
CS	Corticosteroids
DAS	Disease activity score
EMA	European Medicines Agency
EOW	Every other week
EPAR	European Public Assessment Report
EULAR	European League Against Rheumatism
EW	Every week
GOL	Golimumab
IBD	Inflammatory bowel disease
IBDQ	Inflammatory bowel disease questionnaire
IFX	Infliximab
IMM	Immunomodulators
IPAA	Ileal pouch anal anastomosis
ITT	Intention to treat
i.v.	Intravenous
LOCF	Last observation carried forward
LTE	Long term extension
6-MP	6-mercaptopurine
MP	Mercaptopurine
NHS	National Health Service

NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
OL	Open label
OLE	Open label extension
q8w	Every 8 weeks
q12w	Every 12 weeks
PBO	Placebo
PY	Patient years
RA	Rheumatoid arthritis
SAE	Serious adverse event
s.c.	Subcutaneous
SD	Standard deviation
SDAI	Simplified disease activity index
SF-36	Short form-36
SPC	Summary of Product Characteristics
TB	Tuberculosis
TNF-α	Tumour necrosis factor- α
UC	Ulcerative colitis

2. EXECUTIVE SUMMARY

2.1 Background

Ulcerative colitis (UC) is recognised as the most common form of inflammatory bowel disease (IBD) in the United Kingdom. Peak incidence is between 15 and 25 years of age, with a potential second peak between 55 and 65 years. Inflammation in UC typically occurs in the colon and rectum. Symptoms include the development of bloody diarrhoea with or without mucus, abdominal pain, weight loss, fatigue, and an urgent need to defecate. UC can have a substantial impact on the health-related quality of life of patients on account of the young age of disease onset for some patients, the severity of symptoms, and the likelihood of relapse. The burden of UC for the NHS is substantial.

2.2 Objectives

The aim of this assessment is to assess the clinical effectiveness and cost-effectiveness of infliximab, adalimumab and golimumab for the treatment of patients with moderately to severely active UC after the failure of conventional therapy.

The objectives of the assessment are:

- To evaluate the clinical effectiveness of each intervention
- To examine the effect of disease duration on the clinical effectiveness of each intervention (subject to the availability of evidence)
- To evaluate the adverse effect profile of each intervention
- To evaluate the incremental cost-effectiveness of each intervention compared (i) against each other and (ii) against all comparators (including medical and surgical options)
- To estimate the expected net budget impact associated with implementing each intervention
- To identify key areas in which future research may be valuable

2.3 Methods

A systematic review of the literature including network meta-analyses (NMA) was conducted in order to evaluate the clinical effectiveness and safety of infliximab, adalimumab and golimumab in the treatment of moderately to severely active ulcerative colitis after the failure of conventional therapy. A review of the existing cost-effectiveness literature was undertaken. A *de novo* health economic model was constructed by the Assessment Group in order to evaluate the cost-effectiveness of the interventions under assessment.

2.4 Results

2.4.1 Number and quality of studies

A total of ten randomised controlled trials (RCTs) were identified in the clinical effectiveness systematic review. Five, three and two RCTs evaluated the use of infliximab, adalimumab and

golimumab respectively in the treatment of moderate to severely active UC. Nine trials related to adults and one trial was conducted in a paediatric population. All of the adult RCTs (with the exception of one trial, UC-SUCCESS) were performed against placebo. No head-to-head RCTs were identified in which the interventions of interest were assessed against each other.

The risks of bias associated with included RCTs were assessed using the Cochrane risk of bias instrument. Only three RCTs could be considered as being at overall low risk of bias (as allocation concealment, blinded outcome assessment and completeness of outcome data were all judged as low risk). It should be noted that one of the maintenance trials (PURSUIT-Maintenance) re-randomised patients who had previously responded to golimumab induction therapy in two previous trials; the extent of this potential bias on patient outcomes is unclear.

2.4.2 Summary of benefits and risks

The outcome measures pre-specified in the final NICE scope were all addressed by the included trial evidence, with the exception of rates of relapse. Clinical response and remission data based on complete Mayo scores were well reported across trials. Evidence was identified to demonstrate that patients receiving infliximab, adalimumab or golimumab were more likely to achieve clinical response and remission at induction and maintenance time points compared to patients receiving placebo. Patients in the UC-SUCCESS trial who received combination treatment with infliximab and azathioprine experienced the most favourable rates of steroid-free remission compared with infliximab and azathioprine treatment groups. Seven RCTs performed in adult populations contributed data on clinical response and remission at induction or maintenance time points to network meta-analyses.

Based on the NMA, in the induction phase, all treatments were associated with statistically significant beneficial effects relative to placebo, with the greatest effect being associated with infliximab. For patients classified as responders at the end of the induction phase, treatment effects were not statistically significant, although the greatest effect at 8-32 weeks was associated with golimumab 100mg. At 32-52 weeks, only infliximab and golimumab 50mg were associated with beneficial effects on clinical response. For patients classified as being in remission at the end of the induction phase, all treatments except for adalimumab were associated with beneficial treatment effects relative to placebo, with the greatest effect being associated with golimumab 50mg and 100mg, although the effects were not statistically significant at 8-32 weeks. At 32-52 weeks, all treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo, with the greatest effect being associated with adalimumab (the only treatment with statistically significant

effect). Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response and from remission to no response.

Sensitivity analyses were conducted by replacing ULTRA2 anti-TNF-naïve data with ULTRA2 ITT data (sensitivity analysis 1), including Suzuki *et al* (sensitivity analysis 2), and replacing ULTRA2 anti-TNF-α naïve data with ULTRA2 ITT data plus including Suzuki (sensitivity analysis 3).

Available data on hospitalisation outcomes were very limited, but suggested that outcomes may be more favourable for adalimumab-treated and infliximab-treated patients compared with placebo (with no data available from golimumab trials). Data on surgical intervention were also very sparse, with a potential inconclusive benefit for intervention-treated patients compared with placebo. No trials reported whether surgical outcomes were elective or emergency in nature. However, more data are required to demonstrate the impact of interventions on hospitalisation and surgical intervention more conclusively. Data were available from a single trial to support the use of infliximab in induction and maintenance treatment in a paediatric population.

The main safety issues highlighted in the RCT evidence appeared to be generally consistent with those previously discussed in the respective Summary of Product Characteristics (including serious infections, malignancies and administration site reactions). Deaths occurring during and after the study period were described in some trials evaluating golimumab (PURSUIT-Maintenance) and infliximab (ACT trials), of which infection or malignancy commonly appeared to be implicated. This underlines the importance of monitoring and treating serious infections and malignancies in patients receiving immunosuppressive treatment.

Two biosimilars (Remsima and Inflectra) to Remicade were considered as part of the evidence base for infliximab within this assessment. The sponsor submission received from the manufacturers of Remsima (Celltrion) and the European Public Assessment Reports (EPARs) for Remsima and Inflectra indicated that both biosimilars were approved by the European Medicines Agency (EMA) on the basis of reported similar pharmacokinetic and efficacy (demonstrated in ankylosing spondylitis and rheumatoid arthritis patients) profiles to Remicade. No further trials of Remsima or Inflectra were identified in the course of this assessment.

2.4.3 Summary of cost-effectiveness evidence

The manufacturers of adalimumab, infliximab and golimumab submitted economic models to assess the cost-effectiveness of biologic therapies versus conventional treatment. The MSD infliximab submission model indicates that the estimated ICER for infliximab versus standard non-biologic

treatment (colectomy) is £37,682 per QALY gained. The MSD golimumab submission reports an estimated ICER of £27,322 per QALY gained. The AbbVie submission reports a base case ICER of £34,590 per QALY gained. The Assessment Group identified several problems with these models. In particular, none of the models included all relevant treatment options specified in the final NICE scope and each model adopted a short time horizon (10-years). The Assessment Group does not believe that the cost-effectiveness evidence submitted by either manufacturer represents a sufficient basis for informing decision-making.

In order to address the problems identified within the manufacturers' submitted economic models, the Assessment Group developed a *de novo* cost-effectiveness model to assess infliximab, adalimumab, golimumab, conventional non-biologic treatments and elective surgery within the moderate to severe UC population over a lifetime horizon. Underpinning the Assessment Group model is a series of network meta-analyses which synthesise all relevant evidence relating to infliximab, adalimumab, golimumab and conventional non-biologic therapies in the induction and maintenance settings.

The base case analysis of the Assessment Group model suggests that colectomy is expected to produce 14.72 QALYs at a cost of approximately £41,900 over the patient's remaining lifetime. All medical options are expected to produce substantially fewer QALYs at a greater cost than colectomy; hence colectomy is expected to dominate infliximab, adalimumab, golimumab and conventional non-biologic treatments. For some patients, elective colectomy may not be considered an acceptable or preferable option; in circumstances whereby only drug options are considered acceptable, the Assessment Group model suggests that infliximab and golimumab are expected to be ruled out due to dominance, whilst the incremental cost-effectiveness of adalimumab versus conventional non-biologic treatment is expected to be approximately £50,600 per QALY gained.

A separate economic analysis of infliximab, conventional non-biologic treatments and colectomy was undertaken within a paediatric population (mean age=15 years). Where colectomy is an acceptable treatment option, the economic analysis suggests that this option is expected to dominate infliximab and conventional non-biologic treatments. Where colectomy is not an acceptable option, the economic analysis suggests that the incremental cost-effectiveness of infliximab versus conventional treatments is approximately £68,400 per QALY gained. This analysis is however based on adult efficacy evidence; hence it should be interpreted with some degree of caution.

A number of sensitivity analyses were undertaken using the Assessment Group model. These suggested that the results of the economic analysis are largely insensitive to changes in the model assumptions, except for scenarios in which the post-surgery utility value is altered. When utility scores from Swinburn *et al*. are used in the model (rather than those reported by Woehl *et al*),

colectomy produces the lowest QALY gain and conventional management and golimumab are ruled out due to extended dominance. Within this scenario, the incremental cost-effectiveness of adalimumab versus elective colectomy is estimated to be £80,315 per QALY gained, whilst the incremental cost-effectiveness of infliximab versus adalimumab is estimated to be £179,374 per QALY gained. Whilst these results are very different to the Assessment Group's preferred base case analysis, the economic conclusions that should be drawn from this sensitivity analysis are not.

2.5 Discussion

2.5.1 Strengths, limitations of the analyses and uncertainties

The systematic review was based on rigorous methods, with comprehensive searches for evidence, a good level of consistency between reviewers in study selection and double checking of data extraction. Clinical response and remission data were well reported across included trials and study authors were consistent in their use of the complete Mayo score, which aided the comparison of trials. Network meta-analyses were performed to permit a comparison of the efficacy of interventions in terms of clinical response and remission.

The Assessment Group's economic analysis has a number of strengths:

- The treatment pathway represented within the model was based on considerable expert opinion from several leading UC experts.
- The Assessment Group model is underpinned by a complex network meta-analysis across all drug options thereby synthesising relevant efficacy outcomes data within a single network of evidence.
- The model generally adheres to NICE's Reference Case and fully addresses the decision problem set out in the final NICE scope.
- Where appropriate and possible, systematic search methods have been used to identify, select and use evidence to inform the model's parameters (efficacy, HRQoL and colectomy rates).
- The Assessment Group has undertaken extensive sensitivity analyses to examine the impact of alternative assumptions and sources of evidence on the robustness of the results of the model.

The Assessment Group model is also subject to a number of limitations:

- There is considerable uncertainty associated with Assessment Group's extrapolation of short-term trial data (maximum 54 weeks) to a lifetime horizon.
- The model does not consider an explicit sequential pathway of non-biologic treatments; rather during any cycle, a proportion of patients are assumed to receive 5-ASAs, immunomodulators and steroids.

- Evidence relating to complications of colectomy was identified through consideration of approaches used within previous models rather than through a full systematic review; these assumptions were however tested within the sensitivity analyses.

Key uncertainties in this assessment include:

- The optimal duration of intervention treatment in responding patients.
- The maintenance of efficacy outcomes and safety of interventions beyond the limited study lengths available.
- The maintenance of outcomes in responding patients following cessation of anti-TNF-α treatment.

2.5.2 Generalisability of the findings

The trials included in the clinical effectiveness systematic review typically excluded patients with ulcerative proctitis, patients with fulminant/acute severe disease, those with a history of or at imminent risk of bowel surgery, pregnant or lactating women, and patients with diseases of the central nervous system (e.g. demyelinating disease). Furthermore, patients with history of serious infection and/or immunodeficiency were also typically excluded, as were individuals with a history of malignancy or signs of dysplasia. Therefore, the effects of adalimumab, golimumab or infliximab in these UC populations are unknown.

3. BACKGROUND

3.1 Description of health problem

Ulcerative colitis (UC) is recognised as the most common form of inflammatory bowel disease (IBD) in the UK. The incidence of UC is approximately 10 per 100,000 population per year, whilst the prevalence of the disease is approximately 240 per 100,000 population.[1] This is typical for countries with a Westernised lifestyle.[2] Peak incidence is between 15 and 25 years of age, with a potential second peak between 55 and 65 years.[1] The majority (approximately 80%) of incident cases are reported to be of mild or moderate severity. An estimated 132,600 people in England and Wales have been diagnosed with UC.[3] It is distinct from Crohn's disease (CD), which is the other principal form of inflammatory bowel disease.[2]

UC is a chronic disease of unknown cause. It is understood that pathogenesis may result from a change in the colonic environment of a genetically susceptible person and the condition is genetically heterogeneous, having a large number of implicated genes.[4,2] Genetic screening is therefore not currently indicated for UC.[2] However, appendectomy and smoking have been linked with a reduced risk and severity of UC.[2]

Inflammation in UC typically occurs in the colon and rectum. Disease may be limited to the rectum (proctitis), may be left-sided or distal, or may be extensive (pancolitis).[4] Symptoms include the development of bloody diarrhoea with or without mucus, abdominal pain, weight loss, fatigue, and an urgent need to defecate. Extraintestinal manifestations may occur in 10% to 30% of patients at the following sites: skin, eyes, mouth, joints and liver.[2,5] Symptoms may vary according to the degree and severity of bowel inflammation.[1,2] Acute severe exacerbations of UC are characterised by the development of systemic signs of disease (e.g. high temperature, tachycardia, anaemia etc) and require admission to hospital for urgent monitoring and treatment.[4]

Diagnosis of UC is made by medical history, endoscopy and biopsy following the exclusion of potential infectious causes by stool examination.[6] These techniques permit the evaluation of relevant histological features and enable the differentiation of UC from other conditions such as CD.[2] For example, inflammation is characteristically restricted to the mucosal layer of the colon.[2] Diagnostic investigations also enable a determination of disease severity and there is evidence to indicate that severity of disease may be associated with younger age at diagnosis.[7,8] Based on the findings of diagnostic investigations, appropriate treatment can then be identified.

Colectomy by definition removes the source of inflammation in UC and is therefore associated with the relief of UC symptoms but is associated with a range of complications.[2,9] Medical treatments for

UC do not offer the possibility of cure and the disease course follows a relapsing-remitting pattern. The aim of clinical management is to induce and maintain disease remission and to avoid potential complications and the necessity for surgical intervention.[10] Selection of the appropriate therapy to induce remission of UC is determined by a number of factors, including severity and extent of disease. Evidence on prognosis indicates that, in the first decade, remission occurs in most patients and the rate of colectomy after diagnosis is low.[11] Otherwise, reported rates of colectomy among patients with UC are in the region of approximately 5% and 20%,[12,13] although this is an area of considerable uncertainty (some studies in selected populations have reported markedly higher colectomy rates e.g. Gustavvson *et al*.[14]). A range of factors have been suggested as potentially influencing the risk of relapse, including age (and age at first relapse), gender, smoking status, and number of previous relapses.[13]

3.2 Impact of health problem

3.2.1 Significance for patients

Complications of UC, depending on the severity and duration of the disease and age at onset, include severe bleeding and toxic megacolon, extraintestinal manifestations, and osteoporosis.[2] Dysplasia and bowel cancer may also develop. A meta-analysis by Jess *et al.*[15] demonstrated that UC is not associated with an increase in overall mortality. UC can have a substantial impact on the health-related quality of life (HRQoL) of patients on account of the young age of disease onset for some patients, the severity of symptoms, and the likelihood of relapse.[16,17,18,9] The risk of relapse and disease flares is increased by poor adherence to medication regimens.[19,20] Relapse and flares can be unpredictable and require further treatment, thus affecting patients' HRQoL, their ability to perform daily activities including work, and lead to increases in health care costs.[9,21,20]

3.2.2 Significance for the NHS

The burden of UC for the NHS is substantial, particularly with respect to those patients who suffer from poor disease control. A study of the costs of IBD (UC and CD) to the NHS[22] reported in 2004 found that compared with quiescent cases of IBD, disease relapse was associated with a 2–3-fold increase in costs for non-hospitalised cases and a 20-fold increase in costs for hospitalised cases.

3.2.3 Measurement of disease

A range of clinical measures are available for the assessment of disease activity in UC.[23] Of most relevance to this assessment are the modified Truelove and Witts' criteria,[24] the Mayo score[25] and the Paediatric Ulcerative Colitis Activity Index (PUCAI).[26]

Truelove and Witts' Severity Index[24]

The Truelove and Witts' severity index describes the frequency of diarrhoea and whether systemic features of illness, such as high temperature, tachycardia and anaemia, are present or absent in patients (Table 1). When the disease is active, patients are categorised as having mild, moderate or severe disease.

Table 1: Features of the Truelove and Witts' Severity Index (adapted from Ha[27]; Cooney *et al*[23])

Disease classification	Clinical features
Severe disease	Diarrhoea frequency > 6 stools/24 hours with blood
	Temperature > 37.5°C
	Tachycardia > 90 beats/min
	Anaemia (<75% of normal value)
	Erythrocyte sedimentation rate > 30 mm/hour
Moderate disease	Values ranging between mild and severe
Mild disease	Diarrhoea < 4 stools/24 hours, intermittently or non-bloody
	No fever
	No tachycardia
	Normal haemoglobin
	Erythrocyte sedimentation rate ≤ 30 mm/hour

Mayo score

The Mayo score assesses patients' disease in relation to four components: (i) stool frequency; (ii) rectal bleeding; (iii) endoscopic findings, and; (iv) physician's global assessment[25] (see Table 2). Full Mayo scores range from 0 to 12 (with scoring increasing with disease severity). The partial Mayo score, which comprises the non-endoscopic elements of the full Mayo score (i.e. stool frequency, rectal bleeding and physician's global assessment), has been reported to have reasonable correlation with the full Mayo score (Spearman's correlation coefficient ρ=0.70). Partial Mayo scores range from 0 to 9.[28]

Table 2: Features of the Mayo Score (adapted from Ha[27] Cooney *et al.*[23])

Stool frequency	
0	Normal stool frequency for patient
1	1-2 stools more than usual
2	3-4 stool more than usual
3	5 or more stools more than usual
Rectal bleeding	
0	No blood
1	Streaks of blood < 50% of time with stool
2	Obvious blood most of time with stool
3	Blood alone passed
*Endoscopic findings**	
0	Normal/inactive disease
1	Mild disease (erythema, decreased vascular pattern, mild friability)
2	Moderate disease (marked erythema, lack of vascular pattern, friability, erosions)
3	Erosions
Physician's global assessment	
0	Normal
1	Mild
2	Moderate
3	Severe

** Not included in partial Mayo score assessments*

Paediatric Ulcerative Colitis Activity Index (PUCAI)

The PUCAI was developed with the aim of providing a non-invasive assessment instrument for use in paediatric practice and is based on measures of abdominal pain, rectal bleeding, stool consistency, stool frequency, nocturnal stools and activity level (see Table 3). The tool has been described as showing good correlation with physician's global assessment (Pearson's r = 0.91, p<0.001), full Mayo scores (r = 0.95, p<0.001), and endoscopic subscores (r = 0.77, p<0.001).[26]

Table 3: Features of the Paediatric Ulcerative Colitis Activity Index (PUCAI) (adapted from Ha[27])

Variable	Points scored
Abdominal pain	
Absent	0
Able to be ignored	5
Not able to be ignored	10
Rectal bleeding	
None	0
Small amount (<50%) of stools	10
Small amount with most stools	20
Large amount (>50%) of stools	30
Stool consistency	
Formed	0
Partially formed	5
Completely loose	10
Stool frequency (in 24 hours)	
0-2	0
3-5	5
6-8	10
≥ 9	15
Nocturnal stools	
Absent	0
Present	10
Activity level	
No limitations	0
Occasional limitations	5
Severe limitations	10

3.3 Current service provision

3.3.1 Clinical Guidelines

As outlined in NICE Clinical Guideline 166 (*"Ulcerative colitis: Management in adults, children and young people"*),[3] conventional treatment options for moderately to severely active (non-systemic) UC include the use of oral or topical aminosalicylates, corticosteroids and/or immunosuppressants. Recommended conventional treatment options can vary according to the extent and location of colitis. Colectomy may be considered in the event of inadequate control of symptoms and/or poor HRQoL on conventional drug treatment.

3.3.2 Current NICE Technology Appraisal Guidance

Three NICE Technology Appraisals have previously been undertaken. Infliximab was not previously recommended by NICE for the treatment of "subacute" manifestations of moderately to severely active UC (NICE Technology Appraisal Guidance 140).[29] NICE Technology Appraisal 262 (adalimumab for the treatment of moderately to severely active ulcerative colitis) was terminated as no evidence submission was provided by the manufacturer.[30] NICE Technology Appraisal Guidance

TA163 recommended the use of infliximab as an option for the treatment of acute exacerbations of severely active UC only in patients for whom ciclosporin is contraindicated or clinically inappropriate.[31]

3.3.3 Current service cost

Cohen *et al* [32]reports estimates of the direct and indirect costs of UC within the US and Europe based on a systematic review of published cost studies. Cohen reports estimated annual per-patient direct medical costs of UC of between €8,949 and €10,395 in Europe (2008 currency values). The study authors note that hospitalisations associated with UC accounted for 41%-55% of direct medical costs. Indirect costs are also reported to be susbstantial, accounting for between 54% and 68% of total costs in Europe. The total economic burden of UC in Europe was estimated to be in the range €12.5 to €29.1billion.

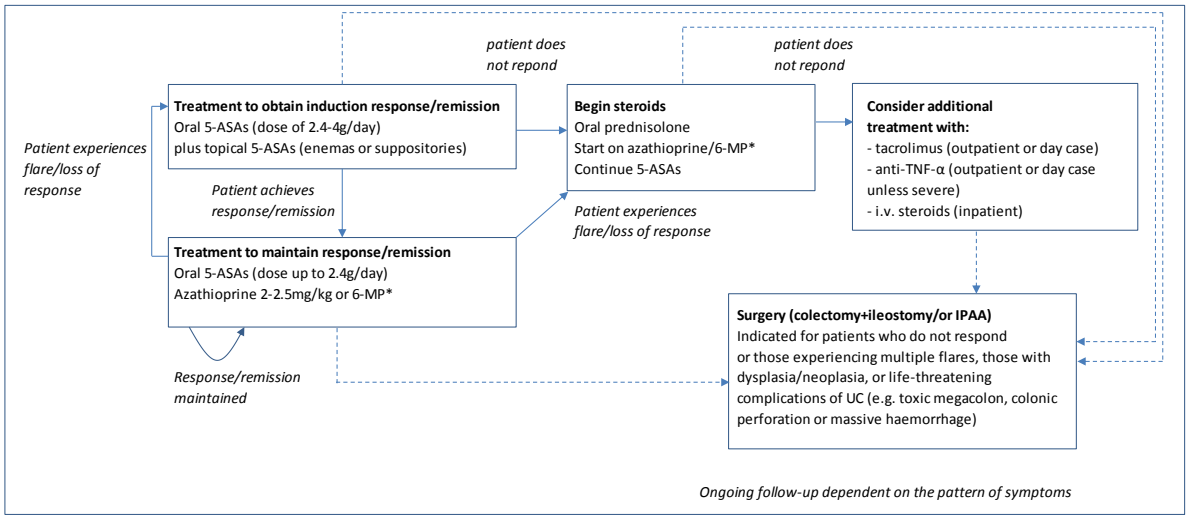
3.3.4 Variation in services and uncertainty about best practice

The optimal duration of treatment with the interventions under assessment is not yet known. The safety and efficacy of the re-administration of interventions following an interruption of treatment has not been fully established. Furthermore, the maintenance of clinical remission following the withdrawal of biologic treatment in responding patients is also unclear. There is no randomised controlled trial (RCT) evidence for the efficacy and safety of switching to a second biologic intervention in patients who are primary or secondary non-responders or in patients who are intolerant to a first biologic intervention.

3.3.5 Current treatment pathway

There does not exist a universally agreed pathway for the second-line treatment of patients with moderate to severe UC. Treatments received by patients may be influenced by the severity of symptoms, the extent and location of inflammation, clinical advice and individual patient choice. Treatments may include a combination of aminosalicylates (5-ASAs - sulfasalazine, mesalazine/mesalamine, balsalazide and olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), and thiopurines (6-mercaptopurine [6-MP] or azathioprine [AZA]), calcineurin inhibitors and surgical intervention (colectomy). The care of people with UC is usually shared between primary care and specialist gastroenterology units working in collaboration with specialist colorectal surgical units.[3] Figure 1 presents a simplified pathway of the main types of treatments used for the management of patients with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids and 6-mercaptopurine (6-MP) or azathioprine, or who are intolerant to or have medical contraindications against such therapies.

Figure 1: Treatment pathway for moderate to severe UC



** Note: Steroids (oral prednisolone) are indicated for inducing response/remission. Azathioprine and 6-MP are indicated as maintenance treatments in patients with two or more flares requiring systemic steroids, where it is not possible to taper steroids, or following acute severe attack. Azathioprine and 6-MP would be started at the same time as oral prednisolone.*

(i) Induction and maintenance of response

Current medical treatments for UC are principally concerned with treating active disease to address symptoms of urgency, frequency of defecation and rectal bleeding, to improve the patient's HRQoL, and thereafter to maintain remission.[3] Treatment usually follows an escalation approach whereby additional drugs are added in order to induce and subsequently maintain response/remission. Initially, patients would most likely be treated using oral and topical 5-ASAs to induce a response. Most commonly, oral 5-ASA treatment involves high-dose oral mesalazine (usually 2.4g-4g/day depending on the particular product used). A dose of up to 2.4g/day mesalazine is used for maintenance. It is very likely that topical 5-ASAs (enemas or suppositories) would also be used during induction; the use of topical 5-ASAs is time-limited (usually 4 weeks maximum) and their efficacy is dependent on the extent of disease and severity of symptoms. If the patient does not respond, achieves but subsequently loses response, or is contraindicated to or unable to tolerate 5-ASAs, treatment is likely to involve the use of oral corticosteroids and immunomodulators. Oral corticosteroids (most likely prednisolone) would be used as a short-term therapy with the intention of inducing a response; corticosteroids are not however used as a maintenance treatment. Prednisolone is typically given at a dose of 40mg/day, with the aim of the dose being tapered by 5mg each week (8 weeks of treatment until the dose is zero). Treatment using immunomodulators, most commonly azathioprine and less commonly 6-MP, would be started at the same time as oral corticosteroids. These are indicated for maintenance rather than induction of response hence patients may receive them on a long-term basis. Patients would likely remain on oral 5-ASA treatment continuously as they may confer other benefits in avoiding cancer, although evidence is conflicting in this respect.[33] If the patient does not respond to

corticosteroids, it is likely that the patient would be considered for treatment using tacrolimus, intravenous steroids or anti-tumour necrosis factor alpha (anti-TNF-α) therapy.

Surgery may be required in emergency scenarios (e.g. in cases of acute severe / fulminant UC) but within the moderately to severely active population, surgery is most likely to be elected by the individual patient. Emergency surgery may be required to ameliorate life-threatening complications of UC such as toxic megacolon, colonic perforation, and massive haemorrhage; it should be noted that surgery might also be used prophylactically to avoid the onset of these complications. More commonly, surgery is elective and is undertaken for severe disease characterised by prior treatment failures and/or frequent UC flares. In some cases, surgery may also be indicated due to the increased risk of colorectal cancer associated with long-standing UC and may also be driven by the identification of pre-malignant dysplasia or malignant neoplasia. Colectomy is associated with postoperative morbidity and a risk of death. Amongst others, complications of surgery may include infertility, transient and chronic pouchitis, wound infections, wound dehiscence and small bowel obstruction.[3]

Patients with less severe disease may be managed either in primary or secondary care. For patients with left-sided or distal UC, follow-up is likely to take place in an outpatient setting, with appointments every 3-12 months depending on the pattern of flares. Follow-up may be consultant-led or IBD nurse-led, but will usually involve a combination of both.

3.4 Description of technology under assessment

3.4.1 Interventions considered in the scope of this report

Three interventions are considered for the adult population (infliximab, adalimumab and golimumab). Only infliximab is licensed for use in children and adolescents. Two biosimilars (Remsima and Inflectra) are also considered as part of the evidence base for infliximab. Interventions are assessed in line with licensed indications, as described in the respective Summary of Product Characteristics (SmPCs) for each intervention.[34,35,36] The interventions under assessment are licensed for the treatment of rheumatoid arthritis (RA), adult CD (infliximab and adalimumab only), paediatric CD (infliximab and adlimumab only), adult UC, paediatric UC (infliximab only), ankylosing spondylitis, psoriatic arthritis and psoriasis (infliximab and adalimumab only).[34,35,36]

3.4.2 Mode of action

Infliximab, adalimumab and golimumab are monoclonal antibodies which inhibit the activity of TNF-α, a key component in the inflammation process.

3.4.3 Marketing licence and administration method

(a) Infliximab (Remicade, Merck Sharp and Dohme)

Infliximab has a UK marketing authorisation for the treatment of moderately to severely active UC in adults, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[34] Infliximab also has a UK marketing authorisation for the treatment of severely active UC in children and adolescents aged 6 to 17 years, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[34]

Infliximab for the treatment of UC is administered by intravenous infusion at a dosage of 5mg/kg followed by additional 5mg/kg infusion doses at 2 and 6 weeks after the initial infusion, then every 8 weeks thereafter.[34] The SmPC states that other concomitant therapies (e.g., corticosteroids and immunosuppressants) should be optimised during infliximab therapy.[34] Infliximab is typically administered intravenously over a 2 hour period as an outpatient or day case appointment. As infliximab treatment is associated with the development of acute infusion reactions, all patients receiving infliximab are required to be observed, in a setting where emergency equipment is available, during the infusion for 1-2 hours post-infusion for safety. Patients may receive pre-infusion treatment with, for example, an antihistamine, hydrocortisone and/or paracetamol. Contraindications to infliximab treatment include a history of hypersensitivity to infliximab or other murine proteins, the presence of tuberculosis or other severe infections such as sepsis, abscesses, and opportunistic infections, and moderate or severe heart failure. Furthermore, women of childbearing potential must use adequate contraception and continue use for at least six months after last receipt of infliximab treatment.

Biosimilar versions of infliximab (Remsima, Celltrion Healthcare; Inflectra, Hospira) are licensed for the same indications as Remicade. The therapeutic indications (including the wording of the licensed indication), dosage and method of administration for Remisima and Inflectra are identical to those for infliximab (Remicade).

(b) Adalimumab (Humira, AbbVie)

Adalimumab has a UK marketing authorisation for the treatment of moderately to severely active UC in adults who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[35] Adalimumab for the treatment of ulcerative colitis is administered subcutaneously according to an induction dose regimen of 160mg at Week 0 and 80mg at Week 2 followed by a

recommended maintenance dosage of 40mg every other week (EOW, increased to 40mg every week [EW] if clinical response is insufficient).[35] Following physician advice, appropriate training and medical follow-up if required, patients may self-inject with adalimumab. The SmPC states that other concomitant therapies (e.g., corticosteroids and immunosuppressants) should be optimised during adalimumab therapy.[35] Contraindications to adalimumab treatment include hypersensitivity to the active substance, the presence of active tuberculosis or other severe infections such as sepsis, and opportunistic infections, and moderate to severe heart failure (NYHA class III/IV). The administration of adalimumab during pregnancy is not recommended.

(c) Golimumab (Simponi, Merck Sharp and Dohme)

Golimumab has a UK marketing authorisation for the treatment of moderately to severely active UC in adults who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[36]

Golimumab for the treatment of UC is administered subcutaneously according to body weight. Patients with body weight less than 80kg receive an initial dose of 200mg, followed by 100mg at week 2, then 50mg every 4 weeks thereafter. Patients with body weight greater than or equal to 80kg receive an initial dose of 200mg, followed by 100mg at week 2, then 100mg every 4 weeks thereafter.[36] Following physician advice and adequate training, patients may self-inject with golimumab. Contraindications to golimumab include hypersensitivity to the active substance, the presence of active tuberculosis (TB) or other severe infections such as sepsis, and opportunistic infections, and moderate or severe heart failure (NYHA class III/IV). The use of golimumab during pregnancy is not recommended.

3.4.4 Criteria for continuing treatment

The SmPC for each intervention describes the use of stopping rules for treatment in non-responders.[34-36]

The SmPC for infliximab states that clinical response should typically be achieved within 14 weeks of treatment (i.e. three doses) and that continued therapy should be carefully reconsidered in patients who show no evidence of therapeutic benefit within 14 weeks. The SmPC also indicates that, for paediatric patients, there is no evidence to support the further use of infliximab in patients who do not respond within the first 8 weeks of treatment.

For adalimumab, the SmPC states that clinical response should be reached within 2-8 weeks of treatment and that treatment should not be continued in patients who fail to respond within this timeframe.

The SmPC for golimumab states that clinical response is expected to be achieved within 12-14 weeks of treatment (i.e. after 4 doses) and that continued therapy should be reconsidered in patients who do not experience therapeutic benefit within this time period.

The SmPCs for each intervention also refer to the requirement to monitor patients closely for infections and to discontinue treatment in patients who develop a serious infection or sepsis.

3.4.5 Current usage in the NHS

Infliximab is currently recommended by NICE as an option for the treatment of acute exacerbations of severely active UC, only in patients in whom ciclosporin is contraindicated or clinically inappropriate. Adalimumab and golimumab do not have recommendations from NICE for use in the treatment of UC. The Assessment Group has received clinical advice to suggest that infliximab, and to a lesser degree, adalimumab, are currently used for the treatment of moderate to severe UC in some larger centres in England and Wales.

3.4.6 Identification of important subgroups

The only subgroup pre-specified in the final NICE scope[37] relates to duration of disease.

3.4.7 Anticipated costs associated with interventions

Table 4 summarises the costs associated with the interventions based on their list prices.[38]

Table 4: Acquisition costs associated with infliximab, adalimumab and golimumab

Drug	Unit type and dose	Price per unit
Infliximab	powder for reconstitution, 100mg vial	£419.62
Adalimumab	40mg prefilled pen or prefilled syringe, 40mg/0.8-mL vial	£352.14
Golimumab	50mg prefilled pen or prefilled syringe	£762.97
	100mg prefilled pen	£1,525.94

4. DEFINITION OF THE DECISION PROBLEM

4.1 Decision problem

The aim of this assessment is to assess the clinical effectiveness and cost-effectiveness of infliximab, adalimumab and golimumab for the treatment of patients with moderately to severely active UC after the failure of conventional therapy.

Interventions

Three interventions are considered within this assessment: infliximab (Remicade®), adalimumab (Humira®) and golimumab (Simponi®). These interventions are described in detail in Section 3.4. Biosimilar versions of infliximab (Remsima®, Celltrion Healthcare; Inflectra®, Hospira) are also licensed for the same indications and are considered as part of the evidence base for infliximab within this assessment report.

Populations (including subgroups)

The assessment considers the following two populations:

(1) Adults aged 18 years and over with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.

As referred to in the final NICE scope[37] and specified in the protocol (Appendix 1) severity of disease in adults would be defined according to the modified Truelove and Witts' severity index (1955) (as described in NICE Clinical Guideline 166).[3]

The following interventions are indicated for use in adults:

- Adalimumab
- Infliximab
- Golimumab

(2) Children and adolescents aged 6 to 17 years (inclusive) with severely active UC, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.

As described in NICE Clinical Guideline 166,[3] severity of UC in children and adolescents was to be assessed using the Paediatric Ulcerative Colitis Activity Index (PUCAI).[26]

The following intervention is indicated for use in children and adolescents:

- Infliximab

The final NICE scope[37] highlighted duration of disease as a potential subgroup of interest; this is examined according to the availability of evidence.

Populations outside of the scope of the appraisal

The following groups were considered to be beyond the scope of the appraisal and therefore are not considered in this assessment report:

- Children with mildly or moderately active UC (as defined by the PUCAI measure)
- Adults with mildly active UC (as defined by the modified Truelove and Witts' [1955] criteria)
- Adults and children with acute severe (systemic) UC.

Relevant comparators

The interventions are compared against each other. Other relevant comparators include standard clinical management options, which (as described in the final NICE scope)[37] could include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors or elective surgical intervention.

Emergency surgical intervention is not considered as a comparator in this assessment as acute severe UC was stated in the final scope as being beyond the remit of the appraisal.

Outcomes

The outcome measures to be considered included:

- Mortality
- Measures of disease activity
- Rates of and duration of response, relapse and remission
- Rates of hospitalisation
- Rates of surgical intervention (both elective and emergency)
- Time to surgical intervention (both elective and emergency)
- Adverse events of treatment (including leakage and infections following surgery)
- Health-related quality of life (HRQoL)

Following discussions during the NICE appraisal scoping process, data relating to mucosal healing were not considered eligible for this assessment.

4.2 Overall aims and objectives of assessment

This assessment addresses the question *"what is the clinical effectiveness and cost-effectiveness of infliximab, adalimumab and golimumab for the treatment of patients with moderately to severely active UC after the failure of conventional therapy as compared against each other and standard clinical management?"*

More specifically, the objectives of the assessment are:

1) To evaluate the clinical effectiveness of each intervention
2) To examine the effect of disease duration on the clinical effectiveness of each intervention (subject to the availability of evidence)
3) To evaluate the adverse effect profile of each intervention
4) To evaluate the incremental cost-effectiveness of each intervention compared (i) against each other and (ii) against all comparators (including medical and surgical options)
5) To estimate the expected net budget impact associated with implementing each intervention
6) To identify key areas in which future research may be valuable

5. ASSESSMENT OF CLINICAL EFFECTIVENESS

A systematic review of the literature including network meta-analyses (NMA) was conducted in order to evaluate the clinical effectiveness and safety of infliximab, adalimumab and golimumab in the treatment of moderately to severely active ulcerative colitis after the failure of conventional therapy.

The systematic review of clinical effectiveness was undertaken in accordance with the general principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[39]

5.1 Methods for reviewing clinical effectiveness

The protocol for this review is registered with PROSPERO (CRD42013006883)[40] and is presented in Appendix 1.

5.1.1 Identification of studies

A comprehensive search was undertaken to systematically identify literature relating to the clinical effectiveness and safety of infliximab, adalimumab and golimumab for treating moderately to severely active UC after the failure of conventional therapy. The search strategy comprised the following main elements:

- Searching of electronic databases
- Handsearching of bibliographies of retrieved papers, key journals and conference proceedings
- Contact with experts in the field

The following electronic databases were searched from inception for published trials and systematic reviews:

- MEDLINE, MEDLINE in-Process and Other Non-Indexed Citations: Ovid. 1946- December 2013
- EMBASE: Ovid. 1974 - December 2013
- Cochrane Library: Wiley Interscience
    - Cochrane Database of Systematic Reviews (CDSR). 1996 - December 2013
    - Database of Abstracts of Reviews of Effects (DARE). 1995 - December 2013
    - Cochrane Central Register of Controlled Trials (CCRT). 1995 - December 2013
    - Cochrane Methodology Register. 1904 - December 2013
    - Health Technology Assessment Database (HTA). 1995 - December 2013
    - NHS Economic Evaluation Database (NHS EED). 1995- December 2013

- CINAHL: EBSCO. 1982 - December 2013
- Web of Science Citation Index: Web of Knowledge. 1900 - December 2013
- Conference Proceedings Citation Index: Web of Knowledge. 1990 - December 2013
- BIOSIS Previews: Web of Knowledge. 1969 - December 2013

The MEDLINE search strategy is presented in Appendix 2. The search strategy combined freetext and MeSH (Medical subject headings) or thesaurus terms relating to *ulcerative colitis*, with freetext and MeSH or thesaurus terms relating to *infliximab, adalimumab or golimumab* combined with highly sensitive filters to retrieve RCTs and systematic reviews. Search terms for infliximab biosimilars were also included. The search strategy was translated across all databases. No date or language restrictions were applied. Literature searches were conducted during December 2013. References were collected in a bibliographic management database and duplicates were removed.

Searches were undertaken to identify unpublished studies (nearing or at completion) relevant to the decision problem within the following research registers**:**

- Clinical Trials.gov (searched December 2013)
- UKCRN Portfolio database (searched December 2013)
- World Health Organisation International Clinical Trials Registry Platform (WHO ICTRP) (searched March 2014)

Proceedings of the following conferences were searched from 2009-2014 (where possible) for recent research:

- Congress of Crohn's and Colitis Conference (ECCO)
- Digestive Disease Week (DDW)
- Gut (British Society of Gastroenterology)

Key journals were identified using the PubMed PubReMiner facility and electronic tables of contents were searched from March 2013 to February 2014 for the following journals:

- Inflammatory Bowel Diseases
- Alimentary Pharmacology & Therapeutics
- Gastroenterology
- Journal of Crohns & Colitis
- American Journal of Gastroenterology

Citation searches were performed on included studies in Web of Science in March 2014.

Manufacturers' submissions received by NICE, as well as any relevant systematic reviews, were also handsearched in order to identify any further potentially relevant clinical trials.

5.1.2 Inclusion and exclusion criteria

Inclusion and exclusion criteria were based on the final NICE scope[41] and were applied as described below.

5.1.2.1 Study selection

The selection of eligible articles was undertaken using a two-stage process. Firstly, in order to assess agreement in the sifting approach between systematic reviewers, a check for consistency was conducted in the early stages of the sifting process. The reviewers (RA and MMSJ) double sifted a total of 940 titles and abstracts. Kappa statistics of 0.888 and 1.000 were obtained, indicating very high strength of agreement.

All remaining titles and abstracts were examined for inclusion by one reviewer (RA and MMSJ each sifted 50% of total citations at title and abstract level). Any citations that clearly did not meet the inclusion criteria (e.g. animal studies, studies unrelated to UC) were excluded. During the second stage of the sifting process, full text articles were examined for inclusion by one reviewer (RJA or MMSJ). Any uncertainty in the eligibility of potentially relevant full text articles was resolved through discussion. Trials retrieved for full paper screening which were subsequently excluded were tabulated (see Appendix 3) together with justification for their exclusion.

5.1.2.2 Inclusion criteria

Studies were included in the review if they met the inclusion criteria outlined below.

a) Interventions

Any of the following interventions were included:

i) For adults (defined by the Assessment Group as aged 18 years and over):

- Adalimumab
- Infliximab
- Golimumab

ii) For children and adolescents aged 6 to 17 years (inclusive):

- Infliximab

Biosimilar versions of infliximab (Remsima®, Celltrion Healthcare; Inflectra®, Hospira) are also licensed for the same indications as Remicade and have been considered as part of the evidence base for infliximab within this assessment.

Studies in which the interventions were assessed in line with licensed indications were included in the systematic review.

b) Populations

i) Adults aged 18 years and over with moderately to severely active (non-systemic) UC (defined as patients with moderately active disease according to the modified Truelove and Witts' criteria [1955][24] only) whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant of or have medical contraindications to such therapies.

ii) Children and adolescents aged 6 to 17 years with severely active (non-systemic) UC (as classified by the Paediatric Ulcerative Colitis Activity Index (PUCAI) measure[26]) whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications to such therapies.

c) Comparators

Relevant comparators included: i) interventions as defined in the protocol for this assessment (see Appendix 1, i.e. infliximab, adalimumab or golimumab compared with each other) and ii) standard clinical management, which may include a combination of aminosalicylates (sulfasalazine, mesalazine/mesalamine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors or elective surgical intervention.

d) Outcomes

Eligible outcomes for consideration were:

- Mortality
- Measures of disease activity
- Rates of and duration of response, relapse and remission
- Rates of hospitalisation
- Rates of surgical intervention (both elective and emergency)
- Time to surgical intervention (both elective and emergency)
- Adverse events of treatment (including leakage and infections following surgery)
- Health-related quality of life

Following discussions during the NICE appraisal scoping process, data relating to mucosal healing were not considered eligible for this assessment.

e) Study design

Randomised controlled trials (RCTs) were eligible for inclusion in the systematic review of clinical effectiveness. Long-term extension studies associated with included RCTs were also included in the review.

Studies published as abstracts or conference presentations were eligible for inclusion only if sufficient details were presented to allow an assessment of the trial methodology and results to be undertaken.

5.1.2.3 Exclusion criteria

The following types of studies were excluded from the review:

- Studies which included adults with mildly active UC (as defined by the modified Truelove and Witts' [1955] criteria[24]), where no separate data were reported for patients with moderate to severe UC
- Studies which included children with mildly or moderately active UC (as defined by the PUCAI measure[26])
- Studies which included adults with (acute) severely active UC as defined by the modified Truelove and Witts' [1955] criteria[24] (representing patients who are systemically ill and are therefore beyond the remit of this appraisal)
- Studies which included adults, adolescents or children with acute severe UC, whose disease is systemic as shown by tachycardia, fever, anaemia or a raised erythrocyte sedimentation rate (representing patients who are excluded as they are outside the remit of this appraisal)
- Studies which included patients with acute severe UC previously hospitalised and treated with intravenous steroids (representing patients in a potentially life threatening medical emergency and excluded as they are outside the remit of this appraisal)
- Studies which included patients with inflammatory bowel disease other than UC (e.g. Crohn's disease) where data were not reported separately for UC patients
- Studies where interventions were not administered in accordance with licensed indications
- Systematic reviews and clinical guidelines (selected systematic reviews identified by the clinical effectiveness searches were used as sources of references)
- Studies which were published only in languages other than English
- Studies based on animal models
- Preclinical and biological studies
- Narrative reviews, editorials and commentaries
- Reports published as abstracts or conference presentations only, where insufficient details were reported to allow an assessment of study quality or results.

5.1.3 Data abstraction strategy

Data relevant to the decision problem were extracted by one reviewer (RA or MMSJ). Data were extracted without blinding to authors or journal. A data extraction form was developed and piloted on two included trials before slight revisions and final use on all included trials. Data relating to study arms in which the intervention treatments were administered in line with their licensed indications were extracted; data relating to the unlicensed use of the interventions were not extracted. All extracted data were double-checked by a second reviewer (MMSJ or CC). The safety data extracted were informed by the SmPCs for each product (available from http://www.medicines.org.uk/emc/).[34,35,36] The key safety issues included such items as the number of patients experiencing infections, number of patients experiencing serious infections, number of patients experiencing malignancy, and the occurrences of infusion-related or injection-site reactions (as appropriate to the mode of administration for each intervention). Study results that were presented only in graphical format were digitised and estimated using Engauge software version 4.1.[42] Where multiple publications of the same study were identified, data extraction was undertaken on all relevant associated publications, and findings were presented together with reference to their published source.

5.1.4 Critical appraisal strategy

The methodological quality of each included study was assessed by one reviewer (RJA or MMSJ). The quality of included studies was assessed using the Cochrane Risk of Bias Tool.[43] This tool addresses specific domains, namely: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data and selective outcome reporting. RCTs were classified as being at 'high risk' of attrition bias where drop-out in any treatment arm was ≥10%.[44] The Assessment Group requested the trial protocols for all included trials from the manufacturers of the products included in this appraisal. These were received for some trials and were used, alongside Clinical Study Reports (CSRs) provided by the manufacturer for some trials and outcomes listed in ClinicalTrials.gov records, in order to inform the selective reporting domain of the Cochrane Risk of Bias tool. All quality assessment findings were double checked by a second reviewer (RJA or MMSJ).

5.1.5 Methods of data synthesis

The extracted data were presented for each study, both in structured tables and as a narrative description.

5.1.5.1 Methods for the estimation of efficacy using network meta-analysis

Network meta-analysis methods are described in full alongside results in Section 5.2.3.3.

5.1.5.2 Supplementary meta-analyses

Where considered appropriate, secondary outcomes of interest were analysed using classical meta-analysis methods. Meta-analysis was undertaken using Cochrane Review Manager software (version 5.2). Outcomes reported as continuous data were estimated using a mean difference (MD) with 95% confidence intervals (95% CIs). Dichotomous outcomes were estimated as risk ratios (RRs) with associated 95% CIs. Where RCTs reported adverse events in sufficient detail, these were analysed as dichotomous data. Clinical heterogeneity across RCTs (the degree to which RCTs appear different in terms of participants, intervention type and duration and outcome type) was considered prior to data pooling. Random-effects models were applied. Effect estimates, estimated in Review Manager as Z-scores, were considered statistically significant at a cutoff of $p<0.05$.

5.2 Results

5.2.1 Quantity of research available

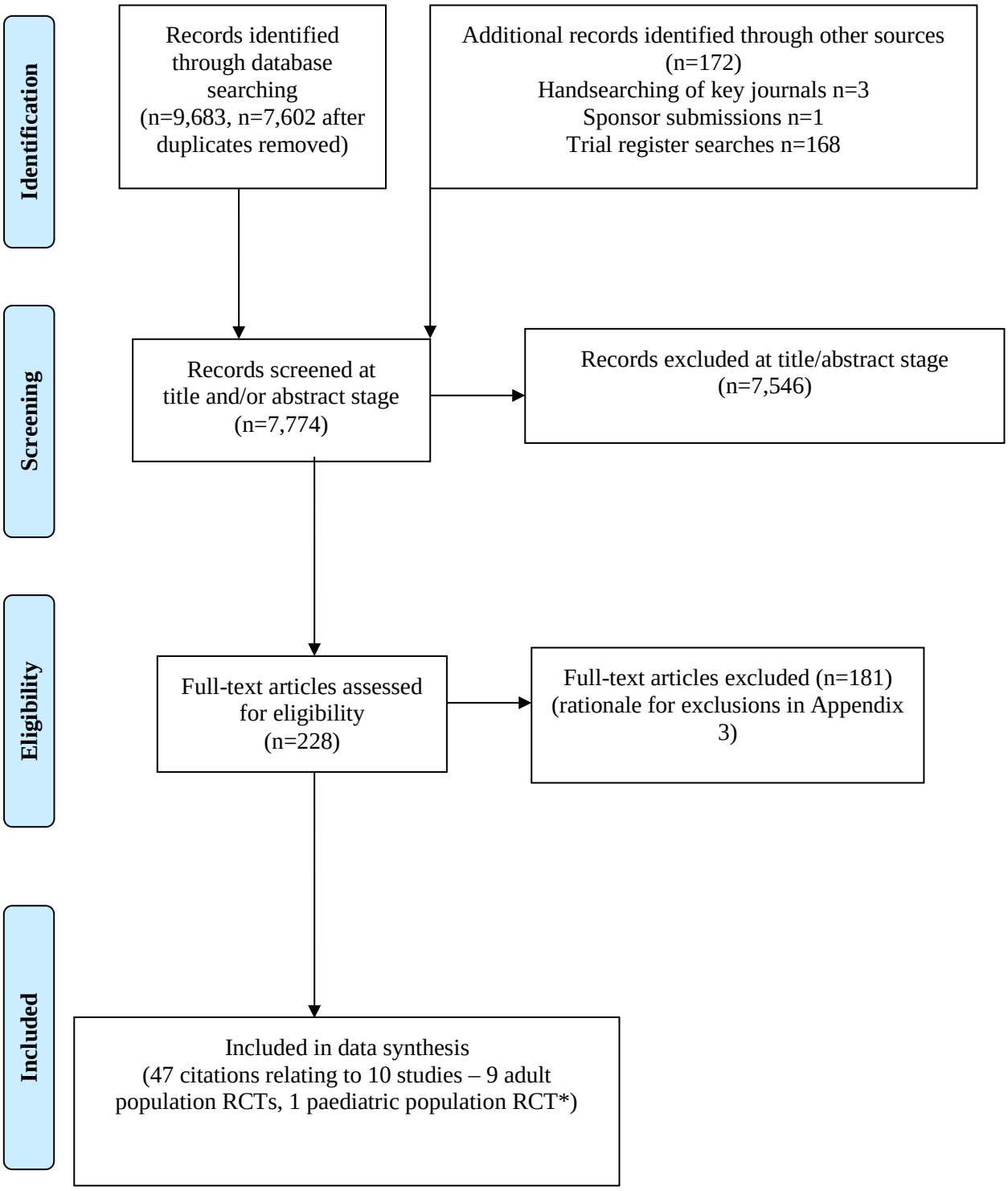
The searches described in Section 5.1.1 yielded 7,774 potentially relevant citations (7,602 from searches of electronic databases after removal of duplicates), 3 from handsearching of key journals, 1 from sponsor submissions and 168 from trial register searches). Of these records, 7,546 were excluded at the title and abstract stage. Full texts of 228 studies were obtained for scrutiny. Of these, 181 citations were excluded (it was not possible to obtain nine studies hence these were excluded, see Appendix 3).

No additional eligible trials that were completed or nearing completion were identified through the trial register searches. Trial NCT01551290 (a study of infliximab versus placebo in Chinese subjects by Xian-Janssen Pharmaceutical Ltd) was stated to be ongoing with an estimated completion date of November 2014. Trial NCT01863771 (a study of golimumab maintenance treatment versus placebo in Japanese patients by Janssen Pharmaceutical) was recruiting as of February 2014. As such, neither trial was judged to be completed or nearing completion.

A total of 47 citations relating to 10 RCTs were included in the review.[45,46,47,48,49,50,51,52,53] The search process is summarised in the form of a PRISMA diagram in Figure 2.

European Public Assessment Reports (EPARs) were available for all included interventions. However, associated FDA reports for interventions could not be identified from the FDA website.

Figure 2: Flow diagram of study inclusion (adapted from PRISMA)[54]



** not including sponsor submissions and EPARs*

5.2.2 Summary of study and population characteristics of included trials

5.2.2.1 Study characteristics

The available comparisons between licensed doses of interventions and placebo are tabulated within the adult population RCTs in Table 5. The trial design characteristics of the included trials are outlined in Tables 6 and 7. The outcome measures pre-specified in the final NICE scope[37] and protocol (see Appendix 1) were all addressed by the included trial evidence, with the exception of rates of relapse. As stated in Section 5.1, data relating to mucosal healing were not eligible for this assessment.

Table 5: Licenced dose comparisons for included adult population RCTs

Trial	Licenced treatment comparisons
ULTRA1[45]	Placebo ADA 160/80mg (licenced induction dose)
ULTRA2[46]	Placebo ADA 160mg at week 0, 80mg at week 2 and then 40mg EOW (licenced maintenance dose) beginning at week 4
Suzuki *et al.*[47]	Placebo ADA 160/80mg (licenced induction dose)
PURSUIT-SC[48]	Placebo GOL 200/100mg (licenced induction dose)
PURSUIT-Maintenance[49]	Placebo GOL 50mg GOL 100mg (licenced maintenance doses)
ACT1[50]	Placebo IFX5mg/kg
ACT2[50]	Placebo IFX5mg/kg
Probert *et al.*[51]	Placebo IFX5mg/kg
UC-SUCCESS[52]	No Placebo IFX5mg/kg AZA IFX5mg/kg /AZA

ADA – adalimumab; IFX – infliximab; GOL – golimumab; AZA - azathioprine

a) Population: Adults aged 18 years and over with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies

A total of nine relevant RCTs were identified which were performed in adult populations. Four RCTs evaluated the use of infliximab (ACT1,[50] ACT2,[50] Probert *et al.*[51] UC-SUCCESS[52]), three RCTs were of adalimumab (ULTRA1,[45] ULTRA2,[46] and Suzuki *et al.*, 2014[47]) and two RCTs were of golimumab (PURSUIT-SC,[48] PURSUIT-Maintenance[49]). Four of these RCTs (ACT1, ACT2, ULTRA1, ULTRA2) had long-term open label extension studies associated with them (ACT1 and ACT2 extension studies,[55] ULTRA3[56]) which were also included as part of the evidence for these

interventions. All of the included RCTs for adults were undertaken against a comparator of placebo, with the exception of UC-SUCCESS which assessed the use of infliximab against active comparators of azathioprine and combination infliximab/azathioprine. No head-to-head RCTs comparing interventions of interest against each other were identified for adults. All RCTs were Phase III (where stated), with the exception of Suzuki *et al.* (Phase II/III) and PURSUIT-SC (Phase II/III). Where stated, all included adult population trials were powered for the primary endpoints of clinical remission (ULTRA1, ULTRA2, Probert *et al.*, UC-SUCCESS) or clinical response (ACT1, ACT2, PURSUIT-SC, PURSUIT-Maintenance). Where the geographical location(s) of study sites were reported, all trials were multicentre, international studies, with the exception of Probert *et al.*, which was performed in the UK and Germany, and Suzuki *et al.*, which was conducted exclusively in Japan. All trials were at least partly industry-funded.

Eight trials included time points for the assessment of the use of interventions in achieving induction of clinical response or remission, of which four assessed infliximab (ACT1, ACT2, Probert *et al.*, UC-SUCCESS), three assessed adalimumab (ULTRA1, ULTRA2, Suzuki *et al.*) and one assessed golimumab (PURSUIT-SC). Six trials reported outcomes at time points for the evaluation of the use of interventions in the maintenance of clinical response or remission, consisting of three infliximab trials (ACT1, ACT2, UC-SUCCESS), two adalimumab trials (ULTRA2 and Suzuki *et al.*) and one golimumab trial (PURSUIT-Maintenance).

None of the included RCTs applied Truelove and Witts' disease severity criteria[24] in their eligibility criteria (as referred to in the final NICE scope for this appraisal[37] and as specified in the protocol [Appendix 1]). All included trials applied the Mayo score (except Probert *et al.* where the score was specified simply as UC symptom score/UCSS) to classify the disease severity of potential participants (note – the UCSS was confirmed to be equivalent to Mayo score by Professor C. Probert, May 2014). The included trials required a Mayo score of 6-12 (with evidence of endoscopic disease) for participant eligibility. Mayo scores of 6-12 were described in the included trial literature as moderate to severe disease and were also subsequently confirmed following clinical advice as representing moderate to severe disease (note - a*d hoc* searches were performed to attempt to identify evidence relating to the relationship between the Truelove and Witts' and Mayo disease severity indices, however no evidence published in full text in English could be identified). Included trials required a varying range of prior use of conventional therapy for eligibility, as described in Tables 6 and 7. The UC-SUCCESS trial, which specified patients to be either AZA-naïve or free from AZA treatment for at least 3 months before enrolment, was a borderline inclusion in the clinical effectiveness systematic review since the wording of the population in the scope and the licensed indications required prior use

of AZA or 6-MP. However, since the trial reported the use of a stated (albeit low) proportion of prior immunosuppressant use, this trial was included in the clinical effectiveness systematic review for completeness. This trial was not however eligible for subsequent inclusion in meta-analyses or network meta-analyses. Suzuki *et al.* included Japanese patients aged 15 years and above (adalimumab is not licensed in the paediatric population), but the mean ages of participants across treatment arms at baseline was 41.3 - 42.5 years.

The COMET initiative (http://www.comet-initiative.org), which promotes the use of core outcome sets in clinical trials, referenced the work by Cooney *et al.*, 2007[23] in classifying the use of outcome measures in UC clinical trials. Whilst acknowledging the very broad range of available disease severity/activity measures available, all adult population trials included in the assessment were consistent in their utilisation of the Mayo score as a measure of clinical response and/or remission. The included trial by Probert *et al.* applied the UC symptom score (UCSS) in the evaluation of clinical remission at induction. This score is equivalent to the full Mayo score: the components of the UCSS are consistent with the elements assessed within the Mayo score (i.e. stool frequency, rectal bleeding, sigmoidoscopic appearance and physician's global assessment) and also is referenced using the citation quoted for the Mayo score (Schroeder *et al.*, 1987[25]). None of the included studies utilised the modified Truelove and Witt's criteria[24] (as referred to in the NICE appraisal scope[37] and as specified in the assessment protocol) in their outcome assessments.

As recommended in the Committee for Human Medicinal Products (CHMP) guideline[57] on the development of new medicinal products for the treatment of UC patients with confirmed UC were eligible for the included trials. Severity of disease was defined by clinical and endoscopic evaluation, as recommended in the CHMP guideline. Although the interventions of interest in this assessment were developed for the treatment of patients not responding/intolerant to previous immunomodulatory therapy, the Assessment Group did not consider that adequate definitions of inadequate response/intolerance were included in trials, as recommended by the CHMP guideline. The guideline recommended that, for refractory populations, a minimum duration and dose of previous baseline medication should be defined; this was not the case in the included trials. In addition, intolerance was not defined by minimum criteria of severity in the trials. In terms of study duration, it was recommended that induction studies should be 8 to 12 weeks, but could be shorter based on the pharmacodynamics properties of the study drug. All induction trials assessed efficacy at 8 weeks, with the exception of the PURSUIT-SC golimumab trial, which was a 6 week study. All maintenance studies were at least one year in length, as recommended in the CHMP guideline.

Adalimumab

ULTRA1 was a multicentre Phase III RCT in adults undertaken across the USA, Puerto Rico, Canada, Western Europe, and Eastern Europe. In the original protocol, 186 participants were randomised and in the amended protocol 390 were randomised (130 per group including placebo). Length of treatment was 12 weeks in the original protocol and 8 weeks in the amendment. Outcomes were reported at week 8. ULTRA2 was a multicentre Phase III RCT in adults undertaken across North America, Europe, Australia, New Zealand, and Israel. Five hundred forty-two participants were randomised to two groups including placebo. Outcomes were reported at week 8 and week 52. Suzuki *et al*. was a 52-week Phase II/III trial in Japanese adults. Two hundred seventy-four participants were randomised to three treatment groups including placebo. Outcomes were reported at 8 weeks and 52 weeks. The two induction adalimumab groups (one licenced dose and the other unlicensed) were combined as one active treatment group for outcomes at 52 weeks. ULTRA3 was the 52-week open label extension study to ULTRA1 and ULTRA2.

Golimumab

PURSUIT-SC was a Phase II/III multicentre RCT in adults reporting outcomes at week 6. The trial was performed across 217 sites (Eastern Europe 400 patients, North America 278 patients, Asia Pacific and South Africa 204 patients, and Western Europe and Israel 183 patients). This was a dose-ranging study with 169 patients randomised to four groups including placebo. PURSUIT-Maintenance was a Phase III RCT in adults across 251 sites across Eastern Europe (477 patients), North America (323 patients), Asia Pacific and South Africa (237 patients), and Western Europe and Israel (191 patients). Overall, 1228 patients who were responders to golimumab induction therapy in two previous golimumab induction therapy trials (including PURSUIT-SC) were randomised to licenced maintenance doses of 50mg golimumab, 100mg golimumab or placebo.

Infliximab

ACT1 was a multicentre Phase III RCT conducted across 62 sites. Three hundred and sixty four adult patients were randomly assigned to licenced and unlicensed induction doses or placebo. ACT2 was a multicentre Phase III RCT across 55 sites. Three hundred and sixty four adult patients were randomly assigned to licenced and unlicensed maintenance doses or placebo. Probert *et al*. was an RCT undertaken across four centres in the UK and Germany. Forty three adult participants were randomised to infliximab or placebo and outcomes assessed at week 6. UC-SUCCESS was a multicentre RCT undertaken in adults. Two hundred thirty-nine participants were randomised to infliximab, AZA, or combination therapy (with no placebo group included). Outcomes were assessed at weeks 8 and 16.

b) Population: Children and adolescents aged 6 to 17 years (inclusive) with severely active UC, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies

A single Phase III open-label RCT was identified for the paediatric population (T72, Hyams *et al.*, 2012[53]) which evaluated the use of infliximab in maintenance therapy. All patients received the licensed infliximab induction regimen before being randomised to one of two infliximab maintenance regimens. Outcomes were reported at week 30 and week 54. No placebo-controlled or head to head RCTs were identified for children and young people. The absence of a placebo or non-infliximab control group in the included RCT made it difficult to consider the effectiveness of infliximab in paediatric patients as compared against conventional UC therapies. This industry-funded trial had the primary endpoint of clinical response and was conducted in the USA, Netherlands, Canada and Belgium. Eligible patients were six to 17 years of age with moderately to severely active UC and a Mayo score of 6-12 points with endoscopic evidence of disease. Therefore, whilst infliximab is licensed in this age group with severe disease only (as reflected in the scope population), this trial was included in consideration of limited paediatric RCT evidence.

Table 6: Trial design characteristics of included clinical effectiveness studies in adults

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
Adalimumab									
ULTRA1 (NCT00385736, M06-826) Reinish *et al.*, 2011[45]	Multicentre, randomised, double-blind, PBO-controlled trial Phase III	Inclusion: Adult ambulatory patients, moderately to severely active UC, Mayo score 6 -12 with endoscopy subscore of 2 -3, despite concurrent and stable treatment with oral corticosteroids and/or immunomodulators. Concurrent therapy not required if failed to respond/could not tolerate previous corticosteroid/ immunomodulator treatment. Exclusion: Ulcerative proctitis, previous receipt of anti-TNF agent or biological agent, receipt of intravenous corticosteroids within 14 days prior to screening/during screening; receipt of cyclosporine,	USA, Puerto Rico, Canada, Western Europe, and Eastern Europe 94 study centres: (USA, 34; Puerto Rico, 3; Canada, 5; Western Europe, 32; and Eastern Europe, 20)	Original study protocol: ADA 160mg/80mg s.c. = 93 randomised, PBO s.c. = 93 randomised. Protocol amended to include ADA 80/40mg group. ADA 160/80mg s.c. = 130 randomised ADA 80/40mg s.c.= 130 randomised PBO s.c. = 130 randomised ADA	No	Clinical remission (Mayo score ≤2 with no individual subscore >1) at week 8 assessed in the ITT-A3 (amendment) population.	Week 8 Included in NMA? Yes	NA Included in NMA? No (8 week study)	Abbott

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
		tacrolimus, mycophenolate mofetil, or methotrexate within 60 days of baseline.		160/80mg group received ADA 160mg at week 0, ADA 80mg at week 2, ADA 40mg at weeks 4 and 6. ADA80/40mg group received ADA 80mg at week 0, ADA 40mg at week 2 and ADA 40mg at weeks 4 and 6. PBO group received same number of s.c. injections as patients in ADA treatment groups.					
ULTRA2	Multicentre, randomised,	Inclusion: Adults with moderately	North America,	PBO 260 (14 excluded	Yes	Proportion of patients	Week 8	Weeks 32 and 52	Abbott

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
(NCT00408629, M06-827) Sandborn *et al.*, 2012[46]	double-blind, PBO-controlled trial Phase III	to severely active UC for ≥3 months and Mayo score of 6 12 (endoscopy subscore ≥ 2), despite concurrent therapy with steroids and/or azathioprine or 6-mercaptopurine. Exclusion: previous treatment with adalimumab; receipt of intravenous corticosteroids intravenous corticosteroids within 2 weeks of screening; receipt of cyclosporine, tacrolimus, or mycophenolate mofetil within 1 month of baseline; or receipt of any investigational agent within 30 days/5 half-lives before baseline.	Europe, Australia, New Zealand, and Israel 103 study centres	due to site non-compliance) ADA 160/80/40mg 258 (10 excluded due to site non-compliance) Patients received ADA s.c. 160mg at week 0, 80mg at week 2 and 40mg eow from week 4 or matching PBO and followed through week 52.	Patients with inadequate response permitted to switch to open label ADA (40mg eow) from week 12. Patients with inadequate response at 2 visits on open label ADA 40mg EOW permitted to escalate to 40mg EW. Data handled using nonresponder imputation methods.	achieving clinical remission at week 8 and proportion of patients achieving clinical remission at week 52	Included in NMA? Yes	Included in NMA? Yes	
ULTRA3 (M10-223) Reinisch *et al.*, 2013[56]	Long term, single arm, open label extension study including patients	Inclusion Patients in both studies who completed the 52-week visit had the option of enrolling in the extension study (M10-223)	See ULTRA1 and ULTRA2	Patients continued to receive OL ADA (EOW or EW dosing permitted)	Yes Patients who had inadequate response or responded and then	NR	NA	Evaluation of ADA maintenance regimens	Abbott

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
	from ULTRA1 and ULTRA2 (currently ongoing)	Exclusion Patients not responding to weekly ADA from Study M06-826 or M06-827		ADA 40mg EOW or EW (N=588)	experienced disease flare eligible for ADA dose increase to 40mg EW (no earlier than the week 12 visit) or the week 2 visit if already receiving OL ADA.				
Suzuki *et al.*, 2014 (NCT00853099, M10-447) Suzuki *et al.*, 2014[47]	Multicentre, randomised, double-blind, PBO-controlled trial Phase II/III	Inclusion: Japanese patients ≥15 years, moderately to severely active UC, Mayo score 6-12 with endoscopy subscore ≥2 despite concurrent treatment with stable doses of oral corticosteroids and/or immunomodulators. Patients previously treated with corticosteroids or immunomodulators during past five years and had failed to respond or who could not tolerate treatment eligible Exclusion:	Japan 65 study centres	PBO = 96 ADA 160/80mg = 90 ADA 80/40mg = 87 Patients received s.c. ADA 160mg at week, 80mg at week 2 and 40mg EOW from week 4, or ADA 80mg at week 0,	Yes Patients with inadequate response to study drug or flare at or after week 8 permitted to enter rescue arm with 4 weeks of blinded ADA (either 160mg initially and 80mg 2 weeks later for PBO group, or 40mg initially and 2 weeks later for patients in either ADA group)	NR	Week 8 Included in NMA? Sensitivity analysis only (on basis of exclusively Japanese population and population eligibility age patients ≥15 years)	Weeks 32 and 52 Included in NMA? Sensitivity analysis only (on basis of exclusively Japanese population and population eligibility age patients ≥15 years)	Abbott

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
		Patients with prior treatment with anti-TNF therapies or other biologics,discontinuation of oral corticosteroids ≤ 2 weeks before baseline; receipt of corticosteroid injection, ciclosporin, tacrolimus, or mycophenolate mofetil ≤ 4 weeks before baseline.		40mg at week 2 and 40mg EOW from week 4 or PBO	followed by open label ADA 40mg EOW (with option to escalate to 80mg EOW if inadequate response/flare ≥ 8 weeks later. Data handled using nonresponder imputation.				
Golimumab									
PURSUIT-SC (NCT00487539) (Program of UC Research Studies Utilizing an Investigational Treatment - Subcutaneous) Sandborn *et al.*, 2014a[48]	Multicentre, randomised, double-blind, PBO-controlled trial Integrated Phase II and Phase III trial	Inclusion: Patients with moderate to severe UC, Mayo score of 6-12, with endoscopic subscore ≥ 2), inadequate response to/failed to tolerate 1 or more of following: oral 5-ASA, oral corticosteroids, AZA, and/or 6-MP; or corticosteroid dependent. Exclusion: Patients with colitis limited to 20 cm of colon; patients with earlier use of: biologic	Eastern Europe, North America, Asia Pacific, South Africa, Western Europe, Israel 217 sites (Eastern Europe 400 patients, North America 278 patients, Asia Pacific and South Africa 204 patients, and Western	Phase II PBO = 42 plus 31 enrolled whilst Phase II data being analysed Phase II GOL 200/100mg = 42 plus 31 enrolled whilst Phase II data being analysed Phase III	No	Phase III primary endpoint was clinical response at week 6.	Week 6 Included in NMA? Yes	NA Included in NMA? No (6 week study)	Janssen Research & Development

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
		anti-TNF agent(s) natalizumab or other agents targeting alpha-4 integrin, B-cell depleting agents (rituximab), or T-cell depleting agents (alemtuzumab, visilizumab) within 12 months of first study drug dose (or continued B- or T-cell depletion > 12 months after completing treatment with lymphocyte-depleting agents); oral corticosteroids at dose > 40mg prednisone or equivalent per day; receipt of cyclosporine, tacrolimus, sirolimus, or mycophenolate mofetil within 8 weeks before first study agent infection	Europe and Israel 183 patients)	PBO =258 Phase III GOL 200/100mg = 258 Patients received s.c. GOL or PBO at weeks 0 and 2.					
PURSUIT-Maintenance (NCT00488631) Sandborn *et al.*, 2014b[49]	Randomised -withdrawal, PBO-controlled, double-blind multicentre trial. Patients who	Inclusion: Patients had completed 1 of 2 GOL induction studies (PURSUIT-SC or PURSUIT-IV). Patients eligible for induction studies had moderate to severe UC	Eastern Europe, North America, Asia Pacific and South Africa, and Western Europe and Israel	PBO randomised = 156 GOL 50mg = 154 GOL 100mg = 154	Yes Induction responders who lost clinical response permitted to modify	Clinical response maintained through week 54 among GOL induction responders.	NA Included in NMA? No (maintenance trial only)	Weeks 30 and 54 Included in NMA? Yes	Janssen Research & Development

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
	responded to GOL induction therapy (n=464) randomised at baseline visit in 1:1:1 ratio to SC PBO, GOL 50mg or GOL 100mg Phase III	and Mayo score 6-12 with endoscopic subscore ≥ 2. Patients had inadequate response to/had failed to tolerate 1 or more of following: oral 5-ASA, oral corticosteroids, immunosuppressives (AZA or 6-MP) or were corticosteroid-dependent. Exclusion: Patients with isolated proctitis excluded from induction studies.	251 sites across Eastern Europe (477 patients), North America (323 patients), Asia Pacific and South Africa (237 patients), and Western Europe and Israel (191 patients).	Randomised patients received s.c. PBO, GOL 50mg or GOL 100mg every 4 weeks through week 52 (efficacy analysis population). PBO-induction responders and PBO- or GOL-induction nonresponders eligible but not randomised. PBO induction responders received PBO every 4 weeks through week 52. GOL-induction or	treatment. PBO group to GOL 100mg every 4 weeks, GOL 50mg re-randomised to GOL 50mg or GOL 100mg every 4 weeks and GOL 100mg re-randomised to GOL 100mg or GOL 200mg (GOL 200mg dose subsequently discontinued and patients on GOL 200mg decreased to GOL 100mg). Proportions of subjects who underwent dose adjustments were 33.8% in the golumumab 50mg group, 28.5% in the golimumab 100mg group and 48.7% in				

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
				PBO-induction nonresponders received GOL 100mg every 4 weeks through week 12, assessed at week 16 and discontinued from study if disease activity unimproved. Nonrandomised patients included in demographic, PK and safety summaries only.	the placebo group.				
Infliximab									
ACT1 (Active UC Trial 1) (NCT00096655) Rutgeerts *et al.*,	Multicentre, randomised, double-blind, PBO-controlled trial Phase III	Active UC with Mayo score of 6 - 12 points and moderate-to-severe active disease on sigmoidoscopy despite concurrent treatment with corticosteroids alone or in combination	62 sites Geographical locations NR	PBO - 121 IFX 5mg/kg - 121 IFX 10mg/kg - 121 Received agent at	No	Clinical response at week 8	Week 8 Included in NMA? Yes	Weeks 30 and 54 Included in NMA? Yes	Schering-Plough

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
2005[50]		with AZ or 6-MP included Patients with diagnosis of indeterminate colitis, Crohn's disease or clinical findings suggestive of Crohn's disease; positive tuberculin skin tests; previously exposed to IFX or any other anti-TNF agent excluded.		weeks 0, 2, 6, 14, 22, 30, 38 and 46:					
ACT2 (Active UC Trial 2) (NCT00036439) Rutgeerts *et al.*, 2005 [50]	Multicentre, randomised, double-blind, PBO-controlled trial Phase III	Same as ACT1	55 sites Geographical locations NR	PBO - 123 IFX 5mg/kg - 121 IFX 10mg/kg - 120 Received agent at weeks 0, 2, 6, 14, 22, 30, 38 and 46: Received agent at weeks 0, 2, 6, 14, 22, 30, 38 and 46:	No	Clinical response at week 8	Week 8 Included in NMA? Yes	Week 30 Included in NMA? Yes	Schering-Plough
ACT1 and ACT2 extension studies	Long-terms extension studies	Inclusion: Patient eligibility described for ACT	See ACT1 and ACT2	229 randomised patients in	Yes Patients	NR	NA	Evaluation of long-term IFX maintenance	As for ACT1 and

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
Reinisch *et al.*, 2012[55]	(open label) of ACT Phase III trials. Study design identical for ACT1 and ACT2 extension studies.	studies. Patients who (in opinion of investigator) could benefit from continued treatment eligible to enter extension study after completing main study treatment and assessments through weeks 46 and 54 (ACT1) or weeks 22 and 30 (ACT2). Exclusion: Patients who received experimental medication to treat UC after completion of main study ineligible.		IFX group entering extension studies. Participating patients continued to receive blinded treatment to which they had been randomised. Sites unblended to treatment after week 54 ACT1 and extension week 24 (E24) in ACT2 analyses completed (and PBO patients discontinued at this point and not included in analyses).	receiving IFX 10mg/kg permitted to lower dose to 5mg/kg. Patients losing response while receiving IFX 5mg/kg permitted to raise dose to 10mg/kg.			to week 152	ACT2

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
Probert *et al.*, 2003 NCT number NR Probert *et al.*, 2003[51]	Multicentre, randomised, double-blind, placebo-controlled trial Phase NR	Inclusion: Patients with UC symptom score (UCSS) of ≥ 6 or more and a sigmoidoscopy score of ≥ 2 on Baron scale, failed to respond to conventional treatment with glucocorticoids. Exclusion: Patients who had received cyclosporin, any therapeutic agent used to directly reduce TNF, or any investigational drug within three months of enrolment, as well as those who had recently commenced treatment (within the last three months) with 6-MP or AZA.	UK and Germany 4 study centres	PBO = 20 IFX 5mg/kg = 23 Patients received i.v. IFX 5mg/kg or PBO at week 0 and second identical infusion at week 2. At week 6, all patients with continued active UC offered open label IFX 10mg/kg.	Yes, from week 6 only (all patients with continued active UC offered open label IFX 10mg/kg)	Clinical remission at week 6	Week 6 (clinical remission only; no clinical response data available for week 6) Included in NMA? No (excluded from NMA as definition of clinical remission inconsistent with other trials (i.e. total Mayo score ≤ 2 but does not specify no individual subscore > 1 as in other trials)	NA Included in NMA? No (no maintenance time points)	Schering-Plough and BMBF Competence Network (Germany)
UC-SUCCESS (NCT00537316) Panacionne *et al.*, 2014[52]	Multicentre, randomised, double-blind (double-dummy),	Inclusion: Patients with moderate to severe UC defined as Mayo scores of 6 to 8 and 9 to 12,	62 study centres Geographical locations NR	AZA = 80 IFX = 79 IFX/AZA = 80	Yes Nonresponders to AZA at week 8 had IFX	Corticosteroid (CS)-free remission at week 16	Week 8 Included in NMA? No (borderline	16 weeks Included in NMA? No (16 weeks of	Schering-Plough

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
	placebo-controlled trial Phase III	respectively. Patients responded inadequately to course of corticosteroids +/- mesalamine within past 12 weeks. Patients required to be either AZA-naive or free from AZA treatment for at least 3 months before enrolment. Prohibited medications at study entry included methotrexate and calcineurin inhibitors (tacrolimus, cyclosporine).		Patients received i.v. IFX 5mg/kg at weeks 0, 2, 6 and 14 + oral once daily PBO capsules, or oral AZA 2.5mg/kg daily + PBO i.v. on IFX schedule or combination therapy with both drugs	rescue infusions at weeks 8, 10 and 14 while continuing AZA. Nonresponders considered treatment failures.		inclusion and partial Mayo response only at week 8)	treatment only)	

Table 7: Trial design characteristics of included paediatric population clinical effectiveness studies

Trial identifier (NCT number), primary publication details	Trial design	Inclusion / exclusion criteria	Geographical location of study sites	Treatment groups and numbers randomised	Open label escape allowance?	Primary outcome	Assessment of induction	Assessment of maintenance	Study sponsor
Infliximab									
Hyams (NCT00336492, C0168T72) Hyams *et al.*, 2012[53]	Randomised, Multicenter, Open-Label Study Phase III	Inclusion: Patients 6–17 years old, with moderately to severely active UC, Mayo score 6–12 points and endoscopy subscore ≥ 2), failed to respond to adequate treatment/experienced medical complications/adverse effects from 5-ASAs immunomodulators (6-MP/AZA) or oral/i.v. corticosteroids. Exclusion: Patients with acute severe extensive UC and those who previously used other investigational drugs or any TNF antagonist	USA, Netherlands, Canada and Belgium 23 study centres	All patients received induction regimen of IFX 5mg/kg at weeks 0, 2 and 6. 45 patients who achieved clinical response at week 8 (primary endpoint) randomised to receive: IFX 5 mg/kg q8w = 22 IFX 5 mg/kg q12w = 23	Yes Patients losing response during maintenance eligible to increase IFX dose and/or frequency to set regimens. IFX 5mg/kg q8w group to 10 mg/kg q8w. IFX 5 mg/kg q12w group to 10 mg/kg q8w, with those losing response between 8 and 12 weeks after previous infusion to 5 mg/kg q8w.	Clinical response at week 8	All patients received induction with IFX 5mg/kg at week 0, 2 and 6 and clinical response and clinical remission assessed at week 8 (no PBO control for induction).	Week 54 (no PBO control for maintenance)	Janssen Research & Development

5.2.2.2 Quality of included evidence

All of the included trials were considered to be at low risk of selection bias as all trials reported an appropriate method for generating the randomisation sequence. Likewise, the majority of trials reported adequate information that allocation was concealed and were considered to be at low risk of bias for this domain. This was with the exception of two trials where there was no information reported to make a judgement. These trials were therefore classified as being at unclear risk of bias (Hyams *et al.*, Suzuki *et al.*). Eight of the ten trials were considered to be at low risk of performance bias because there was reporting to indicate that participants and personnel were blinded to participants' treatment allocation. Two trials were considered at unclear risk of bias for this domain; one because there was no clear statement in the trial report (Hyams *et al.*) and one because the treatment regimen differed for non-responders at week 8 in AZA arm which could break the blinding (UC-SUCCESS). Blinding of the outcome assessment was reported by five trials, all of which were considered at low risk of bias for this domain. The remaining five trials included no statement in the trial report and were considered at unclear risk for this domain (ACT1, ACT2, Hyams *et al.*, UC-SUCCESS, Suzuki *et al.*).

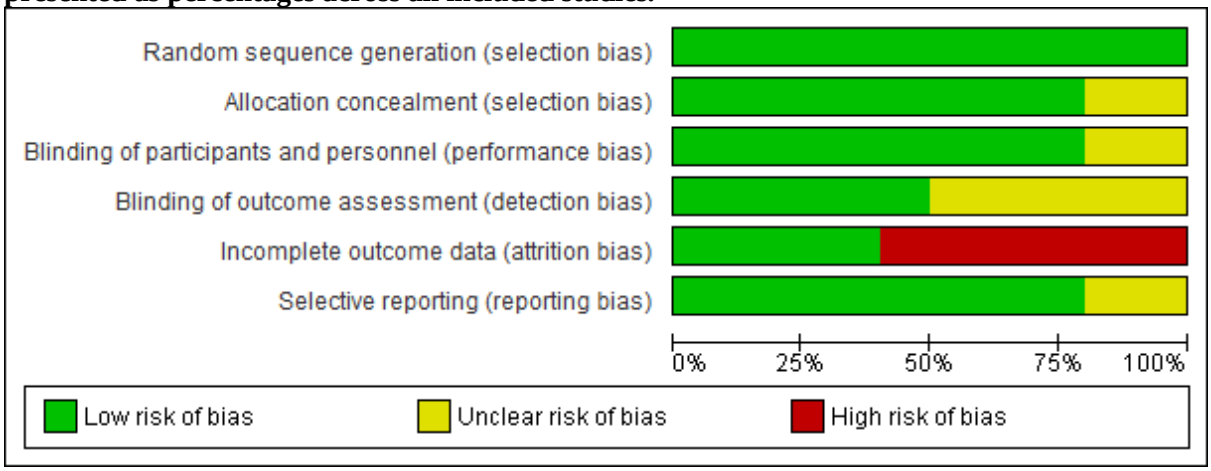
All included trials were reported according to the intention-to-treat (ITT) principle. However, for two trials (ACT1 and ACT2), although ITT was reported, >50% patients in the placebo group and >30% patients in infliximab groups did not complete the trial. Similarly, in another infliximab trial (Hyams *et al.*), the numbers of patients withdrawing from the study were unbalanced across groups with >50% patients withdrawing from the every 12 week dose group. In the UC-SUCCESS trial, there was also a high level of attrition and an imbalance between treatment groups (AZA, 34%; IFX, 18%; IFX/AZA, 21%). In one of the adalimumab trials (ULTRA2), although ITT analysis was undertaken, there was a high level of attrition and an imbalance between treatment groups (placebo, 50%; adalimumab, 59%). In the golimumab maintenance trial (PURSUIT-Maintenance), withdrawal >10% was evident across all treatment groups. These trials were all considered to be at high risk of bias for this domain. Of note, the trial of adalimumab reported by Suzuki *et al.* was considered at low risk of attrition bias for the induction phase (see Figures 3 and 4). A high risk of attrition bias was evident for the maintenance phase (>10% withdrawing). The maintenance active treatment group comprised participants receiving both licenced and unlicensed doses of adalimumab during induction (data not used in this report). Details of the numbers of participants withdrawing and reasons for withdrawal by trial are presented in Appendix 4. The extent of reporting of the reasons for withdrawals was variable between studies.

Selective outcome reporting was assessed based on ClinicalTrials.gov records, trial protocols and Clinical Study Reports (CSRs) where provided by the manufacturers. Adequate data were available across ClinicalTrials.gov records and CSRs (available for some trials only) to compare outcomes with those reported in the associated peer-review publications for all included trials with the exception of

Probert *et al*. Stated primary outcomes were compared between published reports and trial protocols (for those RCTs where trial protocols were provided by manufacturers). Where trial protocols were available, stated primary outcomes were found to be consistent with published reports, [redacted] With the exception of Probert *et al.* and Suzuki *et al.*, all included RCTs were considered to be of at low risk of bias for this domain. Probert *et al.* and Suzuki *et al.* were judged as being at unclear risk of selective reporting bias.

Figure 3: Risk of bias summary - Review authors' judgements about each risk of bias item for each included study

Figure 4: Risk of bias graph - Review authors' judgements about each risk of bias item presented as percentages across all included studies.



5.2.2.3 Population characteristics

The baseline characteristics of participants in the included RCTs are presented in Tables 8 and 9. In addition to comparator arm data, only data relating to licensed doses of interventions are presented.

a) Population: Adults aged 18 years and over with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies

Mean and median reported ages of participants were considered consistent across included adult population trials, ranging from 37 to 42.5 years. Mayo scores at baseline were also consistent across trials and spanned from 8.1 to 8.9. Average proportions of male participants ranged from 41% to 73% and the majority of included patients (where reported) were Caucasian in ethnicity (79.5% to 95.9%), with the exception of the Suzuki *et al.* study, which included exclusively Japanese patients. Mean and median disease duration of participants ranged from 59 months (4.9 years) to 8.5 years. Conventional UC medications at baseline were variable between the included trials. In no included study had all participants previously been trialled on corticosteroids and AZA or 6-MP, as required by the wording used in the final NICE scope[37] population and the wording of the EMA licensing for infliximab, adalimumab and golimumab. Whilst it is noted that AZA and 6-MP may be used more typically in clinical practice as maintenance therapies, due to their longer initiation of effect, it is therefore debatable whether the included trial populations would represent patients who had failed or were intolerant to previous conventional therapies. All trials related to anti-TNF-α naïve populations, with the exception of ULTRA2 (which permitted the inclusion of anti-TNF-α experienced patients) and PURSUIT-Maintenance (in which patients responding to prior golimumab induction therapy were randomised to golimumab maintenance regimens or placebo). Data at induction were reported according to anti-TNF-α experience. Data relating to patients who were anti-TNF-α naive for

maintenance time points were requested and received from the manufacturer of adalimumab (Abbvie). Weight and smoking status were both relatively poorly reported across included studies.

b) Population: Children and adolescents aged 6 to 17 years (inclusive) with severely active UC, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies

The included trial population (Hyams *et al*) averaged 15 years of age in both treatment groups, was 43.5 to 45.5% male and mostly Caucasian in ethnic origin (82.6% to 90.9%) with average disease duration of 1.1 to 1.8 years. Patients had a median Mayo score of 7.5 to 8.0 and a median PUCAI score of 50 to 57.5, where a PUCAI of score of ≥ 65 would indicate severe disease (and therefore were a mixture of patients with moderate and severe disease, whilst infliximab is licensed in paediatric patients with severe disease only). Participants were required to have had prior use of at least one conventional therapy (with 61% to 64% receiving corticosteroids, 50% to 57% immunosuppressants and 46% to 52% aminosalicylates at baseline).

Table 8: Population characteristics of included clinical effectiveness studies in adults

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
Adalimumab									
ULTRA1	Mean NR Median (range) PBO = 37 (18-72) ADA 160/80mg = 37 (18-75)	PBO = 63.1% ADA 160/80mg = 63.8%	NR	PBO = Extensive colitis, 73/130 (56.2%); left-sided colitis, 42/130 (32.3%); other, 15/130 (11.5%), duration median (range) 5.35 (0.3-34.1) ADA 160/80mg = Extensive colitis, 60/130 (46.2%); left-sided colitis, 61/130 (46.9%); other, 9/130 (6.9%), duration median (range) 6.06 (0.2-34.4)	PBO = 8.7 (1.56) ADA 160/80mg = 8.8 (1.61)	Mean NR Median (range) PBO = 0.32 ADA 160/80mg = 0.33	PBO = CS, 55/130 (41.5%); IMM, 18/130 (13.8%); CS +IMM, 34/130 (26.1%); no CS or IMM, 23/130 (17.7%); aminosalicylates, 98/130 (75.4%) ADA 160/80mg = CS, 48/130 (36.9%); IMM, 28/130 (21.5%); CS + IMM, 23/130 (17.7%); no CS or IMM, 31/130 (23.8%); aminosalicylates, 105/130 (80.8%) Dosages of concomitant UC medication(s) stable throughout study	PBO = 78.7 (17.4) ADA 160/80mg = 75.5 (14.2)	NR
ULTRA2	PBO = 41.3 (13.22) ADA 160/80/40 mg = 39.6 (12.47)	PBO = 152/246 (61.8%) ADA 160/80/40mg = 142/248 (57.3%)	NR	PBO = Pancolitis, 120/248 (48.8%); descending colon, 96/246 (39.0%); other, 30/246 (12.2%), duration 8.5 (7.37) ADA 160/80/40mg = Pancolitis, 120/248 (48.4%);	PBO = 8.9 (1.75) ADA 160/80/40mg = 8.9 (1.50)	PBO = 1.3 ADA 160/80/40 mg = 1.5	PBO = Corticosteroids 140/246 (56.9%); Azathioprine/6-MP 80/246 (32.5%); Aminosalicylates 155/246 (63.0%); Azathioprine/6-MP and/or steroids 175/246 (71.1%); Azathioprine/6-MP and steroids 45/246 (18.3%) (41.1%)	PBO = 77.1 (17.31) ADA 160/80/40 mg = 75.3 (17.71)	NR

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
				descending colon, 96/248 (38.7%); other, 32/248 (12.9%), duration 8.1 (7.09)			ADA 160/80/40mg = Corticosteroids 150/248 (60.5%); Azathioprine/6-MP 93/248 (37.5%); Aminosalicylates 146/248 (58.9%); Azathioprine/6-MP and/or steroids 193/248 (77.8%); Azathioprine/6-MP and steroids 50/248 (20.2%) Prior anti-TNF treatment PBO = 101/246 (41.1) ADA 160/80/40mg = 98/248 (39.1) Concomitant UC medications held stable except steroids, which could be tapered after week 8 at discretion of investigator for patients with satisfactory clinical response		
Suzuki *et al.*, 2014	PBO = 41.3 (13.6) ADA 160/80/40 mg = 42.5 (14.6)	PBO = 70/96 (72.9%) ADA 160/80/40mg = 61/90 (67.8%)	Exclusively Japanese	PBO = Pancolitis, 59/96 (61.5%); descending colon, 35/96 (36.5); other , 2/96 (2.1%); duration 7.8 (6.6) ADA 160/80/40mg = Pancolitis, 63/90 (70.0%); descending colon,	PBO = 8.5 (1.6) ADA 160/80/40mg = 8.6 (1.4)	Mean NR Median (range) PBO = 0.34 (0.05-8.72) ADA 160/80/40	PBO = 5-ASAs, 89/96 (92.7%); Immunomodulators (AZA, 6-MP), 52/96 (54.2%); systemic corticosteroids, 58/97 (60.4%) ADA 160/80/40mg = 5-ASAs, 83/90 (92.2%); Immunomodulators (AZA, 6-MP), 41/90 (45.6%); systemic corticosteroids, 57/90 (63.3%)	PBO = 60.8 (14.1) ADA 160/80/40 mg = 60.1 (12.3)	Tobacco non-smoker PBO = 55/96 (57.3%) ADA 160/80/40mg =

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
				27/90 (30.0); other 0/96 (0%); duration 7.8 (7.1)		mg = 0.22 (0.05-6.28)	Changes in doses of UC concomitant medications not permitted during study (other than CSs). After 8 weeks, patient responders permitted to taper CS dose.		50/90 (55.6%)
Golimumab									
PURSUIT-SC (all randomised patients) Sandborn *et al.*, 2014a	PBO = 39.0 (13.04) GOL 200/100mg = 40.0 (13.54)	PBO = 175/331 (52.9%) GOL 200/100mg = 180/331 (54.4%)	PBO = 263/331 (79.5%) GOL 200/100mg = 271/331 (81.9%)	PBO = n = 330, Limited to left side of colon, 188/330 (57.0); extensive = 142/330 (43.0); duration 6.0 (6.65) GOL 200/100mg = n = 331, Limited to left side of colon, 193/331 (58.3); extensive = 138/331 (41.7), duration 6.4 (6.17)	PBO = 8.3 (1.50) GOL 200/100mg = 8.6 (1.53)	PBO = 1.1 GOL 200/100mg = 1.1	PBO = Patients receiving any UC medication (%)= 310/331 (93.7), corticosteroid (excl budesonide)= 134/331 (40.5) , ≥ 20mg/d PEq= 78/331 (23.6) , < 20mg/d PEq= 56/331 (16.9), Budesonide = 8/331 (2.4), Immunomodulatory drugs= 106/331 (32.0), 6-MP/AZA= 102/331 (30.8) , MTX= 4/331 (1.2), Aminosalicylates= 276/331 (83.4) GOL 200/100mg = Patients receiving any UC medication (%)= 302/331 (91.2), corticosteroid (excl budesonide)= 142/331 (42.9) , ≥ 20mg/d PEq= 85/331 (25.7) , < 20mg/d PEq= 57/331 (17.2), Budesonide = 6/331 (1.8), Immunomodulatory drugs= 105/331 (31.7), 6-MP/AZA= 100/331 (30.2),	NR	NR

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
							MTX= 5/331 (1.5), Aminosalicylates= 270/331 (81.6) Patients maintained stable doses of concomitant UC treatments during study		
PURSUIT-M (randomised patients) Sandborn *et al.*, 2014b	PBO = 40.2 (14.05) GOL 50mg = 41.4 (13.84) GOL 100mg = 39.1 (13.11)	PBO = 75/156 (48.1%) GOL 50mg = 77/154 (50.0%) GOL 100mg = 89/154 (57.8)	PBO = 137/156 (87.8) GOL 50mg = 138/154 (89.6) GOL 100mg = 130/154 (84.4)	Disease extent/location NR PBO = Mean 6.9 (6.96), Median 4.2 (IQR NR) GOL 50mg = Mean 6.8 (6.93), Median 4.5 (IQR NR) GOL 100mg = Mean 7.2 (7.04), Median 4.8 (IQR NR)	PBO = 8.3 (1.37) GOL 50mg = 8.1 (1.38) GOL 100mg = 8.5 (1.34)	PBO = 1.0 GOL 50mg = 0.9 GOL 100mg = 0.9	PBO = Any UC medication= 148 (94.9), corticosteroid=83 (53.2) (excl budesonide), ≥ 20mg/day PEq= 59 (37.8), < 20mg/day PEq= 24 (15.4), budesonide=5 (3.2), Immunomodulatory drugs= 52 (33.3), 6-MP/AZA= 51 (32.7), MTX= 1 (0.6), 5-ASA= 125 (80.1) GOL 50mg = Any UC medication= 144 (93.5), corticosteroid= 77 (50.0), ≥ 20mg/day PEq= 52 (33.8), < 20mg/day PEq= 25 (16.2), budesonide 6 (3.9), Immunomodulatory drugs= 47 (30.5), 6-MP/AZA= 45 (29.2), MTX= 2 (1.3), 5-ASA= 128 (83.1) GOL 100mg = Any UC medication= 143 (92.9), corticosteroid= 79 (51.3), ≥ 20mg/day PEq= 55 (35.7), <	NR	NR

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
							20mg/day PEq= 24 (15.6), budesonide=4 (2.6), Immunomodulatory drugs=48 (31.2), 6-MP/AZA= 48 (31.2), MTX= 0, 5-ASA= 119 (77.3) Patients receiving 5-ASAs/immunosuppressants at baseline of induction studies required ot have held stable doses during induction and maintenance. After induction, patients in clinical response receiving concomitant CSs at baseline of PURSUIT-M required to taper CSs.		
Infliximab									
ACT1 Rutgeerts *et al.*, 2005	PBO = 41.4 (13.7) IFX 5mg/kg = 42.4 (14.3)	PBO = 72/121 (59.5%) IFX 5mg/kg = 78/121 (64.5%)	PBO = 111/121 (91.7) IFX 5mg/kg= 116/121 (95.9)	PBO = left side, 66/120 (55.0%); extensive, 54/120 (45.0%); duration, 6.2 (5.9) IFX 5mg/kg = left side, 63/121 (52.9%); extensive, 56/119 (47.1%); duration, 5.9 (5.4)	PBO = 8.4 (1.8) IFX 5mg/kg = 8.5 (1.7)	PBO = 1.7 (2.7) IFX 5mg/kg = 1.4 (1.9)	PBO = CS, 79 (65.3%); ≥ 20mg/day, 54 (44.6%); 5-ASA, 85 (70.2%); IMM, 53 (43.8%); AZA, 36 (29.8%); MP, 17 (14.0%). CS-refractory disease, 38 (31.4%) IFX 5mg/kg = CS, 70 (57.9%); ≥ 20mg/day, 45 (37.2%); 5-ASA, 82 (67.8%); IMM, 66 (54.5%); AZA, 45 (37.2%); MP, 21 (17.4%). CS-refractory disease, 36 (29.8%)	PBO = 76.8 (16.2) IFX 5mg/kg = 80.0 (17.8)	Current smoker PBO = 7/121 (5.8%) IFX 5mg/kg = 2/121 (1.7)

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
							Doses of concomitant medications kept stable apart from CS, tapered by 5mg/wk after week 8 until dose of 20mg/day reached, thereafter dose reduced by 2.5mg/wk until discontinuation		
ACT2 Rutgeerts *et al.*, 2005	PBO = 39.3 (13.5) IFX 5mg/kg = 40.5 (13.1)	PBO = 71/123 (57.7%) IFX 5mg/kg = 76/121 (62.8%)	PBO = 117/123 (95.1) IFX 5mg/kg= 116/121 (95.9)	PBO = left side, 70/120 (58.3%); extensive, 50/120 (41.7%); duration, 6.5 (6.7) IFX 5mg/kg = left side, 70/118 (59.3%); extensive, 48/118 (40.7%); duration, 6.7 (5.3)	PBO = 8.5 (1.5) IFX 5mg/kg = 8.3 (1.5)	PBO = 1.6 (2.9) IFX 5mg/kg = 1.3 (2.3)	PBO = CS, 60 (48.8%), ≥ 20mg/day, 43 (35.0%), 5-ASA, 89 (72.4%), IMM, 54 (43.9%), AZA, 35 (28.5%), MP, 19 (15.4). CS-refractory disease, 36 (29.3) IFX 5mg/kg = CS, 60 (49.6%), ≥ 20mg/day, 40 (33.1%), 5-ASA, 92 (76.0%), IMM, 52 (43.0%), AZA, 41 (33.9%), MP, 11 (9.1). CS-refractory disease, 35 (28.9) Doses of concomitant medications kept stable apart from CS, tapered by 5mg/wk after week 8 unwil dose of 20mg/day reached, thereafter dose reduced by 2.5mg/wk until discontinuation	PBO = 76.1 (17.4) IFX 5mg/kg = 78.4 (17.8)	Current smoker PBO = 6/123 (4.9%) IFX 5mg/kg = 8/121 (6.6%)
Probert *et al.*, 2003	Mean NR Median (IQR) PBO = 40	NR	NR	PBO = extensive UC, 13/20, left-side, 3/20, distal colitis, 4/20; median (IQR) duration 59 (35-96)	Mean (SD) UC severity score PBO = 8.5 (2)	PBO = 12 (10) IFX 5mg/kg = 9 (9)	PBO = AZA use, 7/20 (35%); Prednisolone equivalent (mg/day), mean 28 (7 SD), median 30 (IQR: 25 to 30); duration of steroid treatment (days), median 28 (IQR: 14 to	Mean NR Median (IQR) PBO = 72	NR

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
	(29 -43.5) IFX 5mg/kg = 41 (35.5-50.5)			months IFX 5mg/kg = extensive UC, 14/23, left-side 5/23, distal colitis 4/23; median (IQR) duration 75 (39-141) months	IFX 5mg/kg = 8 (2) Mean (SD) Baron score PBO = 2.4 (0.5) IFX 5mg/kg = 2 (0.5)		45) 28 (11.5–42) IFX 5mg/kg = AZA use, 6/23 (26%); prednisolone equivalent (mg/day), mean 32 (11 SD), median 30 (IQR: 30 to 30); duration of steroid treatment (days), median 28 (IQR: 11.5 to 42) Doses of 5-ASA and AZA/6-MP kept stable during study. Glucocorticoids kept stable during screening then permitted to be changed "according to clinical demands", with goal of reducing daily dose by 5mg prednisolone equivalent per week.	(60-8 as reported) IFX 5mg/kg = 66 (61-78)	
UC-SUCCESS Panaccione *et al.*, 2014	AZA = 40.7 (13.2) IFX 5mg/kg = 38.5 (12.7) IFX 5mg/kg + AZA = 38.0	AZA = 33/79 (41%) IFX 5mg/kg = 42/78 (54%) IFX 5mg/kg + AZA = 48/80 (60%)	NR	Disease extent/duration NR Disease duration AZA = 6.6 (7.8) IFX 5mg/kg = 6.3 (6.5) IFX 5mg/kg + AZA = 5.2 (5.1)	AZA = 8.5 (1.4) IFX 5mg/kg = 8.1 (1.4) IFX 5mg/kg + AZA = 8.6 (1.3)	NR	AZA = Corticosteroid use, 27/79 (34.2%); prior immunomodulatory therapy, 8/79 (10%) IFX 5mg/kg = Corticosteroid use, 31/78 (39.7%); prior immunomodulatory therapy, 8/78 (10.3%) IFX 5mg/kg + AZA = Corticosteroid use, 38/80 (47.5%); prior	NR	NR

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease extent/location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, mean mg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
	(12.2)						immunomodulatory therapy, 8/80 (10.0%) Baseline concomitant treatments kept stable during study. Patients receiving CSs at baseline tapered to 0mg by week 14 unless medically contraindicated.		

Table 9: Population characteristics of included paediatric population clinical effectiveness studies

Trial identifier (NCT number), primary publication details	Age, mean years (SD)	Male participants (%)	Ethnicity, Caucasian (%)	Disease location and disease duration, mean years (SD)	Disease severity at baseline – Mayo score, mean (SD)	CRP at baseline, meanmg/dl	Medications at baseline	Weight kg, mean (SD)	Smoking status
Hyams *et al.*, 2012 (NCT0033649, C0168T72)	Mean NR Median (IQR) Maintenance IFX 5mg/kg q8w = 15.0 (12.0-16.0) Maintenance IFX 5mg/kg q12w = 15.0 (12.0-16.0)	Maintenance IFX 5mg/kg q8w = 10/22 (45.5%) Maintenance IFX 5mg/kg q12w = 10/23 (43.5%)	Maintenance IFX 5mg/kg q8w = 20/22 (90.9%) Maintenance IFX 5mg/kg q12w = 19/23 (82.6%)	Maintenance IFX 5mg/kg q8w = Left side of colon, 6/22 (27.3%); extensive, 16/22 (72.7%); duration median (IQR) 1.8 (0.6-2.4) Maintenance IFX 5mg/kg q12w = Left side of colon, 4/23 (17.4%); extensive, 19/23 (82.6%); duration median (IQR) 1.1 (0.6-1.9)	Mean NR Median (IQR) Maintenance IFX 5mg/kg q8w = 7.5 (7.0-9.0); median PUCAI (IQR) 50.0 (35.0-55.0) Maintenance IFX 5mg/kg q12w = 8.0 (7.0-10.0); median PUCAI (IQR) 57.5 (50.0-65.0)	Mean NR Median (IQR) Maintenance IFX 5mg/kg q8w = 0.3 (0.3-1.5) Maintenance IFX 5mg/kg q12w = 0.3 (0.3-2.2)	Maintenance IFX 5mg/kg q8w = At least 1 concomitant medication, 22/22 (100%); corticosteroids (parenteral or oral), 14/22 (63.6%); ≤1mg/kg prednisone-equivalent, 10/22 (45.5%); ≤1mg/kg prednisone-equivalent, 4/22 (18.2%); corticosteroids (budesonide), 1/22 (4.5%); corticosteroids (rectal), 2/22 (9.1%); immunomodulatory agents, 11/22 (50.0%); 6-MP/AZA, 10/22 (45.5%); MTX, 1/22 (4.5%); aminosalicylates, 10/22 (45.5%); antibiotics, 0/22 (0%) Maintenance IFX 5mg/kg q12w = At least 1 concomitant medication, 23/23 (100%); corticosteroids (parenteral or oral), 14/23 (60.9%); ≤1mg/kg prednisone-equivalent, 10/23 (43.5%); >1mg/kg prednisone-equivalent, 4/23 (17.4%); corticosteroids (budesonide), 0/23 (0%); corticosteroids (rectal), 1/23 (4.3%); immunomodulatory agents, 13/23 (56.5%); 6-MP/AZA, 11/23 (47.8%); MTX, 2/23 (8.7%); aminosalicylates, 12/23 (52.2%); antibiotics, 0/23 (0%) UC therapies to remain stable, CS could be tapered if clinically indicated	Maintenance IFX 5mg/kg q8w = 51.54 (18.294) [median 50.40 (range 26.2 to 91.6, IQR 36.10 to 61.50)] Maintenance IFX 5mg/kg q12w = 52.80 (16.855) [median 52.30 (range 24.5 to 86.4 , IQR 40.30 to 68.60)]	NR

5.2.3 Assessment of effectiveness

5.2.3.1 Population: Adults aged 18 years and over with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies

a) Rates of and duration of response, relapse and remission

Clinical response and remission data were well reported across the included trials for the adult population. It was assumed by the Assessment Group that the numbers of patients who were reported in the trial publications as being in clinical response also included those patients who were in clinical remission. Data relating to transitions of patients between no response, response and remission categories at maintenance time points were requested and received from the manufacturers (MSD and Abbvie). The induction trial data (as reported in the trial publications) and maintenance transition data (received from the manufacturers) from eligible trials were analysed using NMA methods (see Section 5.2.3.3). Definitions of clinical response and clinical remission used in the included trials are presented in Table 10.

Table 10: Definitions of clinical response and remission in adult population RCTs included in the clinical effectiveness systematic review

Trial	Definition of clinical response	Definition of clinical remission	Measurement time points
ACT1 [50]	Decrease from baseline in total Mayo score of ≥ 3 points and ≥ 30%, with accompanying decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore of 0 or 1	Total Mayo score of ≤ 2 points, with no individual subscore > 1	Clinical response and remission assessed at weeks 8, 30 and 54
ACT2 [50]	As above	As above	Clinical response and remission assessed at weeks 8 and 30
Probert *et al.*, 2003 [51]	No definition of response, mean UC "severity score" and improvement reported only	UC symptom score (UCSS) (i.e. Mayo score) Clinical remission = UCSS ≤ 2	Outcomes reported at week 2 and 6
UC-SUCCESS [52]	Decrease in total Mayo score of ≥ 3 points and ≥ 30% decrease from baseline Mayo score	CS-free remission= total Mayo score of ≤ 2 points, with no individual subscore >1 point without the use of CSs	Mayo scores assessed at weeks 0, 8 (partial Mayo) and 16
ULTRA1 [45]	Decrease in Mayo score of ≥ 3 points and ≥ 30% from baseline plus decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore of 0 or 1	Mayo score of ≤ 2 points, with no individual subscore > 1	Mayo scores recorded at weeks 0 and 8
ULTRA2 [46]	Decrease from baseline in total Mayo score of ≥ 3 points and ≥ 30% plus decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore of 0 or 1	Total Mayo score of ≤ 2 points, with no individual subscore > 1	Clinical response and remission measured at weeks 8, 32 and 52/early termination
Suzuki *et al.*, 2014 [47]	Decrease from baseline in total Mayo score of ≥ 3 points and ≥ 30%, with accompanying decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore ≤ 1	Total Mayo score of ≤ 2 points, with no individual subscore > 1	Clinical response and remission assessed at weeks 8, 32 and 52
PURSUIT-SC [48]	Decrease from baseline in Mayo score of ≥ 3 points and ≥ 30% plus decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore of 0 or 1	Mayo score of ≤ 2 points, with no individual subscore > 1	Mayo scores recorded at weeks 0 and 6
PURSUIT-Maintenance [49]	Decrease from baseline value (observed in preceding induction study) in Mayo score of ≥ 3 points and ≥ 30% plus decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore of 0 or 1	Mayo score of ≤ 2 points, with no individual subscore > 1	Mayo scores calculated at weeks 0, 30 and 54

Adalimumab

Four adalimumab trials presented clinical response and remission data (ULTRA1, ULTRA2, ULTRA3 and Suzuki *et al.*).

At week 8, more patients in the adalimumab 160/80mg induction treatment arm of ULTRA1 achieved clinical response (54.6% vs. 44.6%, *p*-value NR) and twice as many reached clinical remission (18.5% vs. 9.2%, $p=0.031$) compared with placebo patients.[45] Subgroup analyses demonstrated that patients with a Mayo score of 10 or above at baseline of ULTRA1 were less likely to achieve remission at week 8 compared with patients with lower baseline Mayo scores.[45] Baseline CRP levels above 10mg/l and baseline weight of 82 kg and above were also linked with lower remission rates in ULTRA1.[45] When baseline prior UC medications were considered, the treatment effect of adalimumab 160/80mg compared with placebo was most pronounced in patients who had received immunomodulator treatment (i.e. AZA/6-MP) at baseline without corticosteroids, and patients who had received no prior aminosalicylates.[45] Clinical response rates at week 8 in the placebo group, when stratified by geographical region, appeared to be higher in Canada and Eastern Europe (compared with US/Puerto Rico and Western Europe) although reasons for this are unclear.[45]

In ULTRA2, patients in the adalimumab 160/80mg induction group were more likely to achieve clinical response (50.4% vs. 34.6%, $p<0.005$) and clinical remission (16.5% vs. 9.3%, $p<0.05$) at week 8 than in the placebo group.[46] Similarly, among patients who had received no prior anti-TNF-α treatment, greater proportions of patients in the adalimumab 160/80mg induction group reached clinical response (50.5% vs. 38.6%, $p<0.005$) and clinical remission (16.5% vs. 11.0%, $p<0.05$) at week 8 than placebo-treated patients.[46] Patients receiving adalimumab as maintenance therapy in ULTRA2 were also more likely at week 52 to be in clinical response (30.2% vs. 18.3%, $p<0.05$) or clinical remission (17.3% vs. 8.5%, $p<0.005$) than subjects in the placebo group.[46] Anti-TNF-α -naïve adalimumab-treated patients were also more likely to achieve clinical response (36.7% vs. 24.1%, $p=0.019$) or remission (22.0% vs. 12.4%, $p=0.029$) at week 52 than those in the placebo group.[46] Patients in the adalimumab group were more likely to achieve sustained response (ITT 21.8%, anti-TNF-α -naïve 26.7, both $p<0.05$ vs. placebo) and sustained remission (ITT 8.1%, anti-TNF-α -naïve 10.7%, both $p<0.05$) compared with placebo group subjects (sustained response: ITT 11.4%, anti-TNF-α -naïve 16.6%; sustained remission ITT 2.4%, anti-TNF-α -naïve 3.4%).[58] At week 52 of ULTRA2, corticosteroid-free remission was achieved by more adalimumab group patients versus placebo (both $p<0.05$).[59] A *post hoc* analysis of ULTRA2 data at week 52 demonstrated that mean days in clinical response (134.58 vs. 94.55, $p<0.001$) and mean days in clinical remission were also greater for adalimumab-treated patients (85.32 vs. 52.87, $p<0.001$).[60] For patients with no prior anti-TNF-α use, stool frequency and rectal bleeding Mayo subscores of 1 or below at week 8 were most likely to be achieved in patients receiving adalimumab than placebo (both $p<0.05$).[46] At week 52, the proportions of patients who had discontinued corticosteroid use and achieved sustained clinical remission at both weeks 32 and 52 (among patients with baseline corticosteroid use) were 10.0% and 1.2% in the adalimumab (no prior anti-TNF-α use) and placebo groups respectively ($p=0.014$).[46] At

week 52, for patients with no prior anti-TNF-α use, 20.3% of the adalimumab group and 6.2% of the placebo group were in corticosteroid-free clinical remission ($p<0.05$).[59]

The open-label extension study ULTRA3 presented the proportions of patients who continued to receive adalimumab and were in clinical response (42.6%) and remission (25.6%) at week 52.[61]

Patients who received adalimumab for induction in the Suzuki *et al.* trial were more likely to be in clinical response (50% vs. 35%, $p<0.05$) by week 8 but not clinical remission (10% vs. 11%, *p*-value NR) compared with placebo group patients.[47] At week 8, a statistically significant greater proportion of patients in the adalimumab arm reached a subscore of 1 or below physician's global assessment domain compared against the placebo arm ($p\leq0.05$); differences in the other Mayo subscores were not statistically significant.[47] Within the Suzuki *et al.* trial, greater proportions of adalimumab maintenance-treated patients were in clinical response (31% vs. 18%, $p<0.05$) and clinical remission (23% vs. 7%, $p<0.01$) through week 52 compared with placebo group patients.[47] At week 52, a greater proportion of subjects in the adalimumab group versus placebo experienced subscores of 1 or below for physician's global assessment and stool frequency subscore (both $p\leq0.05$).[47] The proportions of patients in steroid-free clinical remission at week 52 were 14.2% and 6.9% in the adalimumab and placebo arms respectively (*p*-value NR).[47]

Golimumab

In the PURSUIT-SC induction trial,[48] clinical response and remission data were reported for both Phase II and Phase III. By week 6, in the Phase II (plus additional phase II randomised patients) analyses, more patients receiving golimumab were in clinical response (46.5% vs. 37.7%, *p*-value NR) and remission (18.3% vs. 10.1%, *p*-value NR) as compared against the placebo group. Similarly, more golimumab-treated patients achieved clinical response (51.0% vs. 30.3%, $p<0.0001$) and remission (17.8% vs. 6.4%, $p<0.0001$) than placebo-treated patients by week 6 in the Phase III analyses.

In the PURSUIT-Maintenance study,[49] proportions of patients maintaining clinical response (47.0% vs. 31.3%, $p=0.010$) and in clinical remission (33.1% [golimumab 50mg, $p=0.068$], 33.8% [golimumab 100mg, $p=0.011$] vs. 22.1%) through week 54 were larger for the golimumab groups compared with placebo. PURSUIT-Maintenance patients who maintained clinical response and were corticosteroid-free among those who receiving corticosteroids at maintenance baseline were 38.5% in the golimumab 50mg group ($p=0.026$), 30.5% in the golimumab 100mg group ($p=0.138$) and 20.7% in the placebo group.

Infliximab

By week 8 of the ACT1 trial, more patients treated with infliximab 5mg/kg were in clinical response (69.4% vs. 37.2%, $p<0.001$) and remission (38.8% vs. 14.9%, $p<0.001$) than those who received placebo.[50] At week 54, more infliximab group patients were in clinical response (45.5% vs.19.8%, $p<0.001$) and remission (34.7% vs. 16.5%, $p=0.001$) than placebo-treated subjects.[50] Patients who sustained clinical response at weeks 8, 30 and 54 were 38.8% in the infliximab group and 14.0% in the placebo group ($p<0.001$).[50] Proportions of patients who sustained clinical remission at weeks 8, 30 and 54 were 19.8% and 6.6% in the infliximab and placebo treatment arms respectively ($p=0.002$).[50] Of the 5mg/kg infliximab group, 25.7% were in clinical remission and had discontinued corticosteroids at week 54, compared with 8.9% in the placebo group ($p=0.006$).[50]

In ACT2, more patients in the infliximab 5mg/kg group were in clinical response (64.5% vs. 29.3%, $p<0.001$) and remission (33.9% vs. 5.7%, $p<0.001$) at week 8 compared with placebo.[50] By week 30, more 5mg/kg infliximab-treated patients were in clinical response (47.1% vs. 26.0%, $p<0.001$) and remission (25.6% vs. 10.6%, $p=0.003$) compared with placebo.[50] The proportions of patients who sustained clinical response (41.3% vs. 15.4%, $p<0.001$) and clinical remission (14.9% vs. 2.4%, $p<0.001$) at weeks 8 and 30 were also higher in the infliximab 5mg/kg group compared with patients receiving placebo.[50]

No statistically significant differences were observed between the infliximab and placebo treatment groups through week 6 of the Probert *et al.* trial in terms of clinical remission (as defined by a UCSS score of ≤ 2) (39% vs. 30%, $p=0.76$). Remission rates among patients receiving azathioprine were 67% for infliximab and 33% for placebo groups ($p=0.89$).[51]

A greater proportion of patients in the UC-SUCCESS study who received combination treatment with infliximab plus azathioprine were in steroid-free clinical remission at week 16 (39.74%) compared with the infliximab monotherapy (22.08%, p vs. IFX =0.017) and azathioprine monotherapy (23.68%, p vs. IFX =0.813; p vs. IFX/AZA =0.032) groups.[52]

No included trials reported data on rates or duration of relapse.

Data relating to clinical response and remission are summarised in Table 11.

Table 11: Summarised clinical response and remission data from RCTs in adults

Study name	Treatment arm	Time point	Rates of and duration of response and relapse	Rates of and duration of remission and relapse
ULTRA1	PBO	Week 8	Clinical response: 58/130 (44.6%) (*p* value NR)	Clinical remission, ITT-A3 protocol: 12/130 (9.2%)
ULTRA1	ADA 160/80mg	Week 8	Clinical response: 71/130 (54.6%)	Clinical remission, ITT-A3 protocol: 24/130 (18.5%), *p*-value vs. PBO =0.031
ULTRA2	PBO	Week 52	Patients with response, 45/246 (18.3%) No prior anti-TNF: Clinical response 35/145 (24.1%) Prior anti-TNF Clinical response 10/101 (9.9%)	Patients with remission, 21/246 (8.5%); No prior anti-TNF: Clinical remission 18/145 (12.4%) Prior anti-TNF Clinical remission 3/101 (3.0%)
ULTRA2	ADA160/80mg	Week 52	Patients with response, 75/248 (30.2%) No prior anti-TNF: Clinical response 55/150 (36.7%) Prior anti-TNF Clinical response 20/98 (20.4%)	Patients with remission, 42/248 (17.3%); No prior anti-TNF: Clinical remission 33/150 (22.0%) Prior anti-TNF Clinical remission 10/98 (10.2%)
Suzuki	PBO	Week 8	Full Mayo Score Response, 34/96 (35%)	Full Mayo Score Remission, 11/96 (11%)
Suzuki	ADA160/80mg	Week 8	Full Mayo Score Response, 45/90 (50%); *p*-value vs. PBO ≤ 0.05	Full Mayo Score Remission, 9/90 (10%)
Suzuki	PBO	Week 52	Full Mayo Score Response, 17/96 (18%)	Full Mayo Score Remission, 7/96 (7%)
Suzuki	ADA80/40mg or ADA160/80mg to week 8 then ADA40mg EOW	Week 52	Full Mayo Score Response, 55/177 (31%); *p*-value vs. PBO, ≤ 0.05	Full Mayo Score Remission, 41/177 (23%); *p* vs. PBO, ≤ 0.01
PURSUIT-SC	Phase III PBO	Week 6	Phase III. PBO. Proportion with clinical response, 76/251 (30.3%)	Phase III: Clinical remission, 16/251 (6.4%)
PURSUIT-SC	Phase III GOL 200/100mg phase III	Week 6	Phase III. GOL 200/100mg. Proportion with clinical response, 129/253 (51.0%) (*p*<0.0001)	Phase III: Clinical remission GOL 200/100, 45/253 (17.8) (*p*<0.0001)
PURSUIT-Maintenance	PBO	Week 54	Proportion of patients maintaining clinical response: 31.2%, N=154	Clinical remission, 34/154 (22.1%)
PURSUIT-Maintenance	GOL 50mg	Week 54	Proportion of patients maintaining clinical response: 47.0%, N=151 (*p*=0.010)	Clinical remission. 50/151 (33.1%) (*p*=0.068)
PURSUIT-Maintenance	GOL 100mg	Week 54	Proportion of patients maintaining clinical response, 49.7%, N=151 (*p*<0.001)	Clinical remission, 51/151 (33.8%) (*p*=0.011)
UC-SUCCESS	AZA	Week 16	Data not available	Patients in steroid-free remission: 18/76 (23.68%); *p*-value between IFX, 0.813; IFX/AZA, 0.032
UC-SUCCESS	IFX mg/kg	Week 16	Data not available	Patients in steroid-free remission at: 17/77 (22.08%); *p*-value between IFX/AZA, 0.017
UC-SUCCESS	IFX/AZA	Week 16	Data not available	Patients in steroid-free remission: 31/78 (39.74%)

Study name	Treatment arm	Time point	Rates of and duration of response and relapse	Rates of and duration of remission and relapse
Probert	PBO	Week 6	Data not available	Patients with UCSS score <2: 6/20 (30%). 95% CI for difference with IFX -19 to 34%; *p*=0.76 When remission rates of patients with total disease in each of the two groups were compared, no significant difference was found (*p*=0.9)
Probert	IFX 5 mg/kg	Week 6	Data not available	Patients with UCSS score <2: 9/23 (39%)
ACT1	PBO	Week 8	Proportion of patients with clinical response, 45/121 (37.2%)	Proportion of patients in clinical remission, 14.9% (18/121)
ACT1	IFX 5 mg/kg	Week 8	Proportion of patients with clinical response, 84/121 (69.4%) (*p*<0.001)	Proportion of patients in clinical remission, 38.8% (47/121) (*p*<0.001)
ACT1	PBO	Week 54	Proportion of patients with clinical response, 24/121 (19.8%)	Proportion of patients in clinical remission, 16.5% (20/121)
ACT1	IFX 5 mg/kg	Week 54	Proportion of patients with clinical response, 55/121 (45.5%) (*p*<0.001)	Proportion of patients in clinical remission, 34.7% (42/121) (*p*=0.001)
ACT2	PBO	Week 8	Proportion of patients with clinical response, 36/123 (29.3%)	Proportion of patients in clinical remission, 5.7% (7/123)
ACT2	IFX 5mg/kg	Week 8	Proportion of patients with clinical response, 78/121 (64.5%) (*p*<0.001)	Proportion of patients in clinical remission, 33.9% (41/121) (*p*<0.001)
ACT2	PBO	Week 30	Proportion of patients with clinical response, 32/123 (26.0%)	Proportion of patients in clinical remission, 10.6% (13/123)
ACT2	IFX 5mg/kg	Week 30	Proportion of patients with clinical response, 57/121 (47.1%) (*p*<0.001)	Proportion of patients in clinical remission, 25.6% (31/121) (*p*=0.003)

Consideration was paid to whether it would be appropriate to conduct meta-analysis using the response and remission outcomes within the trials included in the clinical effectiveness review. It was acknowledged that the adalimumab trials differed from the infliximab and golimumab trials in the method of estimation of Mayo scores, in that the infliximab and golimumab trials were based on the average Mayo scores over a consecutive 3 day diary period and the adalimumab trials included scores based on the worst entry over a consecutive 3 day diary period. However, clinical advisors to the Assessment Group did not anticipate that this difference would preclude a synthesis of the evidence. It was further noted by the Assessment Group that there may be potential issues in the consistency of measurement of Mayo scores and levels of placebo response according to physician experience and geographical location. The comparability of the trial data set in terms of prior UC treatment was improved by the requesting and receipt from the manufacturer of adalimumab of anti-TNF-α-naïve maintenance data from ULTRA2. It should also be noted that the PURSUIT-Maintenance trial re-randomised patients who had previously responded to golimumab induction therapy in two previous trials; the extent of this potential bias on patient outcomes is unclear.

Clinical response and remission at induction and maintenance in eligible adult population trials were analysed using NMAs. The results of these analyses are presented in Section 5.2.3.3. For the sake of brevity, all secondary efficacy and safety outcomes data are presented in Appendices 5 and 6.

b) Measures of disease activity

Adalimumab

At week 8 of the ULTRA1 trial, median changes in CRP from baseline were greater in the adalimumab 160/80mg group compared with placebo (-0.77 vs -0.09mg/l).[45] Patients receiving adalimumab 160/80mg in ULTRA1 were also more likely to achieve scores of 1 or below for the Mayo rectal bleeding (p=0.038) and physician global assessment (p=0.035) subscores.[45] Statistically significant changes from baseline in haemoglobin and red blood cells (both p<0.001), total protein and albumin levels (both p <0.01) were observed in the adalimumab group versus placebo in ULTRA1.[62]

In ULTRA2,[46] greater proportions of patients receiving adalimumab achieved Mayo subscores of 1 or below at week 8 than placebo, although only stool frequency and rectal bleeding were statistically significant at the 5% level. Significantly more adalimumab group patients who had not previously received anti-TNF-α treatment reached a rectal bleeding score of 1 or below than placebo (p<0.001).[46]

Golimumab

At week 6 in the Phase II and Phase III components of PURSUIT-SC, mean changes from baseline in Mayo score were – 2.6 (SD=2.73) and – 1.8 (SD=2.96) (Phase II, p=0.219), followed by– 3.1

(SD=2.90) and – 1.6 (SD=2.53) (Phase III, $p<0.0001$) in the golimumab 200/100mg and placebo arms. Mean changes in CRP concentration (mg/l) at week 6 (Phase III) were – 3.35 (golimumab 200/100mg) and + 1.59 (placebo) ($p<0.0001$).

Infliximab

In ACT1, the proportion of patients at week 8 who were not refractory to corticosteroid therapy was higher in the infliximab group compared with placebo (66.7 vs. 37.9%, $p<0.001$).[50] Proportions of patients not refractory to corticosteroids at week 8 of the ACT2 study were 64.8% for infliximab 5mg/kg and 26.4% for placebo ($p<0.001$).[50] As of week 152 of the extension studies, 20 patients remained, of whom 18 (90.0%) had no or mild disease.

Mean improvements in UC symptom scores (UCSS) at week 6 of the Probert *et al.* study were 4 (SD=3) for both placebo and infliximab groups.[51] The mean reduction in daily dose of glucocorticoid was equivalent to 19mg (SD=15) and 14mg prednisolone (SD=12) in the infliximab and placebo groups respectively ($p=0.037$).[51] No statistically significant changes in CRP levels were observed between infliximab and placebo arm patients.[51]

At week 8 of the UC-SUCCESS trial, 65.79% and 36.84% of the AZA arm, 88.31% and 49.35% of the infliximab arm and 85.90% and 52.56% of the infliximab/azathioprine combination arm achieved partial Mayo score decreases of ≥ 1 and ≥ 2 respectively.[52] Week 8 mean changes from baseline in partial Mayo scores were – 2.81 (SD=2.46), - 3.52 (SD=2.25) and – 4.01 (SD=2.04) for azathioprine, infliximab and combination infliximab/azathioprine.[52] Mean changes in total Mayo score from baseline at week 16 were – 3.00 (baseline 8.50) for azathioprine (p vs. IFX/AZA=0.001), - 4.27 (baseline 8.08) for infliximab (p vs. IFX/AZA=0.001), and – 5.28 (baseline 8.54) for combination infliximab/azathioprine.[52]

c) Mortality

Reported deaths for the included trials are presented in Appendix 5.

Adalimumab

No deaths occurred in the ULTRA1 or ULTRA2 adalimumab trials. Deaths were not reported in Suzuki *et al*.

Golimumab

One death occurred in PURSUIT-SC in the unlicensed 400/200mg golimumab induction treatment arm in a patient receiving concomitant prednisolone 20mg with a case of peritonitis and sepsis after surgical complications related to an ischiorectal abscess and subsequent bowel perforation after

surgery. In PURSUIT-Maintenance, no deaths occurred through week 54 in the placebo arm but one death (from pneumonia and heart failure) occurred after week 54 in a patient who had received placebo induction and maintenance. No deaths were observed in the golimumab 50mg group of PURSUIT-Maintenance; however one death was reported after week 54 (in a patient who received GOL 100/50mg induction and GOL 50mg maintenance) due to heart dysfunction in the presence of pronounced atherosclerosis and stenosis affecting aorta, large arteries and coronary arteries. Three deaths were reported through week 54 of PURSUIT-Maintenance in the golimumab 100mg treatment arm due to malnutrition and sepsis (patient receiving 2mg/kg i.v. golimumab induction); cardiac failure with history of thrombosis (patient receiving golimumab 400/200mg s.c. induction); and disseminated tuberculosis in patient who tested positive for latent TB on induction study entry and was receiving isoniazid at time of event (receiving golimumab 200/100mg s.c. induction). Four deaths were reported after week 54 for the golimumab 100mg group in PURSUIT-Maintenance, including one case of myocardial infarction in patient with history of myocardial infarction (placebo s.c. induction and golimumab 100mg maintenance), two deaths due to gallbladder adenocarcinoma with liver metastasis) and due to sepsis (patients receiving golimumab 2mg/kg i.v. induction and golimumab 100mg maintenance) and one death due to accidental nitrous oxide overdose (in a patient receiving golimumab 200/100mg s.c. induction and GOL 100mg maintenance.

Infliximab

The only reported deaths in any of the included infliximab trials occurred in patients recruited into the ACT trials. No deaths occurred through week 54 in ACT1 and ACT2. After week 54, two patients died in the placebo arm of ACT1 (due to suicide and cerebrovascular accident). After 54 weeks, four patients died who received infliximab in the ACT trials (no dose information available, histoplasmosis four weeks after last infusion, listeria encephalitis three years after last infusion, prostate cancer 3.5 years after last infusion, and natural causes ten months after last infusion).

d) Rates of hospitalisation

A total of four included trials reported hospitalisation data for the adult population (ULTRA1 and ULTRA2 for adalimumab, ACT1 and ACT2 for infliximab, no trials for golimumab).

Adalimumab

In ULTRA1, all reported hospitalisation outcome measure data were lower in the adalimumab 160/80mg group than placebo at week 8, indicating more favourable outcomes for the intervention group, including physician visits (p=0.559), emergency room visits (p-value NR), hospital admissions (p-value NR) and days in hospital (p=0.297). None of these differences were statistically significant.[63] Similarly, for the ULTRA2 trial, hospitalisation-related outcome data were also slightly lower for the adalimumab group compared with placebo at week 52, although this was only statistically significant

for physician visits (physician visits p=0.035, emergency room visits p=0.847, hospital admissions p=0.418 and days in hospital p=0.467).[63] A range of hospitalisation-related measures were also reported for ULTRA1 and ULTRA2 data combined. The all-cause hospitalisation incidence rate was lower for adalimumab than placebo (p=0.047), as was the UC-related hospitalisation incidence rate (p=0.002), with a relative risk for UC-related hospitalisation of 0.48 for adalimumab versus placebo (p <0.001).[64]

Golimumab

No included trials reported hospitalisation data for golimumab.

Infliximab

In the ACT1 and ACT2 trials, hospitalisations through week 54 were reported to be lower for the infliximab 5mg/kg group than placebo (ACT1 p=0.061, ACT2 p=0.009).[65]

e) Rates of surgical intervention (both elective and emergency)

Six included trials in the adult population included information on rates of surgical intervention (ULTRA1 and ULTRA2 for adalimumab, PURSUIT-Maintenance for golimumab, and ACT1, ACT2 and Probert *et al.* for infliximab). No trials reported whether surgical outcomes were elective or emergency in nature.

Adalimumab

In ULTRA1, colectomies to week 8 were lower in the adalimumab 160/80mg group than placebo (1.4% vs. 3.6%, p-value NR, elective or emergency NR). Colectomy rates were very slightly lower through week 52 of ULTRA2 in the adalimumab group (4%) vs. placebo (4.9%) (p-value NR, elective or emergency NR).[64,66]

Golimumab

Limited data were available for golimumab that indicated that only 2-3% of golimumab induction responders re-randomised to golimumab 50mg or 100mg in PURSUIT-Maintenance received colectomy at the end of maintenance.[67]

Infliximab

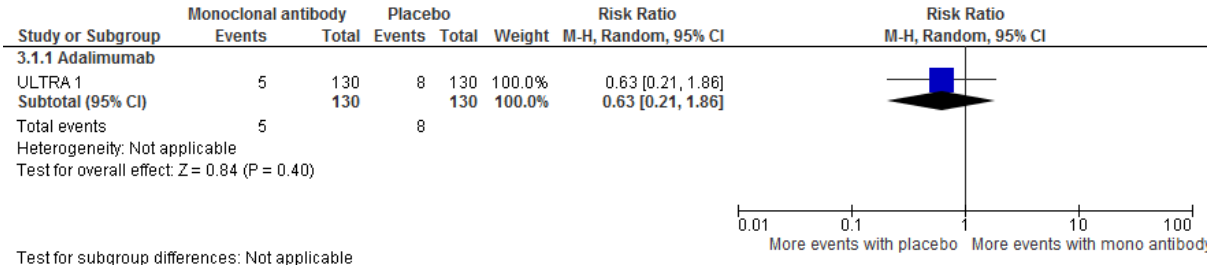
Colectomy and ostomy rates through week 54 of ACT1 were both slightly lower in the infliximab 5mg/kg group (5.8% and 2.5% respectively) than in the placebo group (7.4% and 4.1% respectively) (p-values NR).[65] One patient in each case from the placebo arm was reported as having the outcomes of colectomy and an ostomy (0.7% and 0.7%) through week 54 of ACT2, whilst no patients in the infliximab 5mg/kg group underwent colectomy or ostomy.[65] Limited details were available from the

Probert *et al*. trial to the effect that a single patient in the placebo arm received a colectomy during the intervention period.

Meta-analysis

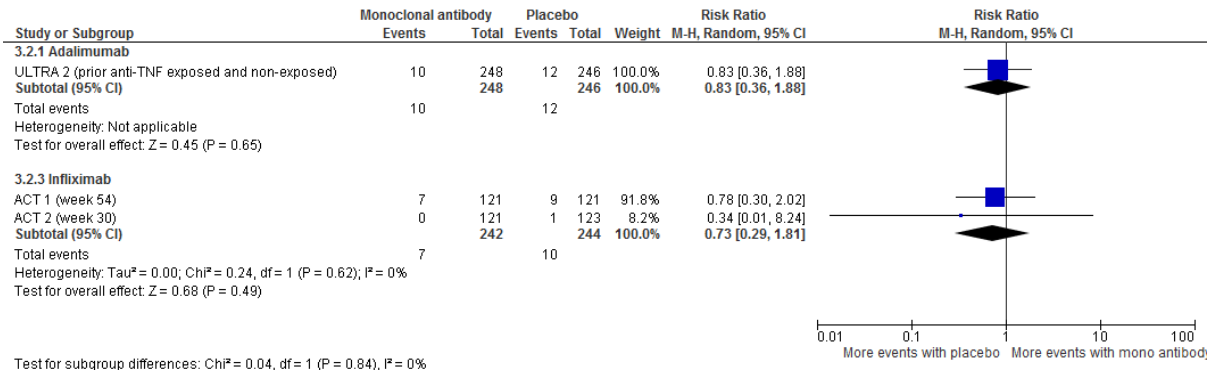
Colectomy rates during induction were reported by one trial (ULTRA1). The between-group difference was not statistically significant (RR=0.63 [random effects] 95% confidence interval, 0.21 to 1.86; *p*=0.40, see Figure 5).

Figure 5: Forest plot of comparison: Colectomy - adults, outcome: Colectomy - induction licenced dose



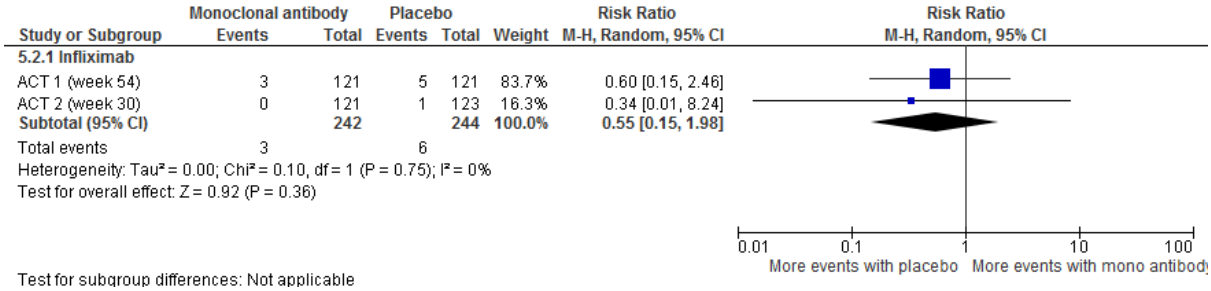
Colectomy rates during maintenance were reported by one trial evaluating the licenced maintenance dose of adalimumab comprising a mixed sample of anti-TNF-α exposed and naïve participants (ULTRA2, 517 participants). The between-group difference was not significant (RR=0.83 [random effects] 95% confidence interval, 0.36 to 1.88; *p*=0.65). Two trials evaluating the licenced maintenance dose of infliximab reported maintenance outcomes at 30 weeks (ACT2) and 54 weeks (ACT1). The pooled effect across these trials (486 participants) was not significant (RR=0.73 [random effects] 95% confidence interval, 0.29 to 1.81; *p*=0.49). The forest plot for these analyses (random effects) is presented in Figure 6.

Figure 6: Forest plot of comparison: Colectomy - adults, outcome: Colectomy - maintenance licenced dose



Ostomy rates during maintenance in adults were reported by two trials evaluating the licenced maintenance dose of infliximab at 30 weeks (ACT2) and 54 weeks (ACT1). The pooled effect across these trials (486 participants) was not significant (RR=0.55 [random effects] 95% confidence interval, 0.15 to 1.98; *p*=0.36). The forest plot for these analyses is presented in Figure 7.

Figure 7: Forest plot of comparison: Ostomy - adults, outcome: Ostomy - maintenance licenced dose



f) Time to surgical intervention (both elective and emergency)

Very limited data were reported from the included trials in the adult population for the outcome of time to surgical intervention. Sandborn *et al.* (2009) combined data from the ACT1 and ACT2 infliximab trials and reported that the cumulative incidence of colectomy through 54 weeks was higher for the placebo group (17%) than for the combined infliximab group (10%) (*p*=0.02) and calculated a hazard ratio of 0.59, indicating a 41% reduction in the risk of colectomy for the combined licensed and unlicensed infliximab groups versus placebo.

g) Health-related quality of life

HRQoL data were available from nine included trials in the adult population (ULTRA1, ULTRA2 and ULTRA3 for adalimumab, PURSUIT-SC and PURSUIT-Maintenance for golimumab, and ACT1, ACT2, UC-SUCCESS and Probert *et al.* for infliximab, see Appendix 7). Data related to HRQoL was measured using IBDQ, SF-36 and the EQ-5D (note - total IBDQ scores can range from 32 [very poor] to 224 [perfect HRQoL]).

Adalimumab

In ULTRA1, the changes from baseline scores to week 8 in IBDQ were very similar for the adalimumab 160/80mg and placebo groups (153 vs. 152, *p*-value NR). Furthermore, the difference between IBDQ mean responses at week 8 in the adalimumab 160/80mg and placebo groups was not statistically significant (70 vs. 75, *p*=0.532). Changes from baseline in SF-36 mental and physical component summary scores were also similar at week 8 in the adalimumab and placebo group (46 vs. 44). ULTRA2 week 52 IBDQ scores were higher in the adalimumab 160/80mg group than placebo, indicating more favourable HRQoL in the adalimumab group (27 vs. 19, *p*<0.05). A greater

proportion of patients experienced an increase in IBDQ of ≥16 points from baseline by week 52 in the adalimumab group compared with placebo (26.2% vs. 16.3%, p <0.05).

Golimumab

In both Phase II and Phase III of the PURSUIT-SC golimumab trial, patients in the golimumab 200/100mg induction arms reported a greater change in IBDQ from baseline to week 6 than placebo groups patients (Phase II, mean 24.9 vs. 14.8 (p-value NS); Phase III mean 27.0 vs. 14.8, p<0.0001). Greater proportions of patients in each golimumab group were also described as achieving "any improvement" to "clinically meaningful improvement" in IBDQ (51.1% vs. 35.2% p<0.001), physical component summary (41.0% vs. 31.6% p=0.01) and mental component summary scores (42.7% vs. 28.5% p<0.001) at week 6.

Infliximab

In the ACT1 trial, changes from baseline in SF-36 physical and mental component summary scores to week 8 were larger for the infliximab 5mg/kg group compared with placebo (both p<0.05). Statistically significant improvements in IBDQ and SF-36 components were evident in the infliximab 5mg/kg treatment arm compared with placebo to week 8 for ACT1 and ACT2 trials combined. The greatest changes from baseline to week 16 in both IBDQ and SF-36 physical function were observed in the infliximab/azathioprine combination treatment arm (p<0.05 vs. AZA, p<0.05 vs. IFX for both outcomes). Improvements in IBDQ and EQ-5D from baseline to week 6 in Probert *et al.* were larger in the infliximab group compared with placebo (p-value NR*).*

h) Adverse events of treatment (including leakage and infections following surgery)

The included trials report data relating to adverse events associated with the interventions under assessment only (i.e. infliximab, adalimumab, and golimumab) and do not report safety outcomes (e.g. leakage and infections) post-surgery. However, whilst the clinical effectiveness systematic review does not take these factors into account, these factors are relevant to the economic analysis (see Section 6).

P-values are provided where available; however the statistical significance of observed differences in safety outcomes was poorly reported across the included trials.

Discontinuations due to adverse events

Adalimumab

Discontinuations due to adverse events at week 8 in ULTRA1 were 5.4% in both adalimumab 160/80mg and placebo groups.[45] Withdrawals due to adverse events were slightly lower for adalimumab than placebo by week 52 of ULTRA2, at 23/257 (8.9%) for adalimumab 160/80mg and 34/260 (13.1%) for placebo.[46] More adverse events leading to discontinuation occurred in the Suzuki

et al. trial in the adalimumab 40mg EOW group versus placebo (N=22 vs. N=6; 22.4/100 patient years [PY] vs. 13.4/100 PY).[47]

Golimumab

Numbers of patients who discontinued study agent through week 6 because of at least one adverse event were relatively low across both golimumab 200/100mg induction (1/331, 0.3%) and placebo (3/330, 0.9%) groups for PURSUIT-SC.[48] Through week 54 of PURSUIT-Maintenance, 8/154 (5.2%) of the golimumab 50mg, 14/154 (9.1%) of the golimumab 100mg and 10/156 (6.4%) of the placebo groups had discontinued study agent due to at least one adverse event.[49]

Infliximab

Through week 54 of ACT1 the numbers of patients with adverse events leading to study drug discontinuation were 10/121 (8.3%) and 11/121 (9.1%) for the infliximab 5mg/kg and placebo groups, respectively. Through week 30 of ACT2, discontinuations due to adverse events occurred in 2/121 (1.7%) and 12/123 (9.8%) of infliximab 5mg/kg and placebo arm patients, respectively.[50] Through week 8 of UC-SUCCESS adverse events leading to discontinuation were highest for azathioprine (6/79, 8%), compared with 2/78 (3%) for infliximab and 3/80 (4%) for combination infliximab and azathioprine.[52]

Number of patients experiencing one or more adverse event

Adalimumab

In ULTRA1, patients reporting at least one treatment-emergent adverse event were 112/223 (50.2%) and 108/223 (48.4%) of the adalimumab 160/80mg induction and placebo groups, respectively.[45] At week 52 of ULTRA2, the proportions of patients reporting any adverse event were similar between groups; 213/257 (82.9%) of the adalimumab 160/80mg arm and 218/260 (83.8%) of the placebo arm. At week 52 in the Suzuki *et al.* study, fewer adverse events occurred (in terms of events per 100 patient years) in the adalimumab 40mg EOW group compared with the placebo group (547.9/100 PY vs. 609.4/100 PY).[47]

Golimumab

By week 6 of PURSUIT-SC, the proportions of patients with at least one adverse event were similar for golimumab 200/100mg induction (124/331, 37.5%) and placebo (126/330, 38.2%).[47] Patients reporting one or more adverse events through week 54 of PURSUIT-Maintenance were 112/154 (72.7%) in the golimumab 50mg, 113/154 (73.4%) in the golimumab 100mg and 103/156 (66.0%) in the placebo treatment arms.[49]

Infliximab

The proportions of patients through week 54 of ACT1 reporting at least one adverse event were 106/121 (87.6%) and 103/121 (85.1%) for infliximab 5mg/kg and placebo, respectively. At week 30 of ACT2, these values were 99/121 (81.8%) and 90/123 (73.2%) for infliximab 5mg/kg and placebo, respectively.[50] Through week 8 of UC-SUCCESS, patients reporting one or more adverse event were higher in the azathioprine group (41/79, 52%) than infliximab (26/78, 33%) or combination infliximab/azathioprine (30/80, 38%).[52]

Number of patients experiencing one or more serious adverse event

Definitions of serious adverse events were poorly reported across included RCTs.

Adalimumab

At week 8 in ULTRA1, the proportions of patients reporting one or more serious adverse events were exactly equivalent, at 5.4% (12/223) in the adalimumab 160/80mg group and 5.4% (12/223) in the placebo group.[45] Proportions of ULTRA2 patients reporting any serious adverse events were also roughly equivalent, with 12.1% (31/257) and 12.3% (32/260) in the adalimumab 160/80mg and placebo groups respectively.[46] At week 52 of the Suzuki *et al.* study, a similar number of events per 100 patient years were classed as serious in the adalimumab 40mg EOW group than in the placebo group (33.6/100 PY vs. 31.3/100 PY).[47]

Golimumab

By week 6 of PURSUIT-SC, the proportion of patients reporting at least one serious adverse event was lower in the golimumab 200/100mg treatment arm (9/331, 2.7%) compared to the placebo group (20/330, 6.1%).[48] More patients in the golimumab 100mg group reported one or more serious adverse event (22/154, 14.3%) than patients in the golimumab 50mg (13/154, 8.4%) or placebo (12/156, 7.7%) groups by week 54 of PURSUIT-Maintenance.[49]

Infliximab

Proportions of patients through week 54 of ACT1 who reported serious adverse events were similar for infliximab 5mg/kg (26/121 (21.5%) and placebo (31/121 (25.6%) groups.[50] At week 30 of ACT2, slightly fewer patients reported serious adverse events in the infliximab 5mg/kg group (13/121 (10.7%) than the placebo group (24/123 (19.5%).[50] Serious adverse events were more frequently reported by week 8 of UC-SUCCESS among patients receiving azathioprine (6/79, 8%) than infliximab (0/78) or combination infliximab and azathioprine (3/80, 4%).[52]

Infections

Adalimumab

The occurrence of infections at week 8 of ULTRA1 was very similar for the adalimumab 160/80mg group (32/223 (14.3%) and the placebo group (35/223, 15.7%).[45] This was also the case at week 52 of ULTRA2, with 45.1% (116/257) and 39.6% (103/260) patients reporting infections within the adalimumab 160/80mg and placebo groups, respectively.[46] At week 8 of Suzuki *et al.*, infections occurred in 18.9% (17/90) and 15.6% (15/96) of the adalimumab 160/80mg and placebo groups.[47]

Golimumab

At week 6 of PURSUIT-SC, 12.1% (40/330) of placebo group patients reported at least one infection, of which 7.0% required treatment (23/330); these values were similar to those in the golimumab 200/100mg induction group (39/331, 11.8%; 15/331, 4.5%).[48] Infections at week 54 of PURSUIT-Maintenance were more common in the golimumab 50mg (60/154, 39.0%; requiring treatment 39/154, 25.3%) and golimumab 100mg (60/154, 39.0%; requiring treatment 44/154, 28.6%) maintenance groups compared with placebo (44/156, 28.2%; requiring treatment 24/156, 15.4%).[48]

Infliximab

Through week 54 of ACT1, infections were slightly more common among patients receiving infliximab 5mg/kg (53/121, 43.8%; requiring treatment 39/121, 32.2%) compared with placebo (47/121, 38.8%; requiring treatment 25/121, 20.7%).[50] At week 30 of ACT2, infections had occurred in 18/121 (14.9%, requiring treatment 17/121, 14.2%) and 29/123 (23.6%; requiring treatment 15/123 (12.2%) of patients receiving infliximab and placebo respectively.[50] Through week 54 of the ACT1 and ACT2 extension studies, infections occurred in 94/242 (39%) of infliximab 5mg/kg and 80/244 (33%) of placebo group patients.[68]

Serious infections

Adalimumab

Reported serious infections were low through week 8 of ULTRA1[45] in both placebo (3/223 [1.3%], 1 pneumonia, 1 sepsis, 1 staphylococcal wound infection) and adalimumab 160/80mg treatment arms (0/223) and remained similarly comparable between treatment arms through week 52 of ULTRA2[46] (adalimumab 160/80mg 4/257, 1.6% vs. placebo 5/260, 1.9%). Serious infections were reported at week 52 of ULTRA3 at a rate of 3.4 events per 100 patient years for patients receiving adalimumab.[63] No serious infections were reported at week 8 of the Suzuki *et al.* trial in the placebo arm, whilst 3 cases occurred by week 8 in the adalimumab 160/80mg group (3/90, 3.3%).[47]

Golimumab

The proportion of patients reporting one or more serious infections were slightly higher at week 6 of PURSUIT-SC in the placebo treatment arm (6/330, 1.8%) compared with golimumab 200/100mg induction (1/331 (pneumonia), 0.3%).[48] By week 54 of PURSUIT-Maintenance, the occurrence of serious infections was marginally higher in the golimumab 50mg (5/154, 3.2%) and golimumab 100mg (5/154, 3.2%) maintenance groups than placebo (3/156, 1.9%).[48]

Infliximab

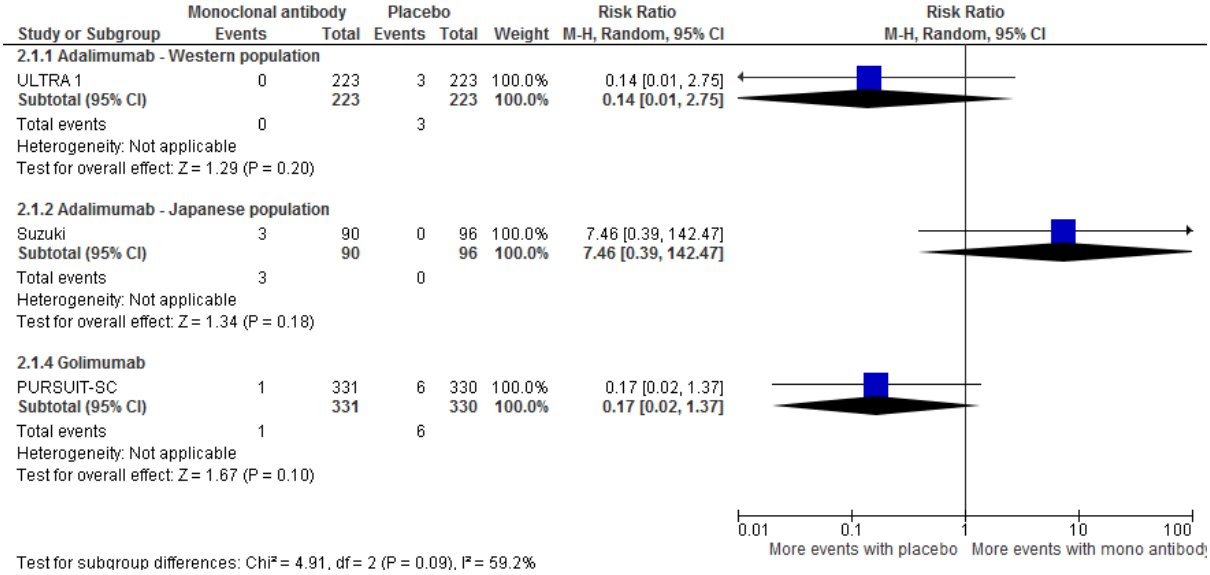
The proportion of patients with serious infections through week 54 of the ACT1 trial was similar between treatment arms (infliximab 5mg/kg 3/121, 2.5%; placebo 5/121, 4.1%).[50] Numbers of patients with serious infections through week 30 of ACT2 were similar for infliximab 5mg/kg (2/121, 1.7%) and placebo (1/123, 0.8%).[50] Through week 54 of the ACT1 and ACT2 extension studies, serious infections occurred in 7/242 (2.89%) of infliximab 5mg/kg and 6/244 (2.46%) of placebo group patients.[68]

Serious infections occurred in very low numbers through week 8 of the UC-SUCCESS trial (azathioprine 1/79, 1%; infliximab 1/78, 1%; combination infliximab/azathioprine 0/80 (0%).[52]

Meta-analysis

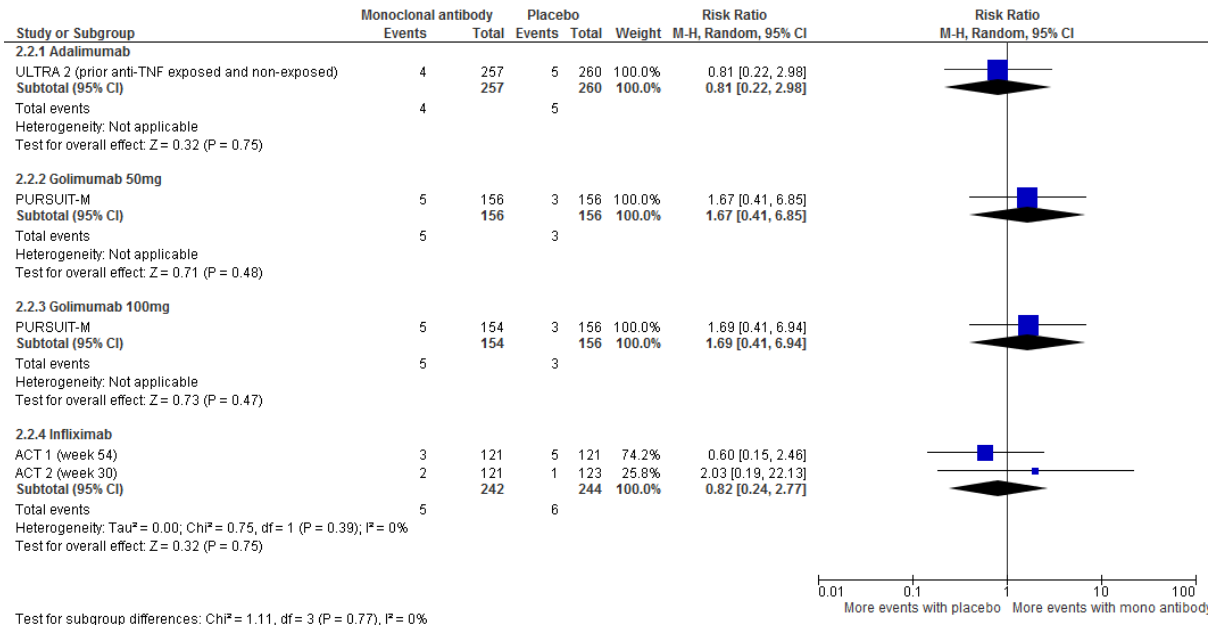
Serious infections associated with the licenced induction dose of adalimumab in were reported by two trials, one in Western populations (ULTRA1, 446 participants) and one in Japanese populations (Suzuki, 186 participants). The between-group difference in both trials was not significant (RR=0.14 [random effects] 95% confidence interval, 0.01 to 2.75; p=0.20; RR=7.46 [random effects] 95% confidence interval, 0.39 to 142.47; p=0.18 respectively). The forest plot for this analysis (random effects) is presented in Figure 8. Serious infections associated with the licenced induction of golimumab in adults were reported by one trial (PURSUIT-SC, 661 participants). The between-group difference was not significant (RR=0.17 [random effects], 95% confidence interval, 0.02 to 1.37; p=0.10).

Figure 8: Forest plot of comparison: Serious infections - adults, outcome: Serious infections - induction licenced dose



Serious infections associated with the licenced maintenance dose of adalimumab were reported by one trial comprising a mixed sample of anti-TNF-α exposed and naïve participants (ULTRA2, 517 participants). The between-group difference was not significant (RR=0.81 [random effects] 95% confidence interval, 0.22 to 2.98; *p*=0.75). Serious infections associated with maintenance dose of golimumab 50mg or 100mg in adults were reported by one trial (PURSUIT-Maintenance). The between-group difference for golimumab 50mg compared with placebo (312 participants) was not significant (RR=1.67 [random effects] 95% confidence interval, 0.41 to 6.85; *p*=0.48). The between-group difference for golimumab 100mg compared with placebo (310 participants) was also not significant (RR=1.69 [random effects] 95% confidence interval, 0.41 to 6.94; *p*=0.47). Two trials evaluating the licenced maintenance dose of infliximab reported maintenance outcomes at 30 weeks (ACT2) and 54 weeks (ACT1). The pooled effect across these trials (486 participants) was not significant (RR=0.82 [random effects] 95% confidence interval, 0.24 to 2.77; *p*=0.77). The forest plot for these analyses is presented in Figure 9.

Figure 9: Forest plot of comparison: Serious infections - adults, outcome: Serious infections - maintenance licenced dose.



Reactivation of TB

Adalimumab

No data relating to the reactivation of TB were reported for ULTRA1 or ULTRA2. Reactivation of TB occurred in a single patient (equating to <0.1 events/100 patient years) by week 52 of ULTRA3.[63] No events occurred in the placebo arm of the Suzuki *et al.*study through week 8, whilst for the adalimumab 40mg every other week group, a single event of reactivation of TB was described (1.0 events/100 patient years).[47]

Golimumab

No cases of reactivation of TB were reported in the PURSUIT-SC trial. In the placebo maintenance group of PURSUIT-Maintenance, one event of reactivation occurred (in a patient who had received unlicensed golimumab 4mg/kg i.v. induction).[48] No cases were reported for patients receiving golimumab 50mg maintenance treatment. However, three cases occurred in the golimumab 100mg maintenance group (1 patient each had received induction regimens of golimumab 400/200mg s.c., 4mg/kg i.v. or 200/100mg s.c.) (including one fatal case).[48]

Infliximab

No cases of reactivation of TB were reported in the ACT1, ACT2, Probert *et al.* or UC-SUCCESS studies.

Reactivation of hepatitis B

Adalimumab

No incidents of reactivation of hepatitis B were reported in any of the included adalimumab trials.

Golimumab

No cases were described in the included golimumab studies.

Infliximab

No events were reported in the included infliximab studies.

Administration reactions (injection site reactions / infusion reactions / serious allergic reactions)

Injection site reactions

Adalimumab

Injection-site reactions were slightly more frequent at week 8 of ULTRA1 among patients receiving adalimumab 160/80mg (13/223 (5.8%) compared with placebo (7/223 (3.1%).[45] Injection-site reactions were also more frequent in the adalimumab 160/80mg group at week 52 of ULTRA2 (31/257, 12.1%) than for placebo (10/260, 3.8%).[46] Patients receiving adalimumab through week 52 of ULTRA3 experienced injection-site reactions at a rate of 10.5 per 100 patient years.[63] Injection-site reactions were more frequent through week 8 of the Suzuki *et al.*trial in the adalimumab 160/80mg group (7/90, 7.8%) than for placebo (2/96, 2.1%).[47]

No serious allergic reactions were described as having occurred in the included adalimumab trials.

Golimumab

At week 6 of the PURSUIT-SC trial, injection-site reactions were more common in patients receiving 200/100mg golimumab induction (11/331, 3.3%) than placebo (5/330, 1.5%).[48] The number of patients reporting one or more injection-site reactions through week 54 of PURSUIT-Maintenance was higher in the golimumab 100mg maintenance treatment arm (11/154, 7.1%) compared with golimumab 50mg (3/154, 1.9%) and placebo (3/156, 1.9%).[48,49]

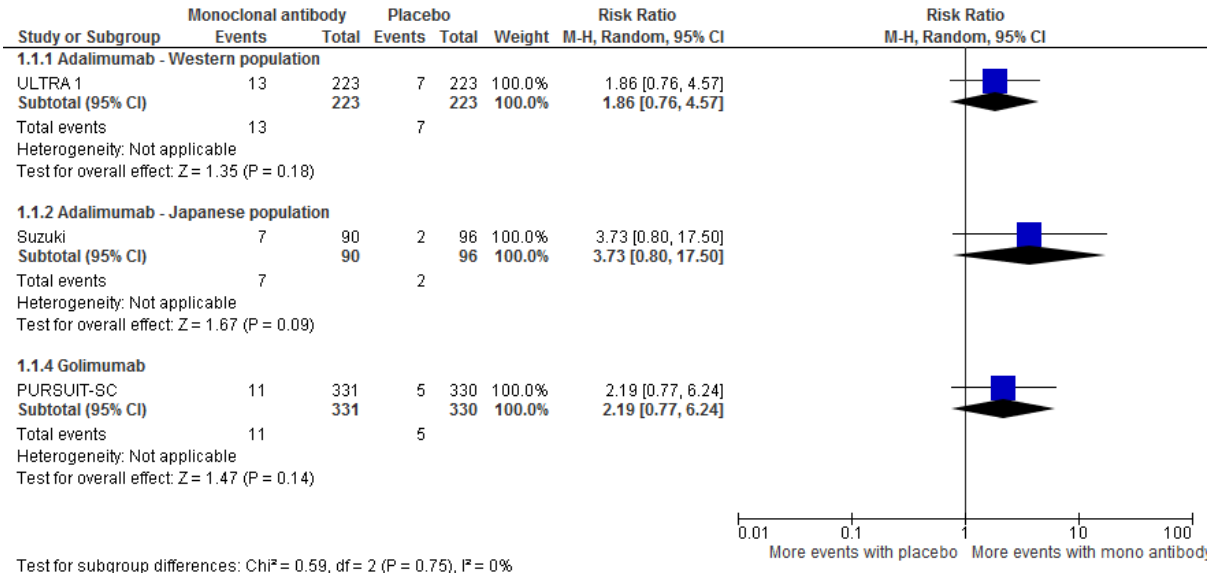
No serious allergic reactions were reported.

Meta-analysis

Injection site reactions associated with the licenced induction dose of adalimumab in were reported by two trials, one in Western populations (ULTRA1, 446 participants) and one in Japanese populations (Suzuki, 186 participants). The between-group difference in both trials was not significant (RR=1.86 [random effects] 95% confidence interval, 0.76 to 4.57; *p*=0.18; RR=3.73 [random effects] 95%

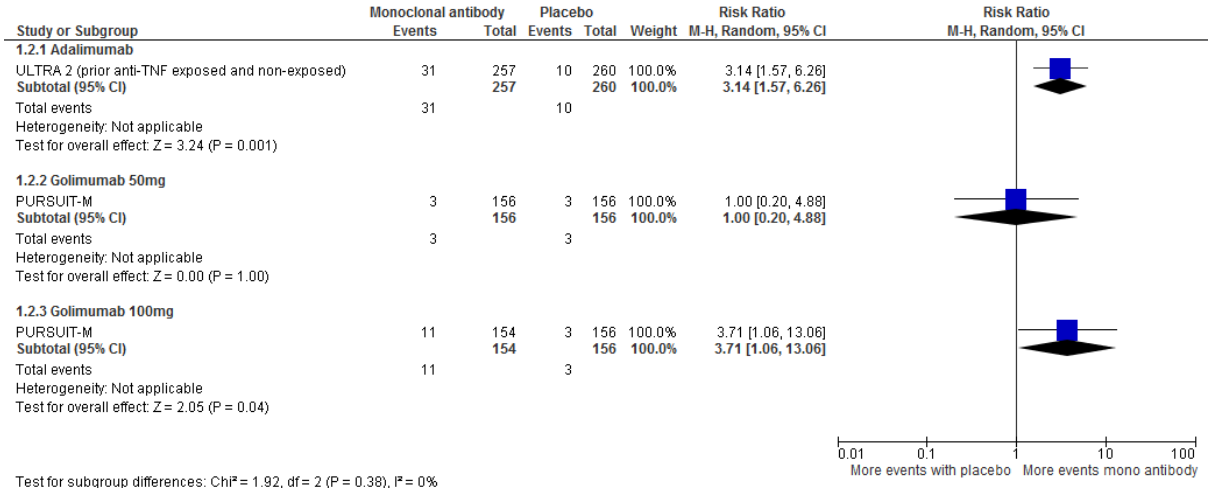
confidence interval, 0.80 to 17.50; *p*=0.09 respectively). The forest plot for this analysis (random effects) is presented in Figure 10. Injection site reactions associated with the licenced induction dose of golimumab in adults were reported by one trial (PURSUIT-SC, 661 participants). The between-group difference was not significant (RR=2.19 [random effects], 95% confidence interval, 0.77 to 6.24; *p*=0.14).

Figure 10: Forest plot of comparison: Injection site reactions - adults, outcome: Injection site reactions (ADA and GOL only) - induction licenced dose



Injection site reactions associated with maintenance doses of adalimumab were reported by one trial comprising a mixed sample of anti-TNF-α exposed and naïve participants (ULTRA2, 517 participants). The between-group difference was significant in favour of placebo (fewer events) (RR=3.14 [random effects] 95% confidence interval, 1.57 to 6.26; *p*=0.001). The forest plot for this analysis is presented in Figure 11. Injection site reactions associated with maintenance dose of golimumab 50mg or 100mg in adults were reported by one trial (PURSUIT-Maintenance). The between-group difference for golimumab 50mg compared with placebo (312 participants) was not significant (RR=1.00 [random effects] 95% confidence interval, 0.20 to 4.88; *p*=1.00). The between-group difference for golimumab 100mg compared with placebo (310 participants) was significant in favour of placebo (fewer events) (RR=3.71 [random effects] 95% confidence interval, 1.06 to 13.06; *p*=0.04).

Figure 11: Forest plot of comparison: Injection site reactions - adults, outcome: Injection site reactions (ADA and GOL only) - maintenance licenced dose



Infusion reactions

Infliximab

Acute infusion reactions occurred in similar numbers of patients in both treatment arms through week 54 of ACT1 (infliximab 5mg/kg 2/121, 9.9%; placebo 13/121, 10.7%).[50] Infusion reactions were slightly higher in ACT2 patients receiving infliximab 5mg/kg (14/121, 11.6%) compared with placebo (10/123, 8.1%).[50]

Infusion reactions were rare through week 8 of UC-SUCCESS (azathioprine 1/79, 1%; infliximab 0/78, 0%; combination infliximab/azathioprine 0/80 (0%).[52] Possible delayed hypersensitivity reactions occurred in 2/242 (1%) of the infliximab 5mg/kg group and 2/242 (1%) of the placebo group through week 54 of the ACT1 and ACT2 extension studies.[68]

No serious allergic reactions were reported.

Heart failure

Adalimumab

Heart failure did not occur in any patients in either adalimumab 160/80mg induction or placebo arms by week 8 of ULTRA1.[45] Only one case of heart failure was reported through week 52 of ULTRA2, which was in a patient receiving adalimumab 160/80mg for induction (1/257, 0.4%).[46] Heart failure was reported at a rate of of 0.2 events per 100 patient years for adalimumab 40mg EOW/EW at week 52 of ULTRA3.[63] Through week 8 of the Suzuki *et al*. trial, no cases of heart failure were reported.[47]

Golimumab

No cases of heart failure were reported for either the golimumab 200/100mg induction or placebo treatment arms through week 6 of PURSUIT-SC or for the golimumab maintenance or placebo groups in PURSUIT-Maintenance.[48,49]

Infliximab

No cases of heart failure were reported in the ACT1, ACT2, ACT1 and ACT extension studies, Probert *et al.* and UC-SUCCESS trials.

Malignancies and lymphoproliferative disorders

Adalimumab

Malignancies were reported at low levels through week 8 of ULTRA1, with 2/223 events (0.9%, 1 basal cell carcinoma, 1 breast cancer) in the placebo group and no cases in the adalimumab 160/80mg group.[45] Two cases of malignancy were reported through week 52 of ULTRA2, both of which were in patients receiving adalimumab 160/80mg.[46] Through week 52 of ULTRA3, events (excluding lymphoma) occurred in the adalimumab 40mg maintenance arm at a rate of 1.0 events/per 100 patient years and at a rate of 0.1 events per 100 patient years for lymphoma.[63] One case of malignancy (1/90, 1.1%) was described in the adalimumab 160/80mg group at week 8 of the Suzuki *et al.* trial.[47]

Golimumab

No cases of malignancy were reported for either the golimumab 200/100mg induction or placebo treatment arms through week 6 of PURSUIT-SC.[48] Whilst one malignancy (1/156, 0.6%) was described by week 54 of PURSUIT-Maintenance in the placebo arm, four cases each were observed in the golimumab 50mg (4/154, 2.6%) and 100mg (4/154, 2.6%) maintenance groups.[48,49]

Infliximab

Two cases of malignancy were reported through week 54 of ACT1 in patients receiving infliximab 5mg/kg.[68] One case of basal cell carcinoma was reported in the placebo arm and one case of rectal adenocarcinoma was described in the infliximab 5mg/kg arm of ACT2 through week 30.[50] No malignancies were described in the UC-SUCCESS trial.[52]

Hepatobilary events / liver enzyme changes

Adalimumab

No cases were described in ULTRA1 or ULTRA2.[45,46] Hepatobiliary events were reported a rate of 0.5 events per 100 patient years in the adalimumab 40mg maintenance arm through week 52 of ULTRA2.[63] By week 8 of the Suzuki *et al.* trial, events occurred in 1/90 (1.1%) of adalimumab 160/80mg and 1/96 (1.0%) of placebo group patients.[47]

Golimumab

No cases were reported for either the golimumab 200/100mg induction or placebo treatment arms through week 6 of PURSUIT-SC or the golimumab maintenance or placebo groups of PURSUIT-Maintenance.[48,49]

Infliximab

No cases of hepatobiliary events were reported in the ACT1 and ACT2 trials.[50] The occurrence of hepatobiliary events was higher in the azathioprine treatment arm (13/79, 16%) compared with the infliximab (3/78, 4%) and combination infliximab/azathioprine (5/80 (6%) treatment groups through week 8 of UC-SUCCESS.[52]

Autoimmune processes (e.g. lupus-like syndrome)

Adalimumab

It was stated that no events of lupus-like syndrome occurred in the adalimumab 160/80mg or placebo treatment arms by week 8 of ULTRA1.[45] One case of lupus-like syndrome (1/257, 0.4%) was reported in a patient receiving adalimumab 160/80mg through week 52 of ULTRA2. [46] No cases were reported through week 8 of the Suzuki *et al*. trial.[47]

Golimumab

No cases of autoimmune processes were reported for either the golimumab 200/100mg induction or placebo treatment arms through week 6 of PURSUIT-SC or the golimumab maintenance or placebo groups of PURSUIT-Maintenance.[48,49]

Infliximab

One patient receiving infliximab 5mg/kg reported experiencing a lupus-like reaction by week 30 of ACT2.[50] No cases of auto-immune processes were described in the UC-SUCCESS trial.[52]

Neurological events

Adalimumab

No cases of demyelinating disease occurred in the adalimumab 160/80mg or placebo treatment arms by week 8 of ULTRA1[45] or by week 52 of ULTRA2.[46] No cases of neurological events were reported through week 8 of the Suzuki *et al*. trial.[47]

Golimumab

No cases were reported for either the golimumab 200/100mg induction or placebo treatment arms through week 6 of PURSUIT-SC or the golimumab maintenance or placebo groups of PURSUIT-Maintenance.[48]

Infliximab

One patient receiving infliximab 5mg/kg reported having optic neuritis through week 54 of ACT1. One patient receiving infliximab 5mg/kg also experienced optic neuritis by week 30 of ACT2.[50,68] No neurological events were described in the UC-SUCCESS trial.[52]

Haematological reactions

Adalimumab

No haematological reactions were described in ULTRA1.[45] One haematological reaction was reported in 5/257 (1.9%) of patients receiving adalimumab 160/80mg by week 52 of ULTRA2.[46] Haematological reactions occurred in 1/90 (1.1%) and 1/96 (1.0%) of patients receiving adalimumab 160/80mg and placebo respectively by week 8 of the Suzuki *et al.* study.[47]

Golimumab

No haematological reactions were reported for either the golimumab 200/100mg induction or placebo treatment arms through week 6 of PURSUIT-SC or the golimumab maintenance or placebo groups of PURSUIT-Maintenance.[48,49]

Infliximab

No haematological reactions were described in ACT1, ACT2, Probert *et al.* or UC-SUCCESS.

5.2.3.2 Population: Children and adolescents aged 6 to 17 years (inclusive) with severely active UC, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies

a) Rates of and duration of response, relapse and remission

Table 12 presents the definitions of clinical response and remission in the included paediatric population RCT.

Table 12: Definitions of clinical response and remission in included paediatric population RCT

Trial	Definition of clinical response	Definition of clinical remission	Measurement time points
Hyams *et al.*, 2012	Decrease in Mayo score of ≥ 3 points and ≥ 30%, with accompanying decrease in subscore for rectal bleeding of ≥ 1 point or absolute rectal bleeding subscore of 0 or 1.	Mayo score of ≤ 2 points, with no individual subscore > 1. PUCAI clinical remission = score < 10.	Mayo scores assessed at weeks 0, 8 and 54. Endoscopy at week 54 optional.

All enrolled patients received induction therapy with infliximab 5mg/kg. At week 8, clinical response was reached by 44/60 patients (73.3%), whilst 24/60 (40.0%) of patients achieved clinical remission.

PUCAI remission rates were evaluated at weeks 30 and 54. A greater proportion of patients in the infliximab 5mg/kg every 8 weeks treatment group achieved PUCAI remission at week 30 (40.0% vs. 19.0%, *p*-values NR) and week 54 (38.1% vs. 18.2%, *p*-values NR) compared with the infliximab 5mg/kg every 12 weeks group. At week 54, PUCAI remission without the use of corticosteroids was reported for 38.5% of the every 8 weeks group and 0% of the every 12 weeks group.[53]

The absence of a placebo/non-infliximab control group limits the comparative evaluation of the efficacy of infliximab in induction and maintenance of clinical response and remission in paediatric patients. A briefing document[69] by Centocor Ltd. to the US Food and Drug Administration (FDA) Gastrointestinal Drugs Committee was produced in June 2011 and considered the evidence available from the Hyams e*t al.*, 2012/T72 study[53] and compared this with the ACT1 and ACT2 trials[50] of infliximab previously conducted in the adult UC population. The briefing document considered efficacy to be similar between T72 and the ACT1 and ACT2 studies during induction (with clinical response and Mayo remission at week 8 induced in 73.3% and 40.0% of paediatric patients and 66.9% and 36.4% of pooled 5mg/kg adult patients from ACT1 and ACT2) and maintenance (with PUCAI remission at week 54 in 38.1% of paediatric subjects in the every 8 weeks group and 34.7% at week 54 of ACT1 (with reported good correlation of 0.75-0.88 between PUCAI and Mayo scores described at baseline and week 8).

b) Measures of disease activity

At week 8 of the Hyams *et al.* study, the median reductions in partial Mayo scores were 4 points for both the infliximab 5mg/kg every 8 weeks group and infliximab 5mg/kg every 12 weeks group.[53] By week 30, the median reduction in partial Mayo score was approximately 3 points for the every 8 weeks group and 1 point for the every 12 weeks group.[53,53]

c) Mortality

No deaths were reported in the Hyams *et al.* trial.

d) Rates of hospitalisation

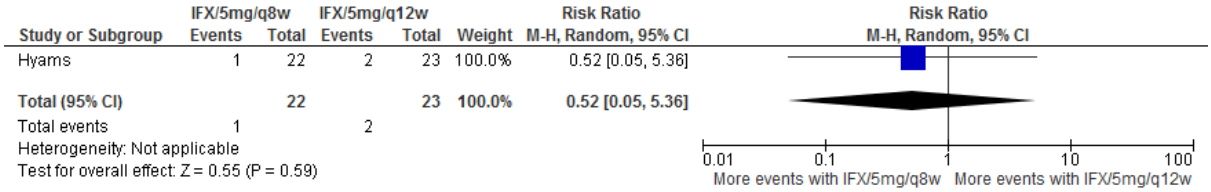
No hospitalisation-related outcome data were reported in Hyams *et al*.

e) Rates of surgical intervention (both elective and emergency)

One of 22 patients (4.5%) in the infliximab 5mg/kg every 8 weeks group required colectomy through week 54 in the Hyams *et al.* trial as compared with two of 23 (8.7%) patients in the infliximab 5mg/kg every 12 weeks treatment arm.

Colectomy rates during maintenance in children were reported by one trial evaluating the licenced dose of infliximab every 8 weeks (q8w) or every 12 weeks (q12w) (Hyams, 45 participants). The between-group at week 54 was not significant (RR=0.52 [random effects], 95% confidence interval, 0.05 to 5.36; *p*=0.59, see Figure 12).

Figure 12: Forest plot of comparison: Colectomy - children, outcome: Infliximab maintenance



f) Time to surgical intervention (both elective and emergency)

No data were reported in the paediatric population for the outcome of time to surgical intervention.

g) Health-related quality of life

No HRQoL data were included in the Hyams *et al*. trial.

h) Adverse events of treatment (including leakage and infections following surgery)

Discontinuations due to adverse events

Through week 54 of the Hyams *et al*. trial, discontinuations due to at least one adverse event were higher in the infliximab 5mg/kg every 12 weeks group than the every 8 weeks frequency group (6/23 [26.1%] vs. 3/22 [13.6%]).[53]

Number of patients experiencing one or more adverse event

All patients in both treatment arms of the Hyams *et al*. study reported at least one adverse event (22/22 [100%] vs. 23/23 [100%]).[53]

Number of patients experiencing one or more serious adverse event

The numbers of patients reporting at least one serious adverse event were similar between the infliximab 5mg/kg every 12 weeks (5/23 [21.7%]) and every 8 weeks (4/22 [18.2%]) treatment arms.[53]

Infections

The occurrence of infections was comparable between infliximab 5mg/kg every 8 weeks (13/22, 59.1%) and every 12 weeks (14/23, 60.9%) treatment groups.[53]

Serious infections

No cases of serious infection were reported in the Hyams *et al.* trial. [53]

Reactivation of TB

No cases were reported.

Reactivation of hepatitis B

No cases were reported.

Administration reactions (injection site reactions / infusion reactions / serious allergic reactions)

The numbers of patients experiencing infusion reactions were similar between treatment groups in the Hyams *et al.* study (4/22, 18.2% vs. 3/23, 13.0%).[53]

5.2.3.3 Subgroups

As stated in the assessment protocol (Appendix 1), the only pre-specified subgroup of interest was duration of disease. However, clinical data reported according to disease duration were very limited. The only studies to evaluate the effect of disease duration on outcomes were ULTRA2[46] and PURSUIT-Maintenance.[48,49]

For ULTRA2, the odds ratios for the proportion of patients in clinical remission at week 8 for adalimumab versus placebo were very similar for patients with disease duration of ≤2 years (OR=1.91, 95%CI=0.4, 8.8, p=0.40) and those with disease duration of > 2 years (OR=1.92, 95%CI 1.1, 3.4, p=0.03). However at week 52, the odds ratio for clinical remission was considerably higher for patients with disease duration > 2 years (OR=3.59, 95% CI=1.9, 6.9, p<0.001) than for patients with a shorter disease duration of ≤ 2 years (OR=0.22, 95%CI=0.04, 1.1, p=0.05).

PURSUIT-Maintenance reported the odds ratios (OR) for comparing the proportion of patients in clinical response in the golimumab maintenance group versus the placebo group for golimumab-induction responders. The odds ratio for proportion of patients in clinical response through week 54 for golimumab 50mg versus placebo treatment arms was slightly higher among patients with longer disease duration (>5 to ≤15 years; OR=2.3, 95% CI 1.0, 5.4, p=0.056) than those with shorter duration of disease (≤5 years; OR=1.4, 95% CI 0.9, 2.7, p=0.533). Similarly, the odds ratio for golimumab 100mg versus placebo groups was also reported to be greater among those with a disease duration of >5 to ≤15 years (OR=2.2, 95% CI 1.0, 4.9, p=0.068) than for patients with disease duration of 5 years or less (OR=1.6, 95% CI 0.8, 3.1, p=0.128). However, it was noted that the 95% confidence intervals for these observations overlapped between estimates.

5.2.3.4 Methods for network meta-analysis

The trials identified in the systematic review formed a connected network such that each trial had at least one treatment in common with at least one other trial. Treatment effects were estimated using NMAs of clinical response and remission as defined by the complete Mayo score.

Selection of evidence contributing to the network meta-analysis

For RCTs to be eligible for inclusion in the NMA they were required to have both clinical response and clinical remission data reported for either an induction (6-8 weeks) or maintenance (approximately 30 weeks or 52-54 weeks) time point. It should be noted that two adult population RCTs evaluating the use of infliximab as an induction treatment (Probert *et al.*, 2003; UC-SUCCESS) were excluded from the adult population NMA; these studies were excluded for other reasons, as described in the table of trial characteristics (see Table 6). The base case analyses utilised data from the anti-TNF-α-naïve population rather than the ITT population in ULTRA2 in order to increase comparability of the dataset. The induction base case also incorporated both Phase II (plus additional analysed patients from Phase II) and Phase III data from PURSUIT-SC. The effect of using the ITT (mixed anti-TNF-α-experienced) population from ULTRA2 was explored in a sensitivity analysis. Since the Suzuki *et al.* trial was conducted in exclusively Japanese patients, this trial was not included in the base case; however, the addition of this trial to the network was explored in a sensitivity analysis. Therefore, three sensitivity analyses were performed for both induction and maintenance phases to assess the robustness of replacing ULTRA2 anti-TNF-α-naïve data with ULTRA2 ITT data (sensitivity analysis 1), including Suzuki *et al.* (sensitivity analysis 2), and replacing ULTRA2 anti-TNF-α-naïve data with ULTRA2 ITT data plus including Suzuki *et al.* (sensitivity analysis 3).

Clinical response and remission data were defined as outlined in Table 10 and were taken from two different sources. Firstly, data relating to clinical response and remission for the use of interventions as induction treatment were extracted directly from the published RCT reports. Secondly, data relating to clinical response and remission for the use of interventions as maintenance treatments conditional on outcomes at previous timepoints were requested and received from the manufacturers of the products under assessment (MSD and Abbvie).

Statistical model for the network meta-analysis

Clinical response/remission can be considered as ordered categorical data with three mutually exclusive categories: (i) no response (ii) response and (iii) remission. The model for the data assumed that the treatment effect was the same irrespective of the category. Data available at 6 weeks and 8 weeks were combined, as were data available at 30 weeks and 32 weeks, and 52 weeks and 54 weeks. The likelihood function for the data is described as follows. Let r_{ikj}represent the number of patients

in arm k of trial i in the mutually exclusive category $j = 1,2, \ldots J$. The responses r_{ikj} will follow a multinomial distribution such that:

$$r_{ikj=1,\ldots,J} \sim \text{Multinomial}(p_{ikj=1,\ldots,J}, n_{ik}), \sum_{j=1}^{J} p_{ikj=1,\ldots,J} = 1$$

The parameters in the model are the probabilities, p_{ikj}, that a patient in arm k of trial i has a response equivalent to category j.

We used a probit link function to map the probabilities, p_{ikj}, onto the real line such that:

$$\theta_{ikj} = \Phi^{-1}(p_{ikj}) = \mu_{ij} + \delta_{i,bk} I_{k \neq 1}$$

so that:

$$p_{ikj} = \Phi(\mu_{ij} + \delta_{i,bk} I_{k \neq 1}).$$

In this model, the effect of treatment was to change the probit score of the control arm by $\delta_{i,bk}$ standard deviations.

The study-specific treatment effects, $\delta_{i,bk} I_{k \neq 1}$, were assumed to arise from a common population distribution with mean treatment effect relative to the reference treatment, which in this analysis was placebo, such that:

$$\delta_{i,1k} \sim N(d_{t_{i1},t_{ik}}, \tau^2)$$

We further assumed that there is an underlying continuous latent variable which has been categorised by specifying cut-offs, z_{ij}, which corresponds to the point at which an individual moves from one category to the next in trial i. The model is re-written as:

$$p_{ikj} = \Phi(\mu_i + z_{ij} + \delta_{i,bk} I_{k \neq 1}).$$

The z_{ij} can be treated as fixed, which would assume that these points are the same in each trial and each treatment. Alternatively, they can be treated as random in which they are assumed to vary according to the trial but that within a trial they are the same such that:

$$z_{ic} \sim N(\nu_c, \sigma_z^2).$$

We used a model in which the z_{ij} were treated as being random because this resulted in a much better fit of the model to the data. Further details of the model are presented in Dias *et al.*[70]

The model was completed by giving the parameters prior distributions. When there are sufficient sample data, we can use conventional reference prior distributions and these will have little influence on the posterior results.

The reference prior distributions used in the analyses were:

- Trial-specific baselines, $\mu_i \sim N(0, 1000)$
- Treatment effects relative to reference treatment, $d_{1t} \sim N(0, 1000)$
- Between study standard deviation of treatment effects, $\tau \sim U(0,2)$
- Population cut-offs, $\upsilon_{c_j} = \upsilon_{c_{j-1}} + \upsilon_c^{'}, \quad \upsilon_c^{'} \sim U(0,5)$
- Between study standard deviation of cut-offs, $\sigma_z^2 \sim U(0,2)$

In both the induction and maintenance phases, there were relatively few studies to allow Bayesian updating of the implausibly vague prior distribution for the between-study standard deviation. Without Bayesian updating, a reference prior distribution that does not represent genuine prior belief will have a significant impact on the results and give posterior distributions that are unlikely to represent genuine posterior beliefs. To allow for this, we used a weakly informative prior distribution (a half normal distribution) for the between study standard deviation such that $\tau \sim HN(0, 0.32^2)$.

To estimate the absolute probabilities of being in each category for each treatment, we combined the treatment effects with an estimate of the placebo "No response" category (baseline model). We used a Binomial likelihood function for the number of patients, r_{ik1} in each study who were classified as having "No response" when treated with placebo for the baseline model such that:

$$r_{ik1} \sim \text{Binomial}(n_{ik}, p_{ik1}).$$

We used a probit link function such that:

$$\Phi^{-1}(p_{ik1}) = \mu_i^{'}.$$

We assumed that the study-specific baselines arose from population of effects such that:

$$\mu_i^{'} \sim N(\mu_b, \tau_b^2).$$

The model was completed by giving the parameters prior distributions such that:

- $\mu_b \sim N(0, 1000)$
- $\tau_b \sim U(0, 2)$

Again, in both induction and maintenance phase there were relatively few studies providing data so a weakly informative prior distribution was used for the between-study standard deviation such that:

$$\tau \sim HN(0, 0.32^2).$$

All analyses were conducted in the freely available software package OpenBUGS.[70] For the baseline and relative treatment effects models, we used a burn-in of 50,000 iterations of the Markov chain and

retained a further 10,000 iterations to estimate parameters. In addition, the network meta-analyses exhibited moderate correlation between successive iterations of the Markov chains so the chains were thinned by retaining every 10th sample.

5.2.3.5 Results of network meta-analyses

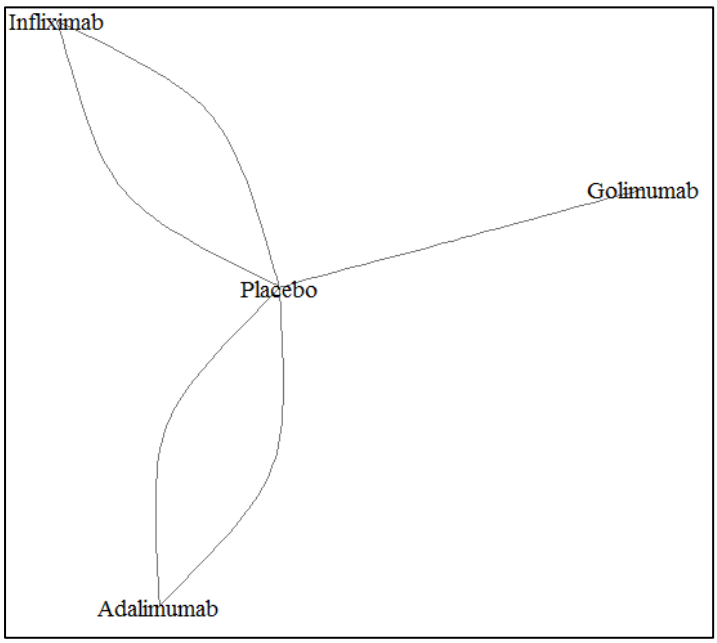
a. Results of network meta-analyses

A summary of the data used in the NMA is provided in Appendix 8. As described earlier, three sensitivity analyses were undertaken to assess the robustness of replacing ULTRA2 anti-TNF-naïve data with ULTRA2 ITT data (sensitivity analysis 1), including Suzuki *et al* (sensitivity analysis 2), and replacing ULTRA2 anti-TNF-α naïve data with ULTRA2 ITT data and including Suzuki (sensitivity analysis 3). The results presented in Sections a.1 to a.12 were derived using weakly informative prior distributions (a half normal distribution) for the between-study standard deviation such that $\tau \sim HN(0, 0.32^2)$. Results using vague reference prior distributions (τ~U(0,2)) are presented in Appendix 9.

a.1 Base case – induction phase

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the induction phase. Data were available from 5 studies comparing two treatments. Figure 13 presents the network of evidence for the base case induction phase.

Figure 13: Base case – network of evidence for the induction phase



Note: solid line indicates a 2-arm trial

Figure 14 presents the effects of each treatment relative to placebo on the probit scale for the base case induction phase. Figure 15 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 18.16, being close to the total number of data points included in the analysis, 20. The between-study standard deviation was estimated to be 0.12 (95% CrI: 0.01, 0.50), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab. All treatment effects were statistically significant at a conventional 5% level. Infliximab was associated with the greatest effect -0.92 (95% CrI: -1.27, -0.56) and was most likely to be the most effective treatment (probability of being the best = 0.93).

Figure 14: Base case - comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))

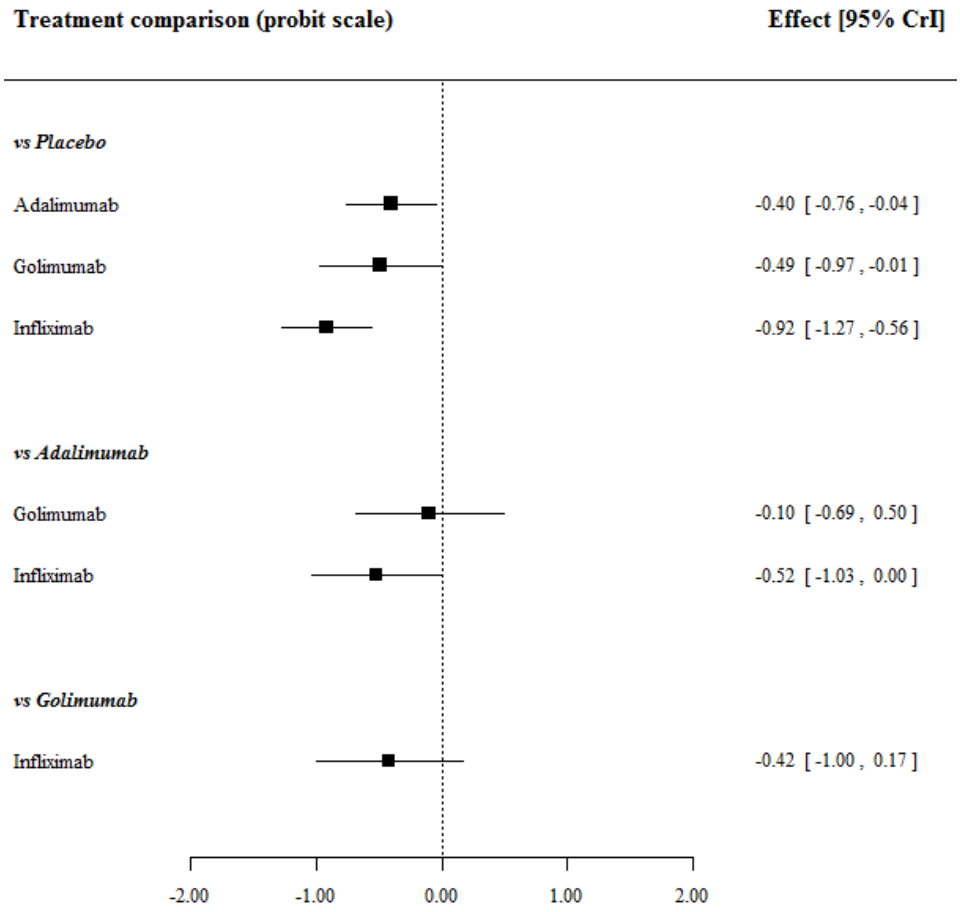
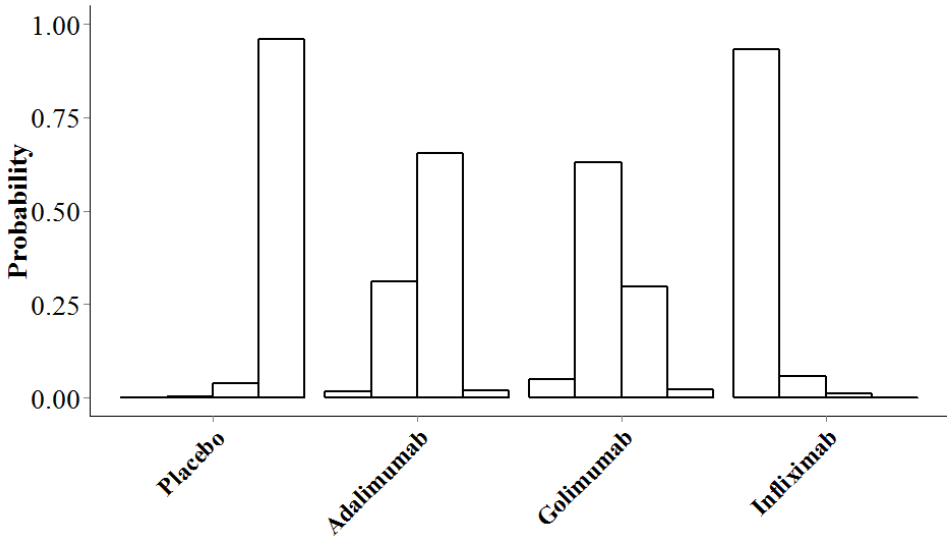


Figure 15: Base case – ranking probability histograms for the induction phase



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 13 presents the probabilities of achieving each of the following categories: no response, response and remission for the base case induction phase. Infiliximab was associated with the highest probability of moving from no response to response and no response to remission, respectively. The effects of adalimumab and golimumab on each transition probability were comparable.

Table 13: Base case – probabilities of being in each category for the induction phase

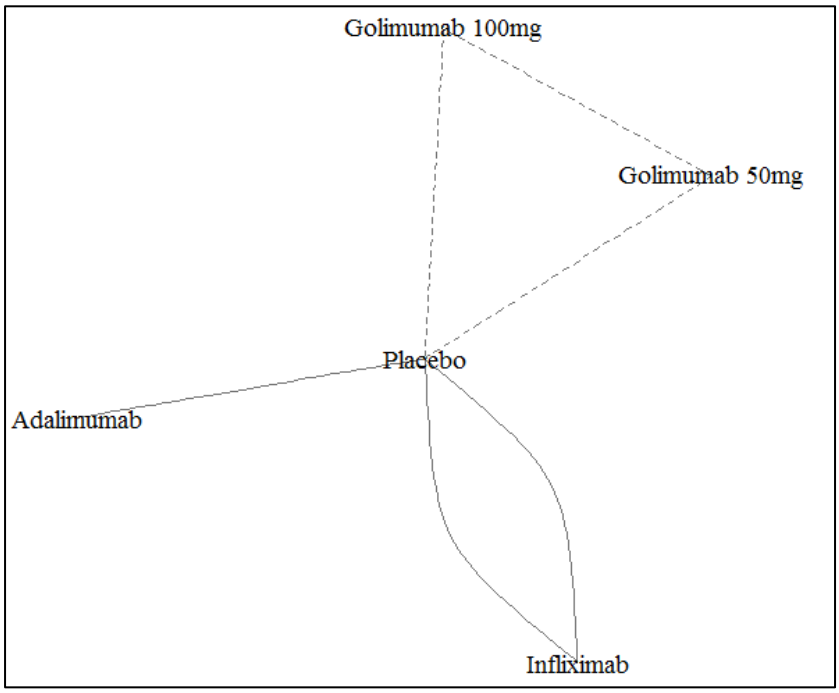
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.640	0.641	0.568, 0.706	0.260	0.260	0.214, 0.308	0.099	0.097	0.062, 0.147
Adalimumab	0.485	0.485	0.330, 0.642	0.324	0.327	0.247, 0.385	0.190	0.185	0.092, 0.322
Golimumab	0.448	0.447	0.262, 0.645	0.333	0.337	0.244, 0.393	0.219	0.212	0.094, 0.390
Infliximab	0.292	0.289	0.170, 0.438	0.351	0.353	0.280, 0.412	0.356	0.352	0.209, 0.523

a.2 Base case – maintenance phase 8-32 weeks

a.2.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 8-32 weeks for patients starting in response. Data were available from 4 studies comparing two or three treatments. Figure 16 presents the network of evidence for the base case maintenance phase at 8-32 weeks for patients starting in response.

Figure 16: Base case – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 17 presents the effects of each treatment relative to placebo on the probit scale for the base case maintenance phase at 8-32 weeks for patients starting in response. Figure 18 presents the probabilities of treatment rankings for this analysis. There was some suggestion that the model did not represent the data well with the total residual deviance, 11.73, being smaller than would be expected given the total number of data points included in the analysis, 18. The probability of observing a value less than 11.73 was 0.139, which means that it could be a chance event. All 4 studies had smaller residual deviances than expected (ULTRA2: deviance 3.0 compared with 4 data points; ACT1: deviance 2.1 compared with 4 data points; ACT2: deviance 2.66 compared with 4 data points; and PURSUIT: deviance 4.0 compared with 6 data points). The between-study standard deviation was estimated to be 0.17 (95% CrI: 0.01, 0.61), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 100mg. However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 100mg was associated with the greatest effect -0.42 (95% CrI: -0.78, 0.29) and was most likely to be the most effective treatment (probability of being the best = 0.47).

Figure 17: Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~HN(0,0.32^2))

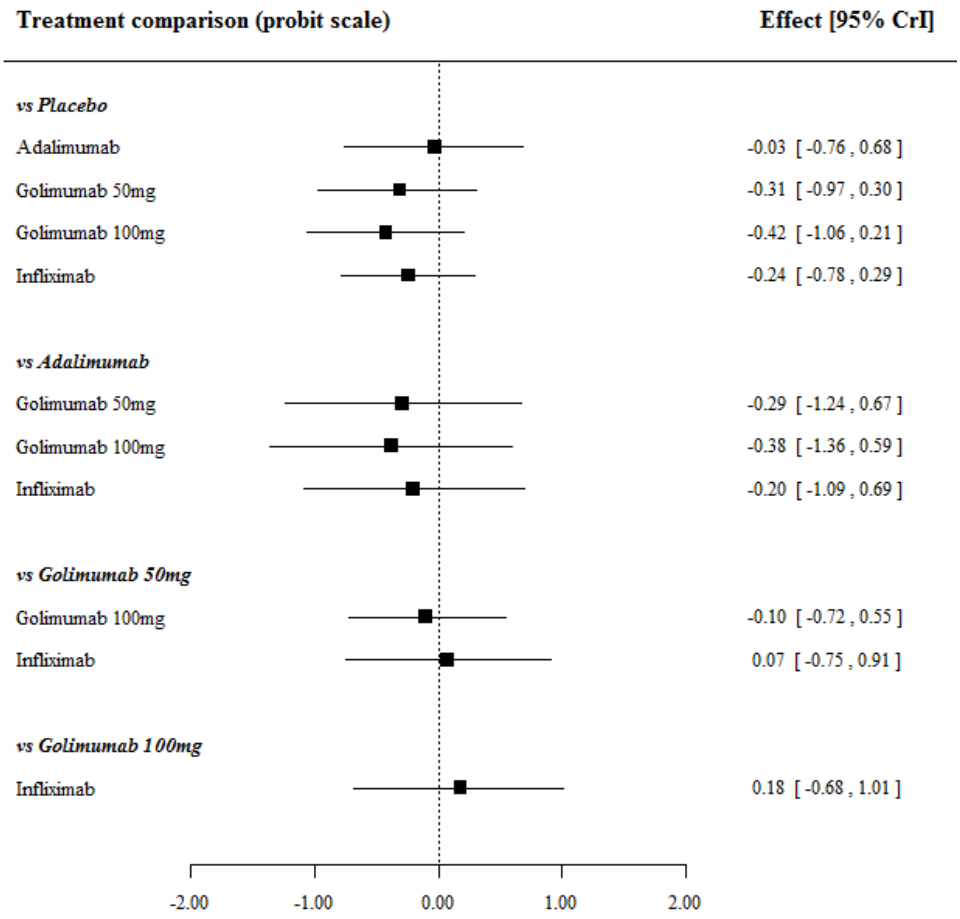
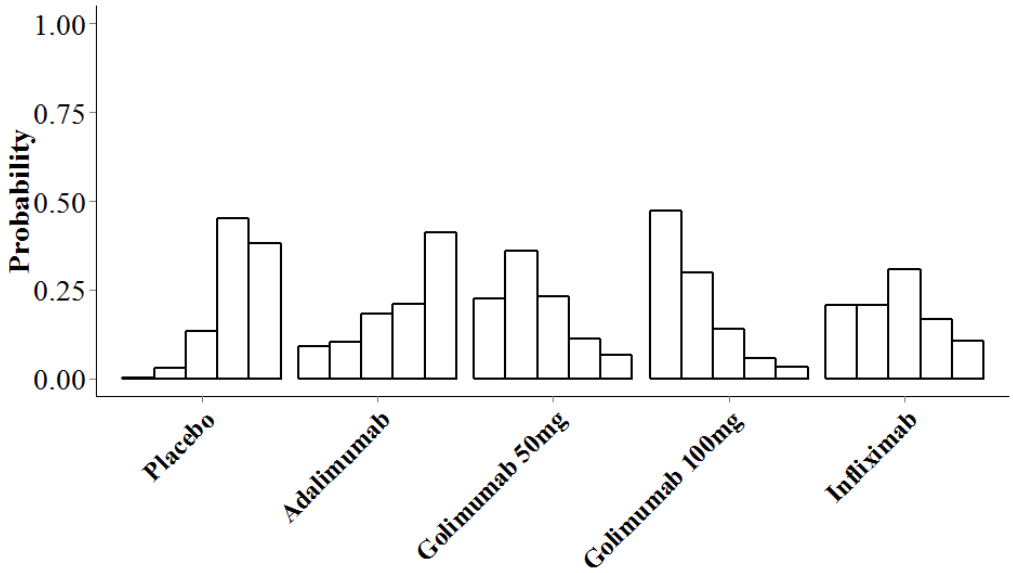


Figure 18: Base case – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 14 presents the probabilities of achieving each of the following categories: no response, response and remission for the base case maintenance phase at 8-32 weeks for patients starting in response. Golimumab 100mg was associated with the highest probability of moving from response to remission and staying in the response state at 8-32 weeks. It was also associated with the smallest probability of moving from response to no response. The probabilities of staying in response were comparable among all treatments at 8-32 weeks.

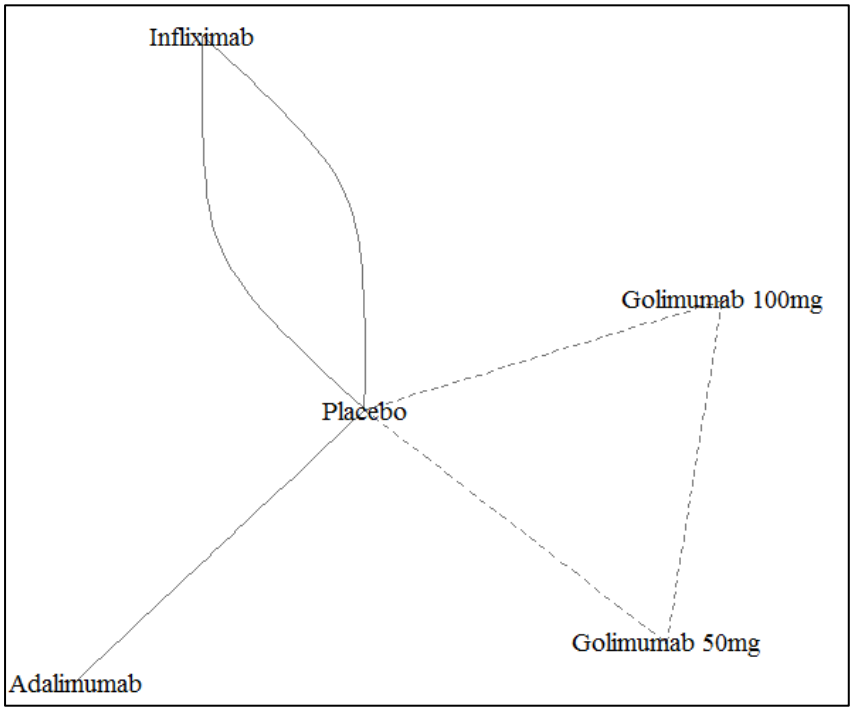
Table 14: Base case – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.524	0.525	0.426, 0.622	0.270	0.270	0.198, 0.341	0.206	0.202	0.117, 0.311
Adalimumab	0.512	0.512	0.230, 0.782	0.261	0.267	0.140, 0.354	0.227	0.211	0.055, 0.493
Golimumab 50mg	0.403	0.399	0.173, 0.660	0.283	0.285	0.176, 0.374	0.313	0.303	0.108, 0.588
Golimumab 100mg	0.368	0.360	0.149, 0.619	0.285	0.288	0.176, 0.377	0.347	0.338	0.129, 0.623
Infliximab	0.432	0.430	0.220, 0.659	0.282	0.283	0.189, 0.371	0.286	0.276	0.109, 0.518

a.2.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 8-32 weeks for patients starting in remission. Data were available from 4 studies comparing two or three treatments. Figure 19 presents the network of evidence for the base case maintenance phase at 8-32 weeks for patients starting in remission.

Figure 19: Base case – network of evidence for the maintenance phase at 8-32 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 20 presents the effects of each treatment relative to placebo on the probit scale for the base case maintenance phase at 8-32 weeks for patients starting in remission. Figure 21 presents the probabilities of treatment rankings for this analysis. The model fitted the data well, with the total residual deviance, 18.20, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.18 (95% CrI: 0.01, 0.64), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except adalimumab were associated with beneficial treatment effects relative to placebo with the greatest effects being associated with golimumab 50mg (-0.63; 95% CrI: -1.36, 0.11) and golimumab 100mg (-0.61; 95% CrI: -1.32, 0.11). However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 50mg and golimumab 100mg were most likely to be the most effective treatments (probability of being the best = 0.47 and 0.42, respectively).

Figure 20: Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))

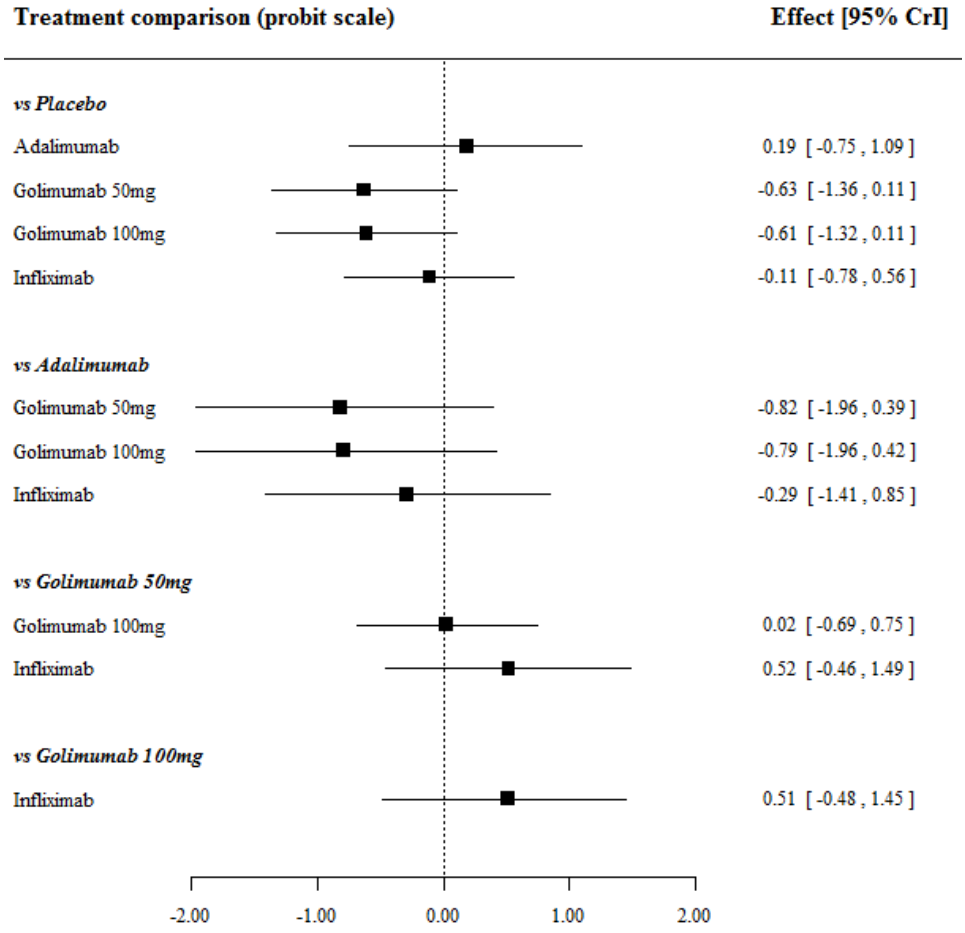
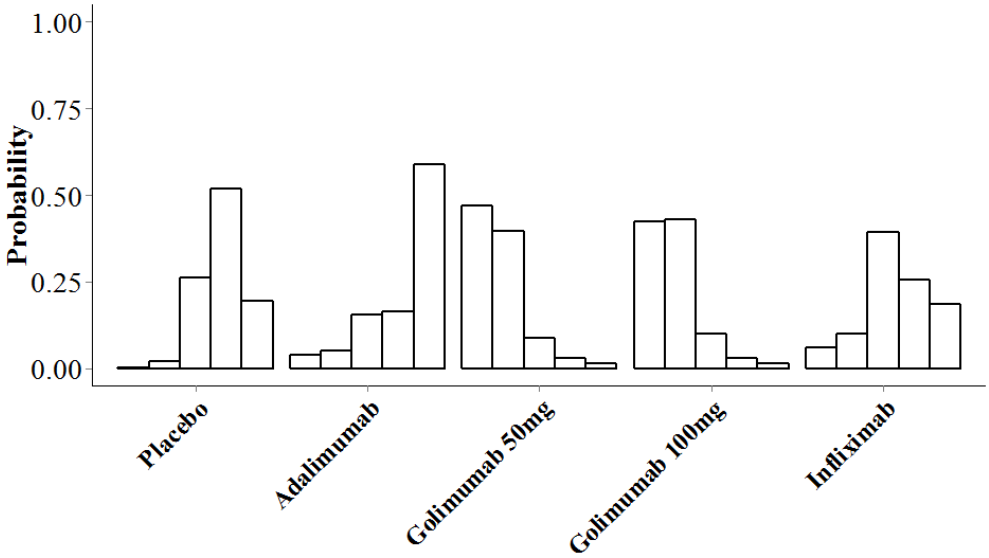


Figure 21: Base case – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 15 presents the probabilities of achieving each of the following categories: no response, response and remission for the base case maintenance phase at 8-32 weeks for patients starting in remission. Golimumab 50mg and 100mg were associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response or remission no response at 8-32 weeks.

Table 15: Base case – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission

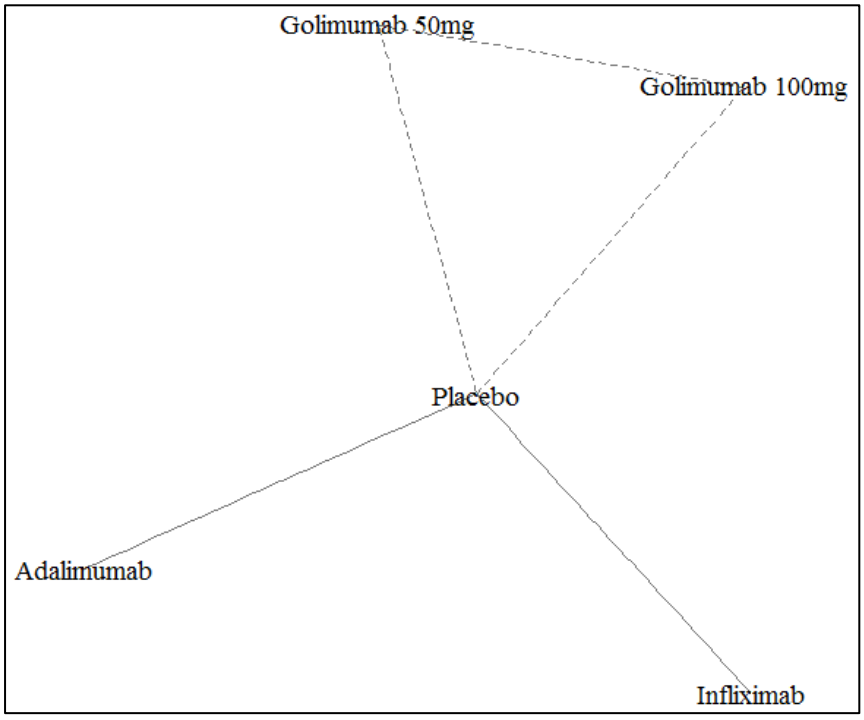
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.353	0.347	0.168, 0.572	0.180	0.174	0.070, 0.316	0.467	0.466	0.225, 0.708
Adalimumab	0.428	0.420	0.099, 0.803	0.166	0.164	0.053, 0.297	0.406	0.392	0.083, 0.804
Golimumab 50mg	0.177	0.152	0.027, 0.457	0.136	0.131	0.028, 0.283	0.687	0.708	0.321, 0.933
Golimumab 100mg	0.182	0.158	0.029, 0.469	0.138	0.134	0.030, 0.285	0.680	0.700	0.322, 0.929
Infliximab	0.325	0.309	0.084, 0.648	0.169	0.165	0.057, 0.304	0.506	0.509	0.178, 0.829

a.3 Base case – maintenance phase 32-52 weeks

a.3.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase for patients starting in response at 32-52 weeks. Data were available from 3 studies comparing two or three treatments. Figure 22 presents the network of evidence for the base case maintenance phase at 32-52 weeks for patients starting in response.

Figure 22: Base case – network of evidence for the maintenance phase at 32-52 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 23 presents the effects of each treatment relative to placebo on the probit scale for the base case maintenance phase 32-52 weeks for patients starting in response. Figure 24 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 12.88, being close to the total number of data points included in the analysis, 14. The between-study standard deviation was estimated to be 0.21 (95% CrI: 0.01, 0.71), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except adalimumab and golimumab 100mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 50mg. However, none of the treatment effects were statistically significant at a conventional 5% level. Infliximab was associated with the greatest effect -0.36 (95% CrI: -1.33, 0.62) and was most likely to be the most effective treatment (probability of being the best = 0.56).

Figure 23: Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))

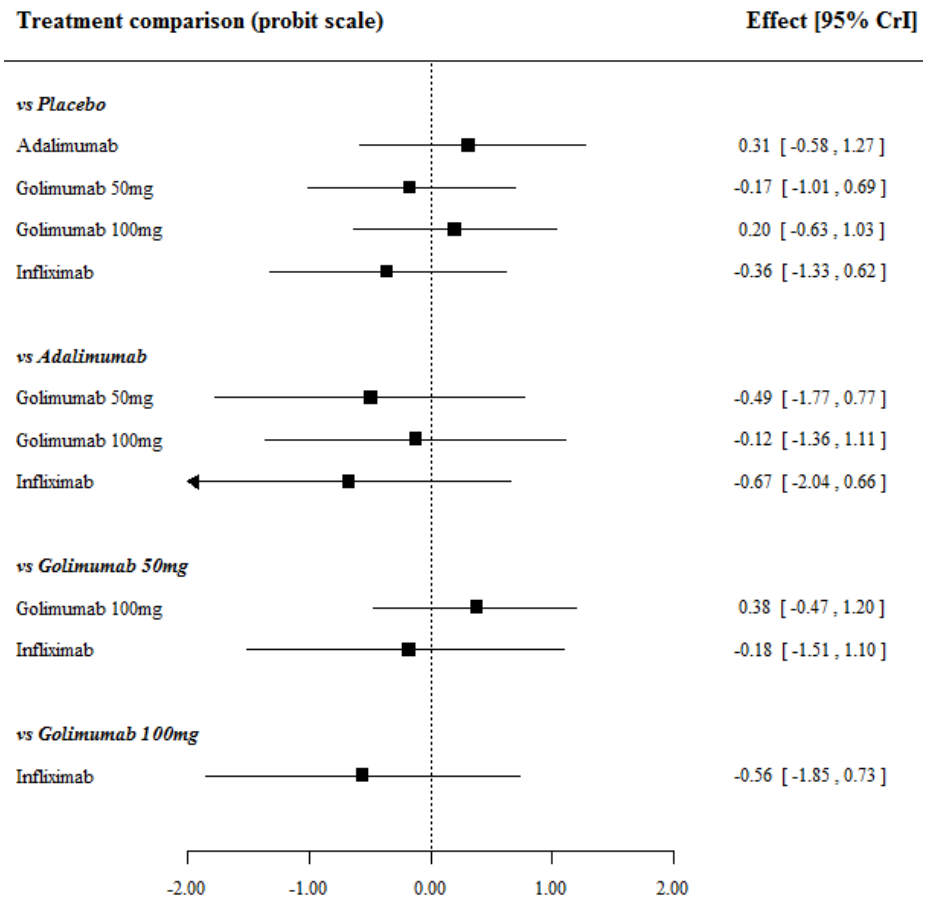
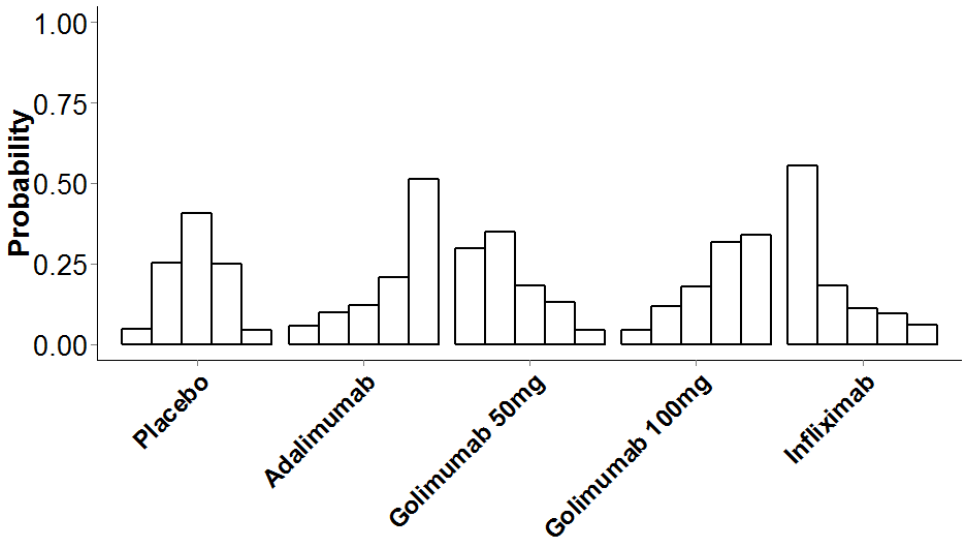


Figure 24: Base case – Ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 16 presents the probabilities of achieving each of the following categories: no response, response and remission for the base case maintenance phase at 32-52 weeks for patients starting in response. Infliximab was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 32-52 weeks. The probabilities of staying in the response state were comparable among treatments at 32-52 weeks.

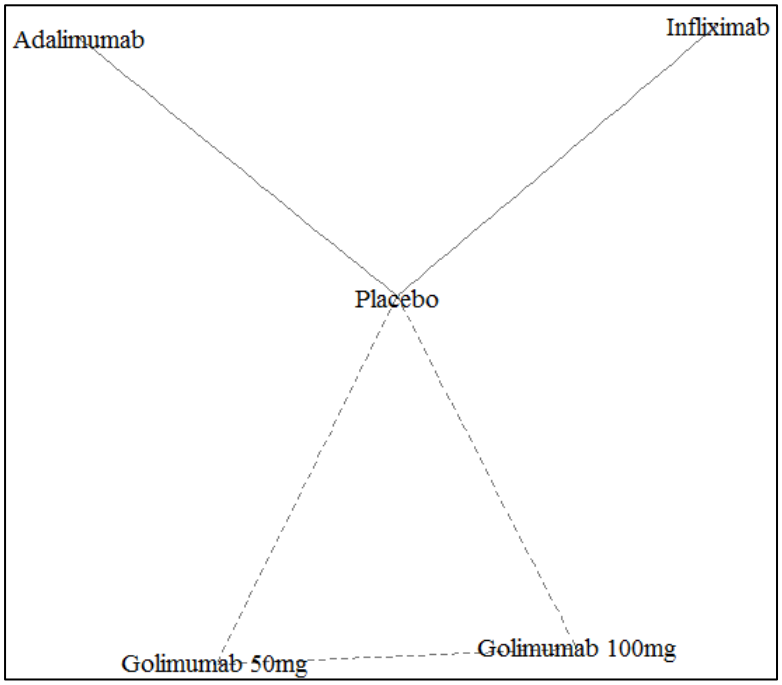
Table 16: Base case – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.338	0.319	0.066, 0.711	0.370	0.378	0.122, 0.604	0.292	0.259	0.027, 0.717
Adalimumab	0.450	0.440	0.063, 0.889	0.327	0.340	0.067, 0.562	0.223	0.167	0.005, 0.716
Golimumab 50mg	0.295	0.258	0.025, 0.750	0.353	0.363	0.081, 0.616	0.352	0.319	0.021, 0.842
Golimumab 100mg	0.410	0.393	0.055, 0.852	0.342	0.353	0.083, 0.581	0.248	0.199	0.009, 0.741
Infliximab	0.250	0.205	0.013, 0.716	0.341	0.353	0.065, 0.621	0.409	0.385	0.029, 0.892

a.3.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 32-52 weeks for patients starting in remission. Data were available from 3 studies comparing two or three treatments. Figure 25 presents the network of evidence for the base case maintenance phase at 32-52 weeks for patients starting in remission.

Figure 25: Base case – network of evidence for the maintenance phase at 32-52 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 26 presents the effects of each treatment relative to placebo on the probit scale for the base case maintenance phase at 32-52 weeks for patients starting in remission. Figure 27 presents the probabilities of treatment rankings for this analysis. The model fitted the data well, with the total residual deviance, 18.46, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.21 (95% CrI: 0.01, 0.72), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with adalimumab. However, only the treatment effects of adalimumab were statistically significant at a conventional 5% level. Adalimumab was associated with the greatest effect -1.04 (95% CrI -1.93, -0.12) and was most likely to be the most effective treatment (probability of being the best = 0.84).

Figure 26: Base case – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))

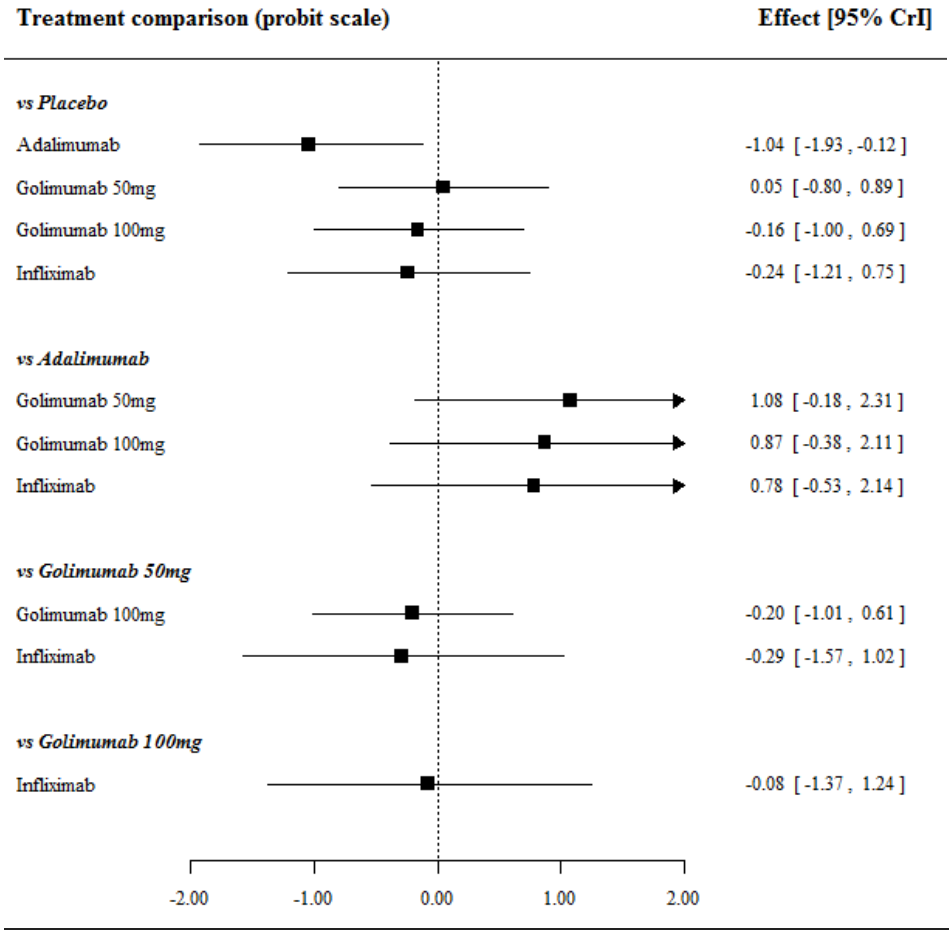
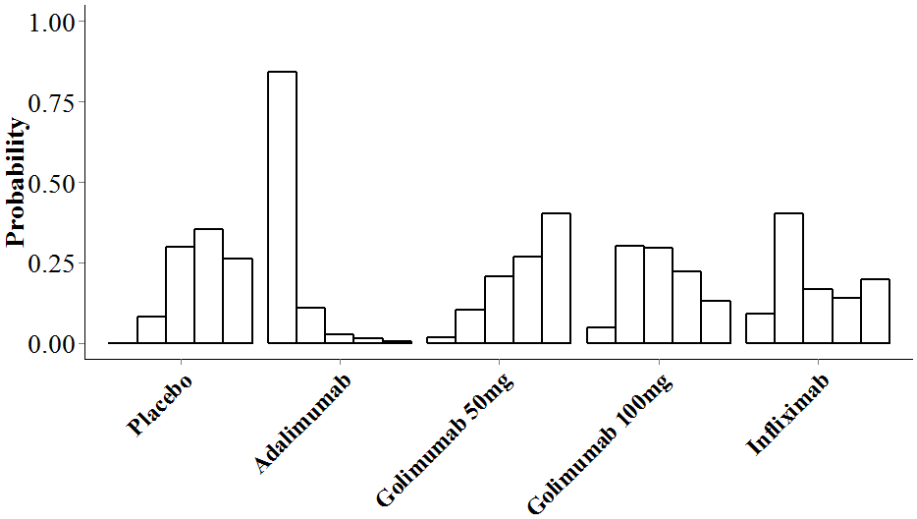


Figure 27: Base case – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 17 presents the probabilities of achieving each of the following categories: no response, response and remission for the base case maintenance phase at 32-52 weeks for patients starting in remission. Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response or from remission to no response at 32-52 weeks.

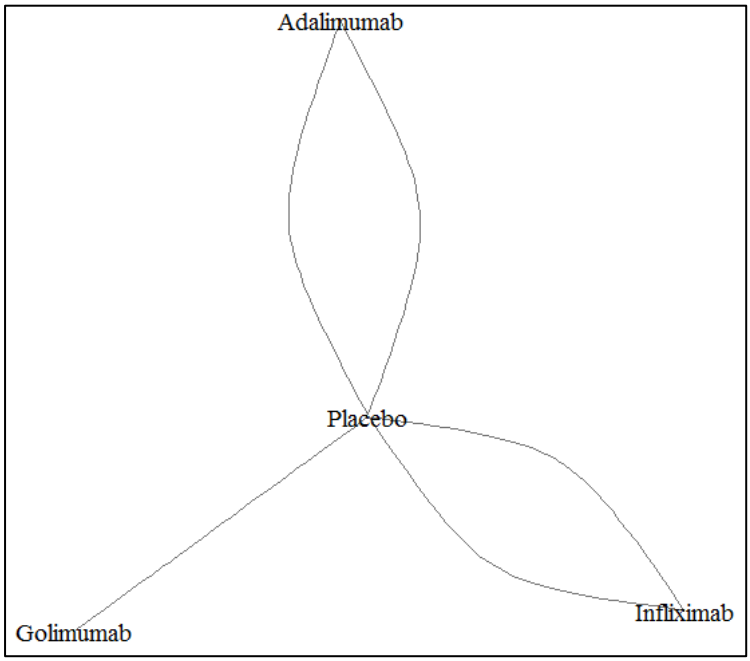
Table 17: Base case – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.301	0.296	0.174, 0.449	0.164	0.147	0.029, 0.449	0.536	0.548	0.237, 0.734
Adalimumab	0.081	0.059	0.005, 0.288	0.084	0.061	0.005, 0.337	0.834	0.874	0.447, 0.985
Golimumab 50mg	0.329	0.314	0.080, 0.664	0.155	0.141	0.024, 0.415	0.515	0.523	0.135, 0.851
Golimumab 100mg	0.266	0.245	0.052, 0.604	0.147	0.132	0.020, 0.417	0.587	0.604	0.169, 0.894
Infliximab	0.247	0.220	0.033, 0.613	0.140	0.126	0.017, 0.413	0.613	0.634	0.174, 0.928

a.4 Sensitivity analysis 1 – induction phase

Sensitivity analysis 1 involved replacing ULTRA2 anti-TNF-naïve data with ULTRA2 ITT data. A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the induction phase. Data were available from 5 studies comparing two treatments. Figure 28 presents the network of evidence for the sensitivity analysis 1 induction phase.

Figure 28: Sensitivity analysis 1 – network of evidence for the induction phase



Note: solid line indicates a 2-arm trial

Figure 29 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 1 induction phase. Figure 30 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 17.08, being close to the total number of data points included in the analysis, 20. The between-study standard deviation was estimated to be 0.11 (95% CrI: 0.01, 0.47), which implies mild heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab (-0.91; 95% CrI: -1.25, -0.57). All treatment effects were statistically significant at a conventional 5% level. Infliximab was most likely to be the most effective treatment (probability of being the best = 0.94).

Figure 29: Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))

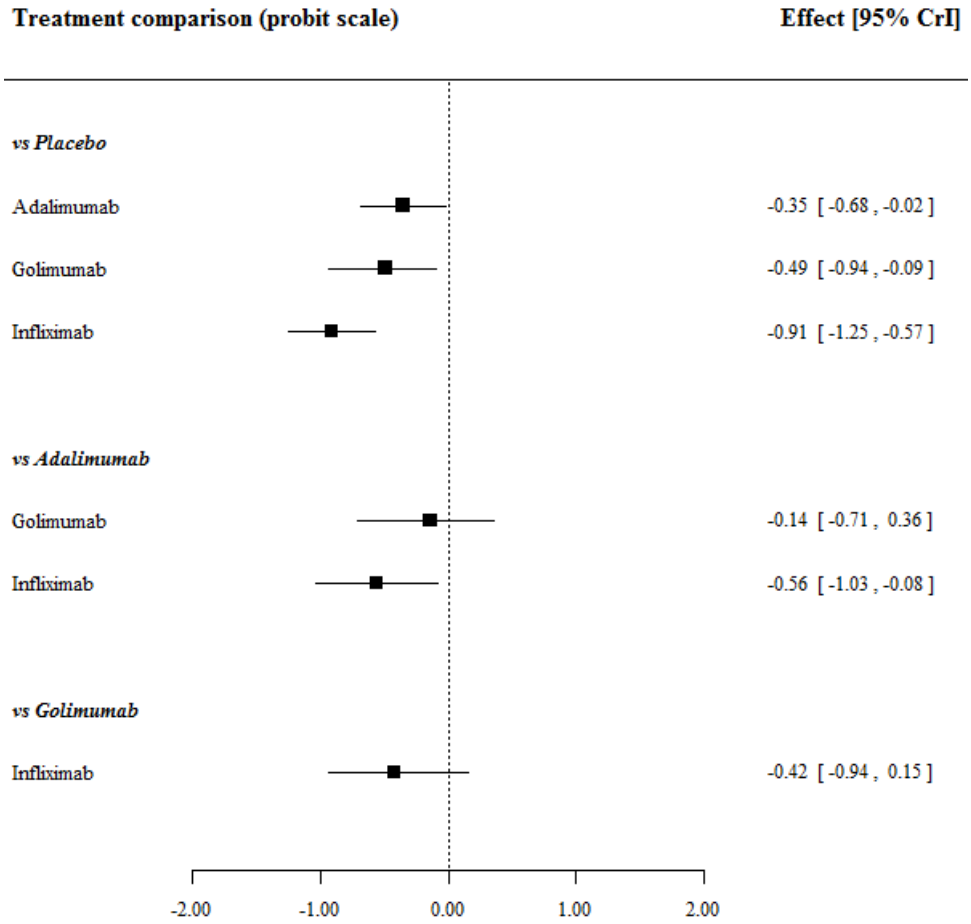
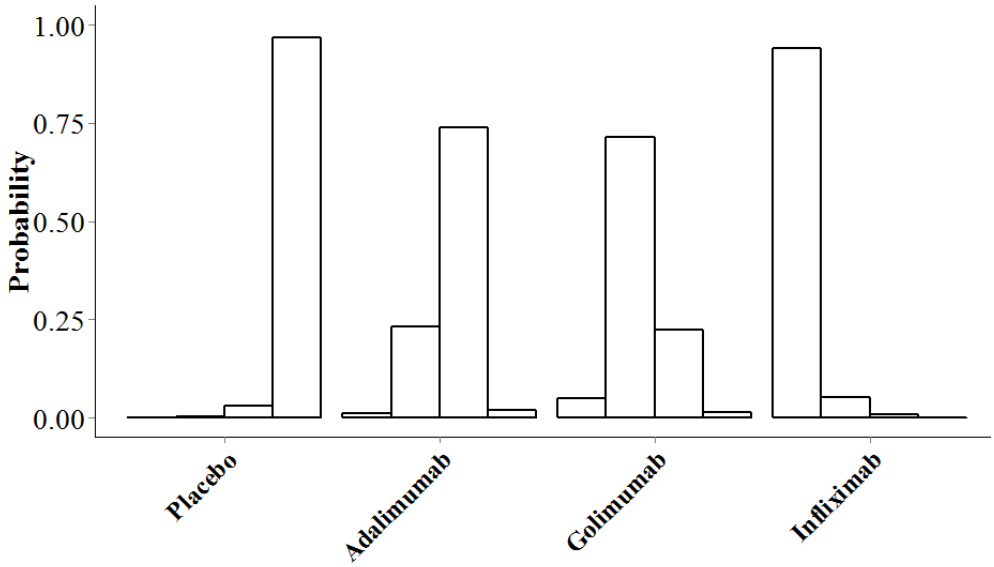


Figure 30: Sensitivity analysis 1 – ranking probability histograms for the induction phase



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 18 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 1 induction phase. Infliximab was associated with the highest probability of moving from no response to response and from no response to remission.

Table 18: Sensitivity analysis 1 – probabilities of being in each category for the induction phase

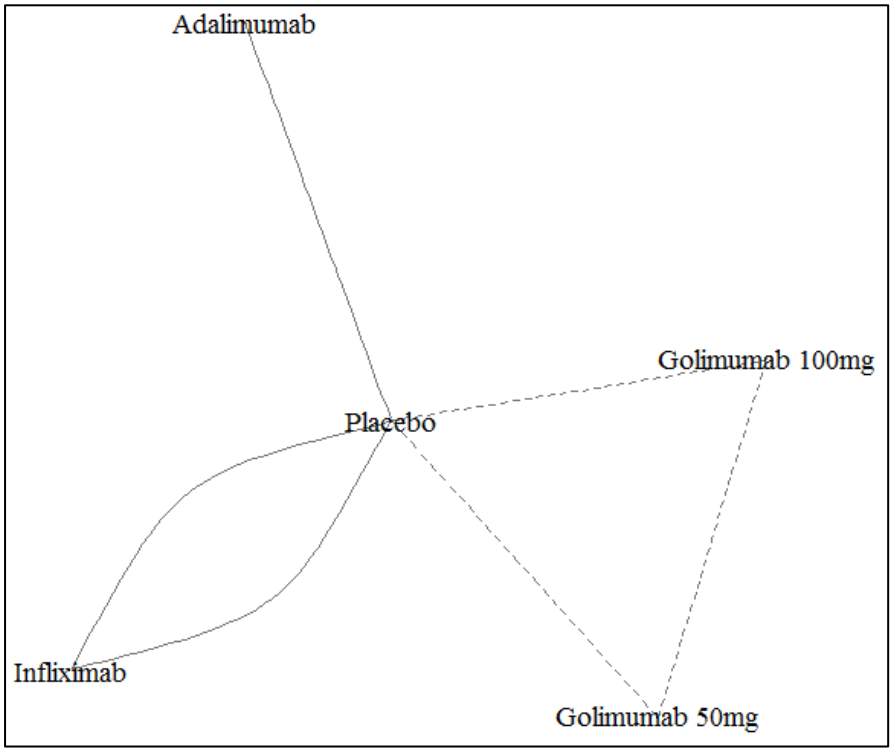
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.649	0.649	0.586, 0.710	0.255	0.255	0.212, 0.298	0.096	0.095	0.062, 0.140
Adalimumab	0.513	0.512	0.372, 0.652	0.315	0.317	0.240, 0.375	0.173	0.169	0.088, 0.286
Golimumab	0.456	0.456	0.283, 0.631	0.330	0.334	0.250, 0.389	0.214	0.207	0.101, 0.368
Infliximab	0.302	0.298	0.180, 0.441	0.351	0.353	0.281, 0.409	0.347	0.345	0.208, 0.506

a.5 Sensitivity analysis 1 – maintenance phase 8-32 weeks

a.5.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase for patients starting in response at 8-32 weeks. Data were available from 4 studies comparing two or three treatments. Figure 31 presents the network of evidence for the aensitivity analysis 1 maintenance phase at 8-32 weeks for patients starting in response.

Figure 31: Sensitivity analysis 1 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 32 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 1 maintenance phase at 8-32 weeks for patients starting in response. Figure 33 presents the probabilities of treatment rankings for this analysis. There was some suggestion that model did not represent the data well with the total residual deviance, 11.54, being smaller than would be expected given the total number of data points included in the analysis, 18. The probability of observing a value less than 11.54 was 0.130, which means that this could be a chance event. Similar to the base case analysis, all 4 studies had smaller residual deviances than expected. The between-study standard deviation was estimated to be 0.17 (95% CrI: 0.01, 0.63), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 100mg (-0.41 95% CrI: -1.06, 0.22). However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 100mg was most likely to be the most effective treatment (probability of being the best = 0.43).

Figure 32: Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32weeks for patients starting in response (SD~HN(0,0.32^2))

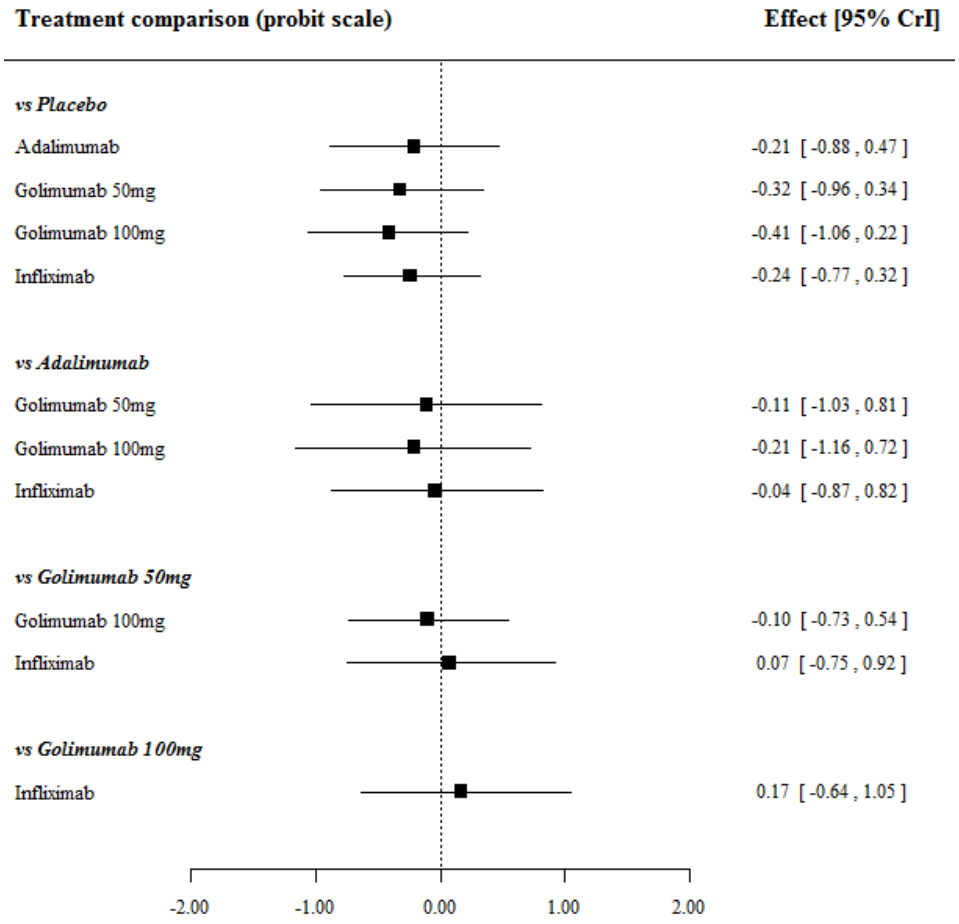
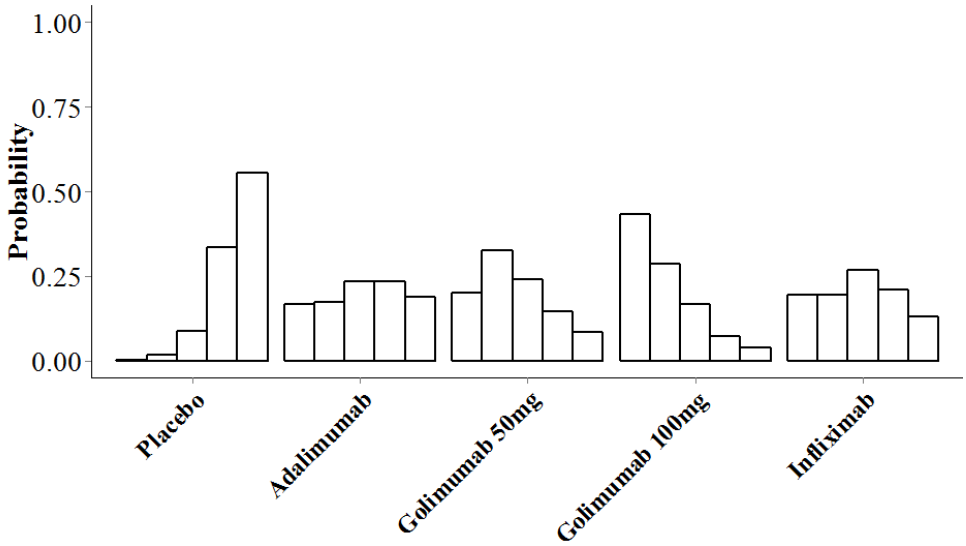


Figure 33: Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 19 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 1 maintenance phase at 8-32 weeks starting in response. Golimumab 100mg was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 8-32 weeks. The probabilities of staying in the response state were comparable among treatments.

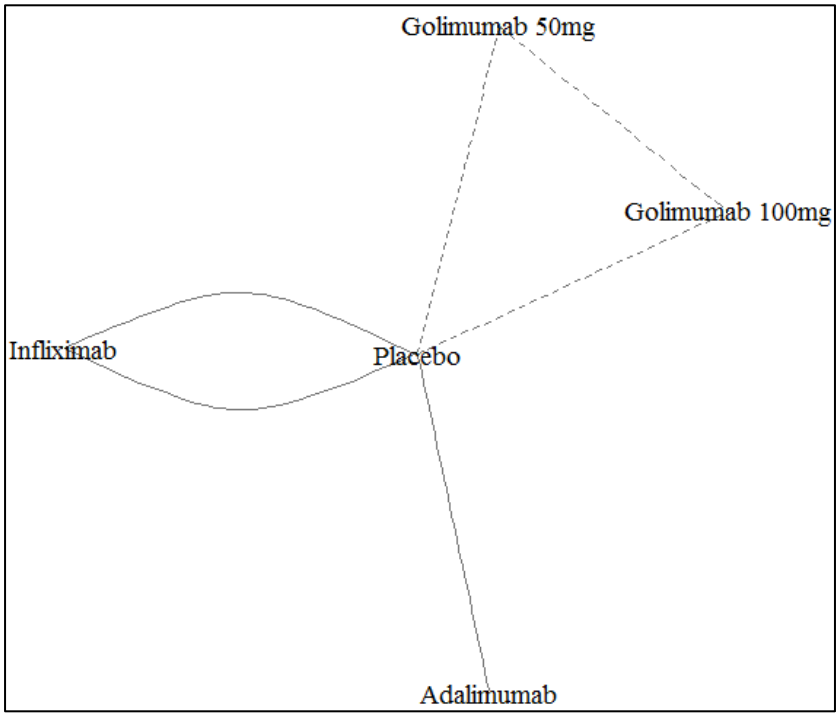
Table 19: Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.548	0.548	0.447, 0.649	0.269	0.267	0.199, 0.361	0.183	0.181	0.085, 0.282
Adalimumab	0.468	0.467	0.210, 0.744	0.279	0.283	0.158, 0.391	0.252	0.240	0.059, 0.525
Golimumab 50mg	0.427	0.423	0.190, 0.688	0.289	0.289	0.176, 0.412	0.284	0.273	0.081, 0.552
Golimumab 100mg	0.390	0.384	0.162, 0.649	0.293	0.292	0.182, 0.421	0.318	0.310	0.098, 0.591
Infliximab	0.453	0.451	0.237, 0.685	0.286	0.287	0.186, 0.403	0.260	0.252	0.082, 0.490

a.5.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 8-32 weeks for patients starting in remission. Data were available from 4 studies comparing two or three treatments. Figure 34 presents the network of evidence for the sensitivity analysis 1 maintenance phase at 8-32 weeks for patients starting in remission.

Figure 34: Sensitivity analysis 1 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 35 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 1 maintenance phase at 8-32 weeks for patients starting in remission. Figure 36 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 15.29, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.19 (95% CrI: 0.01, 0.65), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effects being associated with golimumab 50mg (-0.62; 95% CrI: -1.36, 0.11) and golimumab 100mg (-0.61; 95% CrI: -1.34, 0.13). However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 50mg and golimumab 100mg were most likely to be the most effective treatments (probability of being the best = 0.44 and 0.41, respectively).

Figure 35: Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))

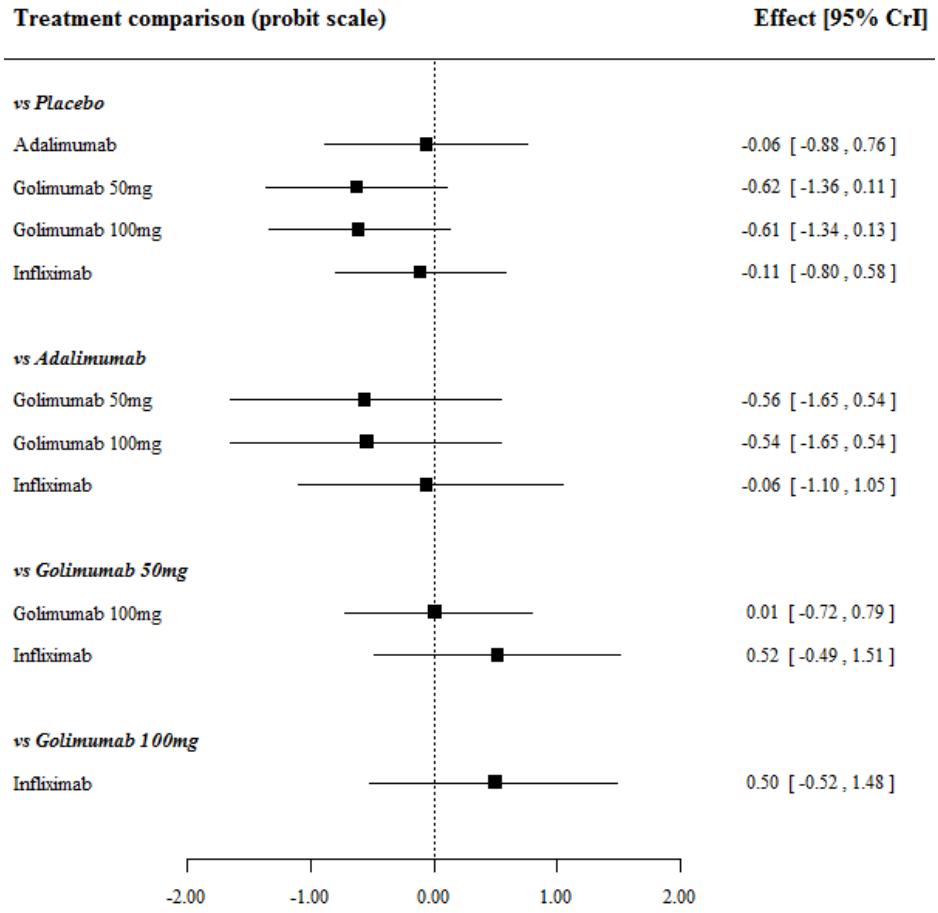
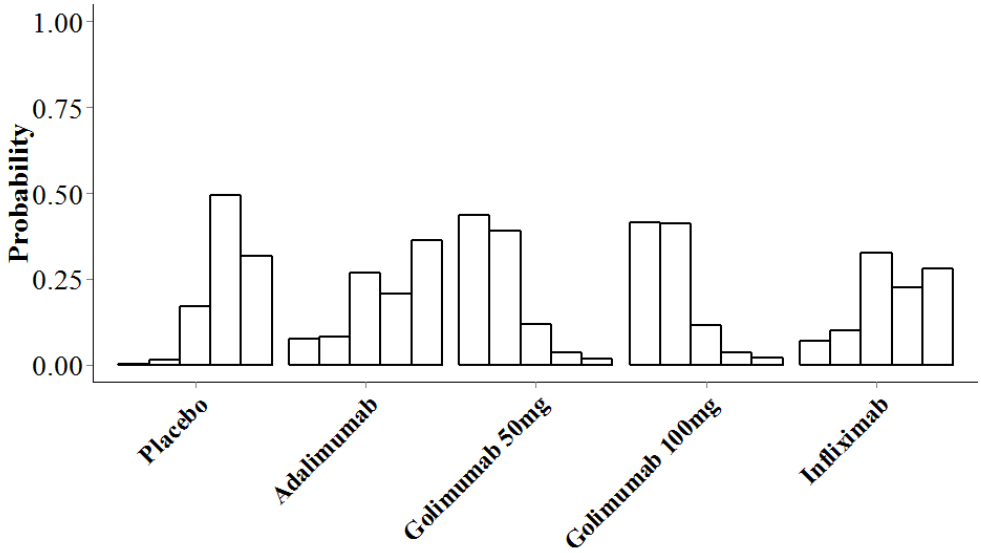


Figure 36: Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 20 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 1 maintenance phase at 8-32 weeks starting in remission. Golimumab 50mg and 100mg were associated with the highest probability of staying in remission, and the smallest probability of moving from remission to no response and from remission to response at 8-32 weeks.

Table 20: Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission

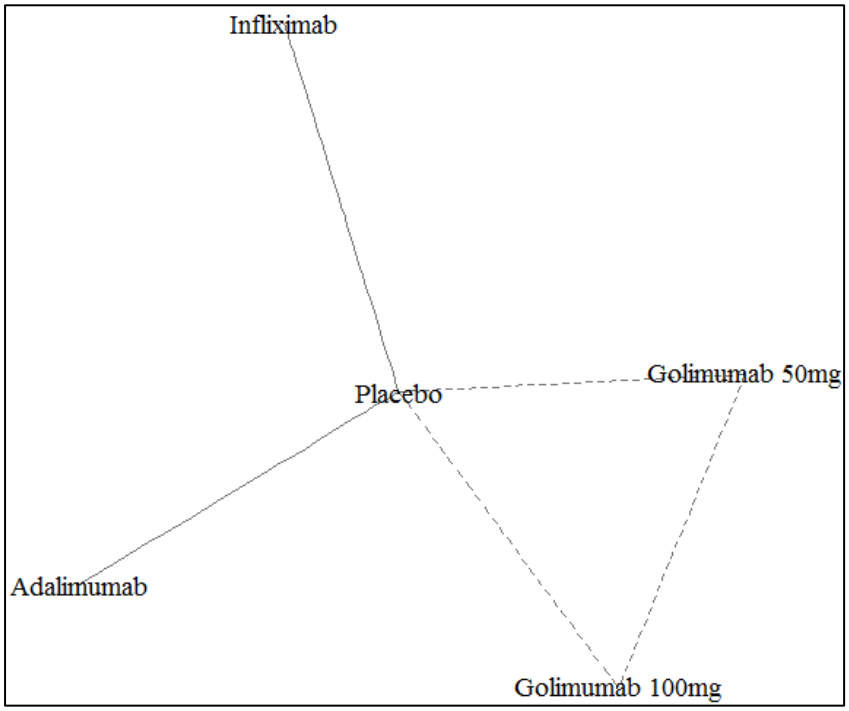
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.392	0.389	0.217, 0.584	0.180	0.175	0.078, 0.309	0.428	0.426	0.218, 0.650
Adalimumab	0.379	0.365	0.096, 0.736	0.167	0.164	0.06, 0.3	0.454	0.450	0.128, 0.807
Golimumab 50mg	0.204	0.182	0.036, 0.493	0.143	0.139	0.036, 0.285	0.653	0.669	0.300, 0.914
Golimumab 100mg	0.207	0.188	0.038, 0.494	0.144	0.140	0.037, 0.287	0.648	0.662	0.303, 0.911
Infliximab	0.359	0.347	0.107, 0.679	0.170	0.166	0.065, 0.301	0.471	0.470	0.165, 0.793

a.6 Sensitivity analysis 1 – maintenance phase 32-52 weeks

a.6.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase for patients starting in response at 32-52 weeks. Data were available from 3 studies comparing two or three treatments. Figure 37 presents the network of evidence for the sensitivity analysis 1 maintenance phase at 32-52 weeks for patients starting in response.

Figure 37: Sensitivity analysis 1 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 38 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 1 maintenance phase 32-52 weeks for patients starting in response. Figure 39 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 12.32, being close to the total number of data points included in the analysis, 14. The between-study standard deviation was estimated to be 0.21 (95% CrI: 0.01, 0.72), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except golimumab 100mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab (-0.37; 95% CrI: -1.30, 0.59). However, none of the treatment effects were statistically significant at a conventional 5% level. Infliximab was most likely to be the most effective treatment (probability of being the best = 0.52).

Figure 38: Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))

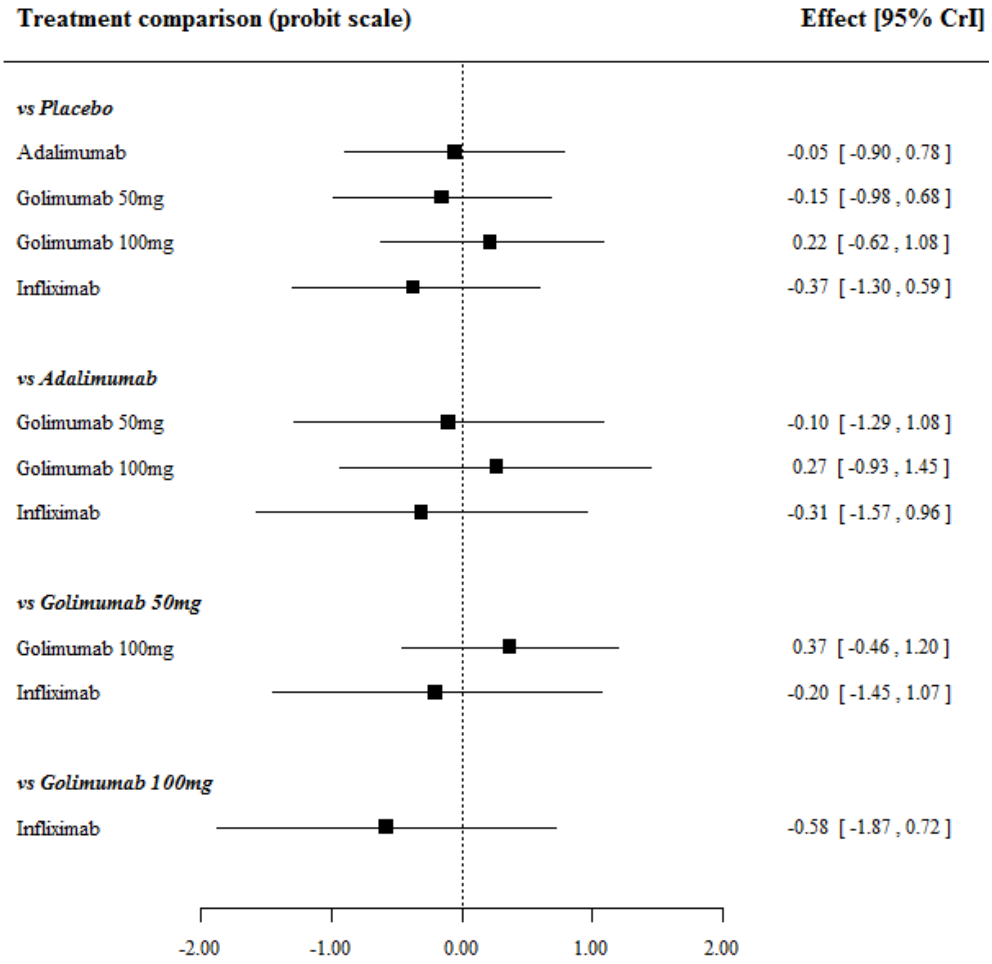
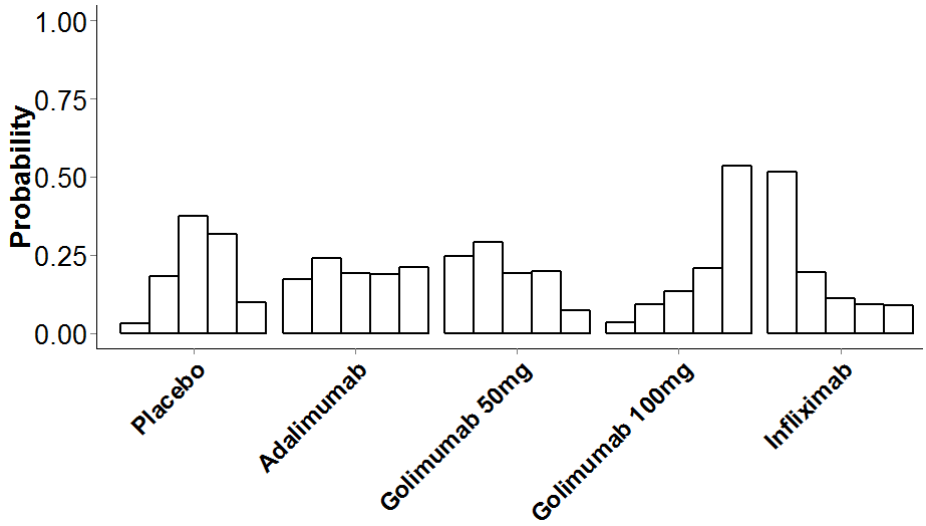


Figure 39: Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 21 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 1 maintenance phase at 32-52 weeks starting in response. Infliximab was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 32-52 weeks. The probabilities of staying in the response state were comparable among treatments at 32-52 weeks.

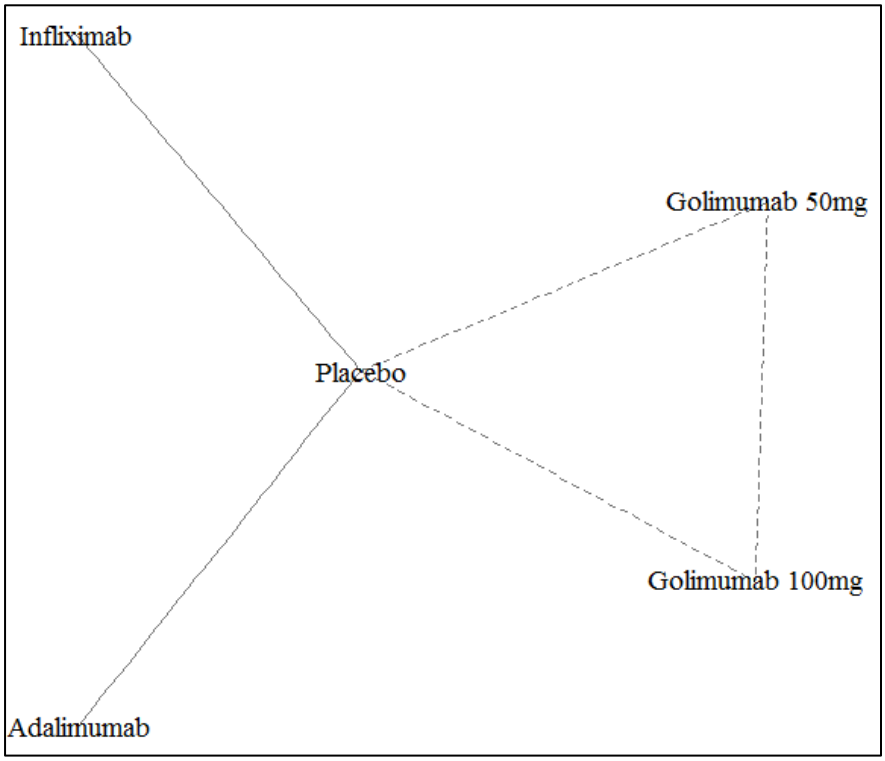
Table 21: Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.338	0.317	0.064, 0.718	0.373	0.378	0.118, 0.628	0.29	0.259	0.019, 0.714
Adalimumab	0.332	0.300	0.031, 0.790	0.354	0.363	0.084, 0.625	0.314	0.276	0.013, 0.812
Golimumab 50mg	0.302	0.269	0.024, 0.769	0.355	0.364	0.077, 0.632	0.344	0.308	0.015, 0.844
Golimumab 100mg	0.417	0.401	0.053, 0.854	0.341	0.352	0.077, 0.595	0.242	0.192	0.006, 0.742
Infliximab	0.249	0.201	0.012, 0.730	0.343	0.353	0.063, 0.643	0.408	0.387	0.022, 0.898

a.6.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 32-52 weeks for patients starting in remission. Data were available from 3 studies comparing two or three treatments. Figure 40 presents the network of evidence for the sensitivity analysis 1 maintenance phase at 32-52 weeks for patients starting in remission.

Figure 40: Sensitivity analysis 1 – network of evidence for the maintenance phase at 32-52weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 41 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 1 maintenance phase at 32-52 weeks for patients starting in remission. Figure 42 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 17.73, being close to the total number of data points included in the analysis, 14. The between-study standard deviation was estimated to be 0.22 (95% CrI: 0.01, 0.72), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with adalimumab (-0.86; 95% CrI: -1.71, 0.00). However, only the treatment effects of adalimumab were statistically significant at a conventional 5% level. Adalimumab was most likely to be the most effective treatment (probability of being the best = 0.78).

Figure 41: Sensitivity analysis 1 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))

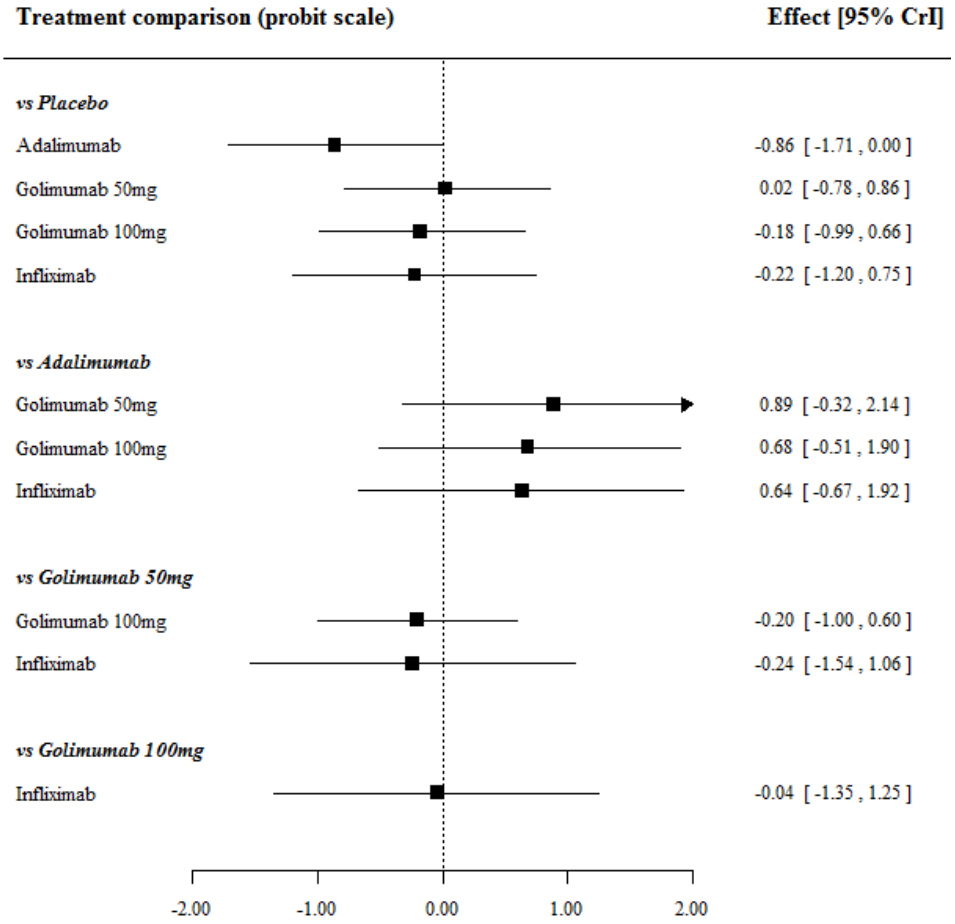
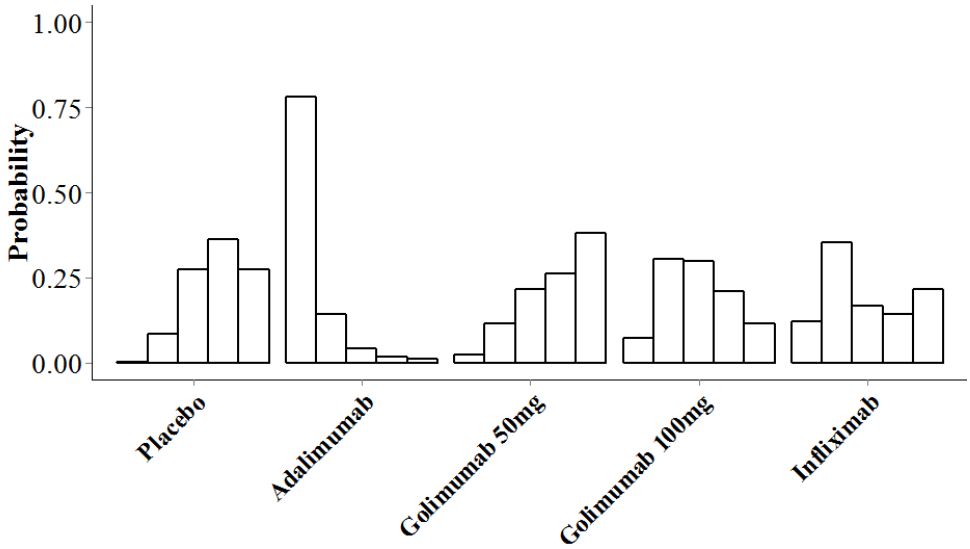


Figure 42: Sensitivity analysis 1 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 22 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 1 maintenance phase at 32-52 weeks starting in remission. Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response or from remission to no response at 32-52 weeks.

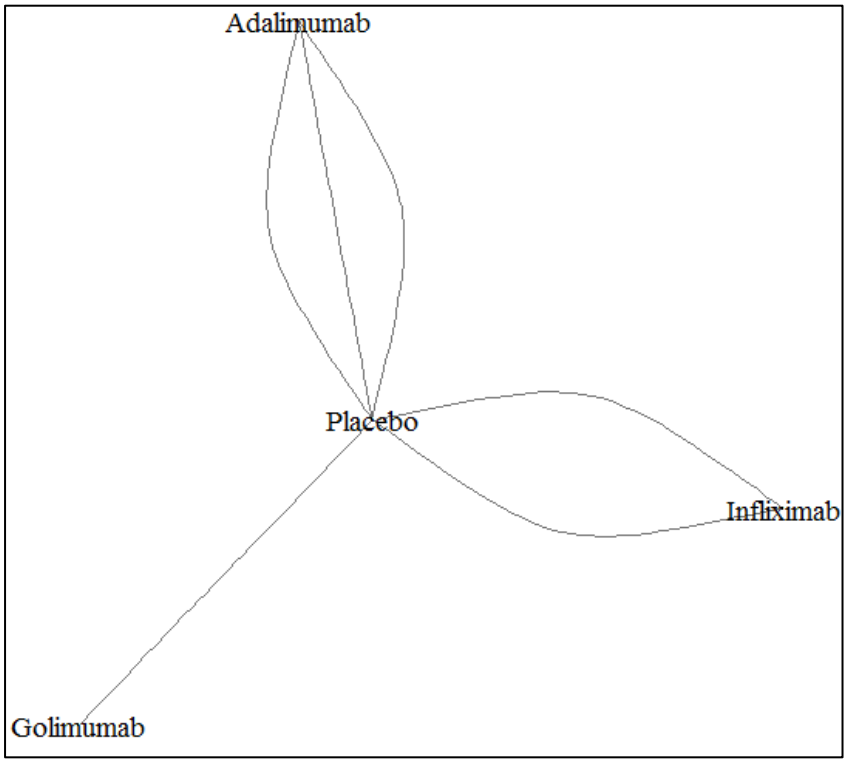
Table 22: Sensitivity analysis 1 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.302	0.299	0.177, 0.448	0.167	0.155	0.030, 0.396	0.530	0.539	0.270, 0.727
Adalimumab	0.104	0.082	0.010, 0.324	0.098	0.081	0.008, 0.308	0.797	0.828	0.426, 0.974
Golimumab 50mg	0.324	0.307	0.082, 0.664	0.158	0.147	0.026, 0.374	0.519	0.526	0.145, 0.842
Golimumab 100mg	0.260	0.239	0.053, 0.591	0.149	0.136	0.022, 0.374	0.591	0.605	0.200, 0.890
Infliximab	0.254	0.225	0.035, 0.620	0.144	0.132	0.019, 0.367	0.603	0.620	0.186, 0.918

a.7 Sensitivity analysis 2 – induction phase

Sensitivity analysis 2 involved including Suzuki *et al*. A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the induction phase. Data were available from 6 studies comparing two treatments. Figure 43 presents the network of evidence for the sensitivity analysis 2 induction phase.

Figure 43: Sensitivity analysis 2 – network of evidence for the induction phase



Note: solid line indicates a 2-arm trial.

Figure 44 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 2 induction phase. Figure 45 presents the probabilities of treatment rankings for this analysis. The model fitted the data well, with the total residual deviance, 24.36, being close to the total number of data points included in the analysis, 24. The between-study standard deviation was estimated to be 0.10 (95% CrI: 0.01, 0.41), which implies mild heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab (-0.92, 95% CrI: -1.24, -0.60). All treatment effects were statistically significant at a conventional 5% level. Infliximab was most likely to be the most effective treatment (probability of being the best = 0.96).

Figure 44: Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))

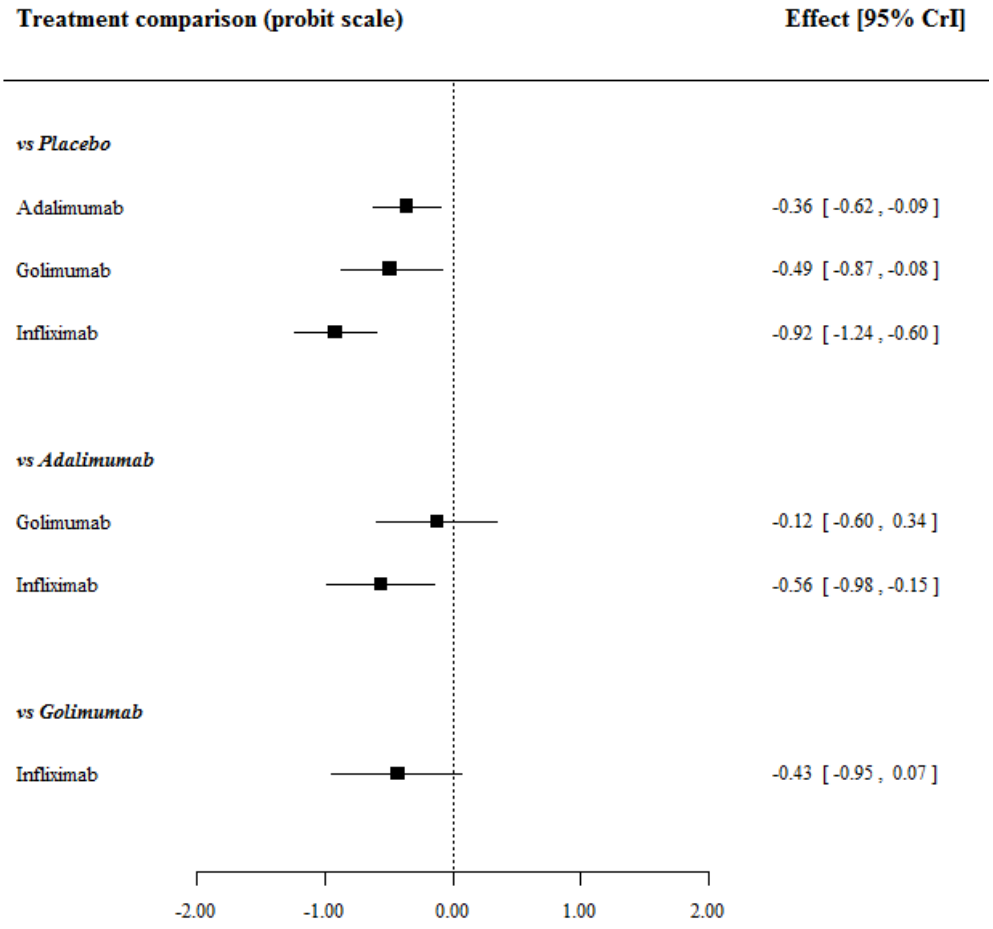
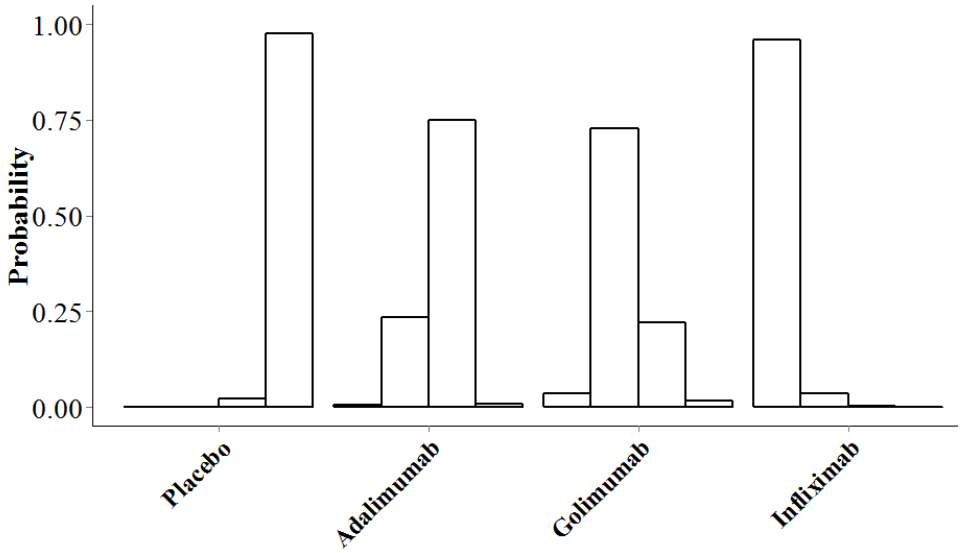


Figure 45: Sensitivity analysis 2 – ranking probability histograms for the induction phase



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 23 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 2 induction phase. Infliximab was associated with the highest probability of moving from no response to response and from no response to remission.

Table 23: Sensitivity analysis 2 – probabilities of being in each category for the induction phase

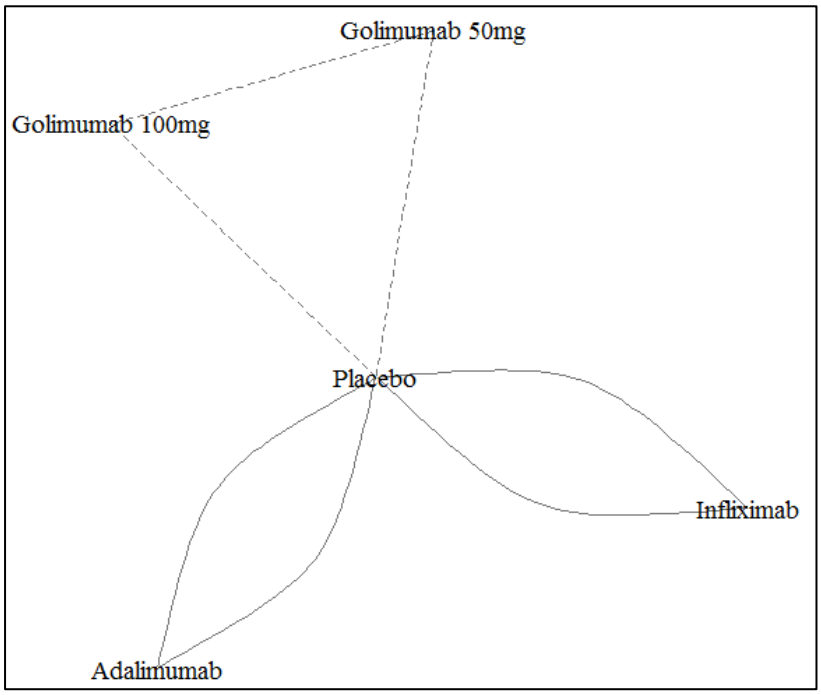
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.642	0.642	0.585, 0.696	0.260	0.263	0.225, 0.301	0.095	0.094	0.065, 0.132
Adalimumab	0.502	0.500	0.384, 0.623	0.326	0.327	0.264, 0.377	0.173	0.170	0.099, 0.264
Golimumab	0.452	0.450	0.298, 0.624	0.339	0.343	0.266, 0.392	0.209	0.204	0.060, 0.344
Infliximab	0.292	0.252	0.183, 0.421	0.360	0.360	0.305, 0.411	0.348	0.347	0.068, 0.488

a.8 Sensitivity analysis 2 – maintenance phase 8-32 weeks

a.8.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase response at 8-32 weeks for patients starting in response. Data were available from 5 studies comparing two or three treatments. Figure 46 presents the network of evidence for the sensitivity analysis 2 maintenance phase at 8-32 weeks for patients starting in response.

Figure 46: Sensitivity analysis 2 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 47 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 2 maintenance phase at 8-32 weeks for patients starting in response. Figure 48 presents the probabilities of treatment rankings for this analysis. There was some suggestion that model did not represent the data well with the total residual deviance, 14.80, being smaller than would be expected given the total number of data points included in the analysis, 22. The probability of observing a value less than 14.8 is 0.129, which means that this could be a chance event. Similar to the base case analysis, all 5 studies had smaller residual deviance than expected. The between-study standard deviation was estimated to be 0.16 (95% CrI: 0.01, 0.58), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 100mg (-0.43; 95% CrI: -1.03, 0.19). However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 100mg was most likely to be the most effective treatment (probability of being the best = 0.44).

Figure 47: Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~HN(0,0.32^2))

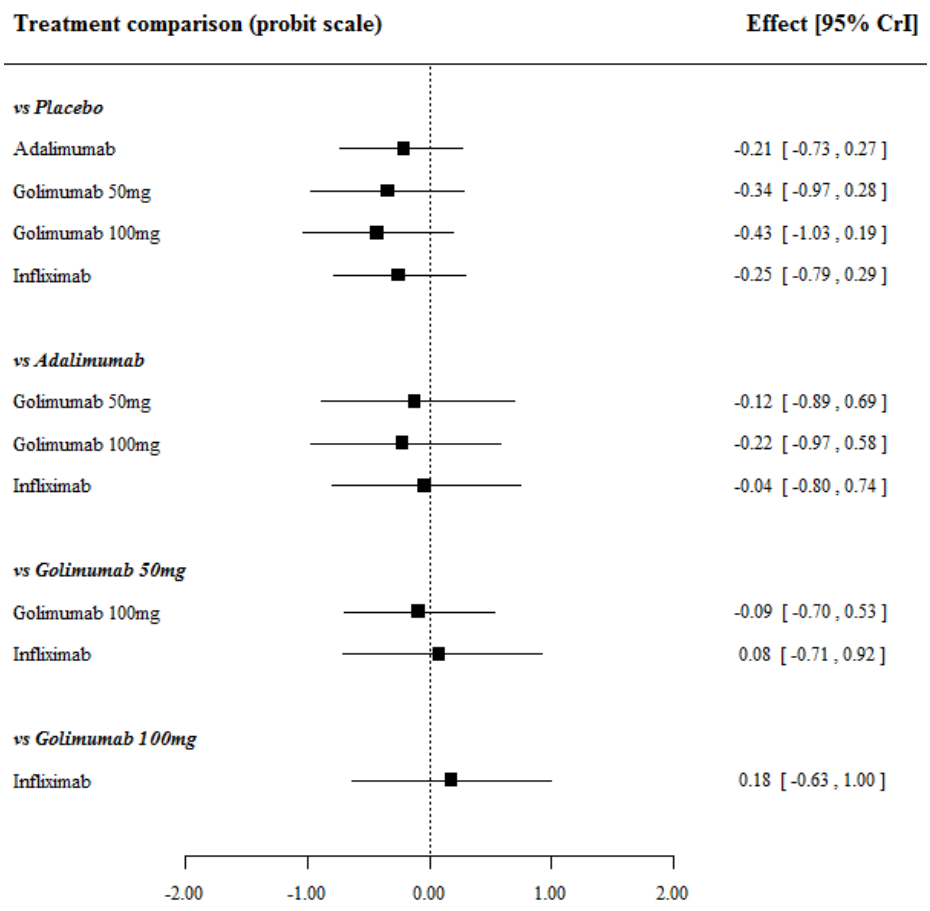
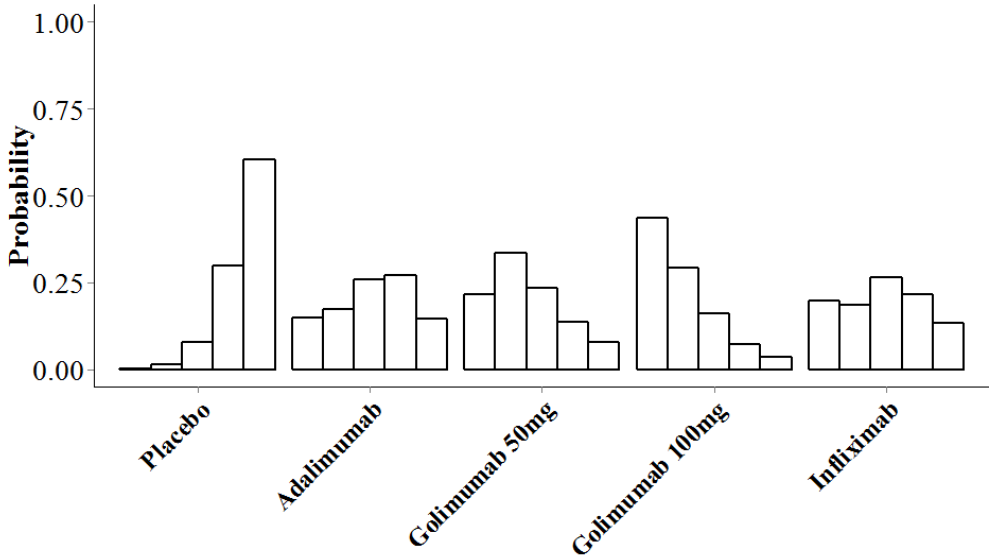


Figure 48: Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 24 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 2 maintenance phase at 8-32 weeks for patients starting in response. Golimumab 100mg was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 8-32 weeks. The probabilities of staying in the response state were comparable among treatments.

Table 24: Sensitivity analysis 2 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response

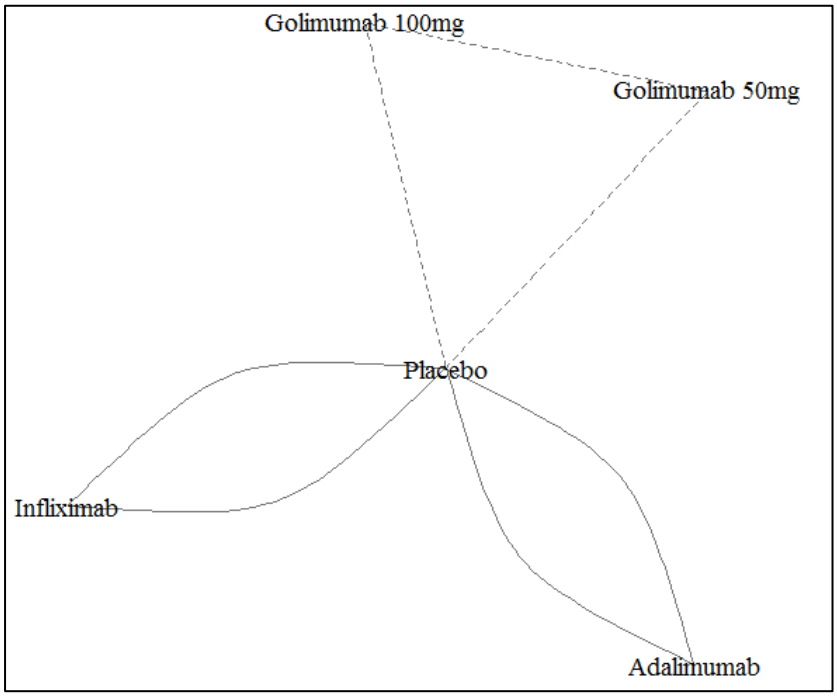
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.525	0.525	0.437, 0.609	0.274	0.274	0.220, 0.329	0.201	0.199	0.130, 0.286
Adalimumab	0.441	0.441	0.238, 0.647	0.286	0.288	0.209, 0.354	0.273	0.263	0.116, 0.485
Golimumab 50mg	0.398	0.393	0.174, 0.649	0.289	0.292	0.199, 0.357	0.313	0.305	0.116, 0.569
Golimumab 100mg	0.363	0.356	0.158, 0.617	0.291	0.293	0.200, 0.360	0.346	0.338	0.132, 0.598
Infliximab	0.429	0.425	0.222, 0.657	0.287	0.289	0.204, 0.352	0.284	0.276	0.110, 0.504

a.8.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 8-32 weeks for patients starting in

remission. Data were available from 5 studies comparing two or three treatments. Figure 49 presents the network of evidence for the sensitivity analysis 2 maintenance phase at 8-32 weeks for patients starting in remission.

Figure 49: Sensitivity analysis 2 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 50 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 2 maintenance phase at 8-32 weeks for patients starting in remission. Figure 51 presents the probabilities of treatment rankings for this analysis. The model fitted the data well, with the total residual deviance, 21.05, being close to the total number of data points included in the analysis, 22. The between-study standard deviation was estimated to be 0.17 (95% CrI: 0.01, 0.60), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except adalimumab were associated with beneficial treatment effects relative to placebo with the greatest effects being associated with golimumab 50mg (-0.61; 95% CrI: -1.30, 0.09) and golimumab 100mg (-0.60; 95% CrI: -1.29, 0.09). However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 50mg and golimumab 100mg were most likely to be the most effective treatments (probability of being the best = 0.46 and 0.44, respectively).

Figure 50: Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))

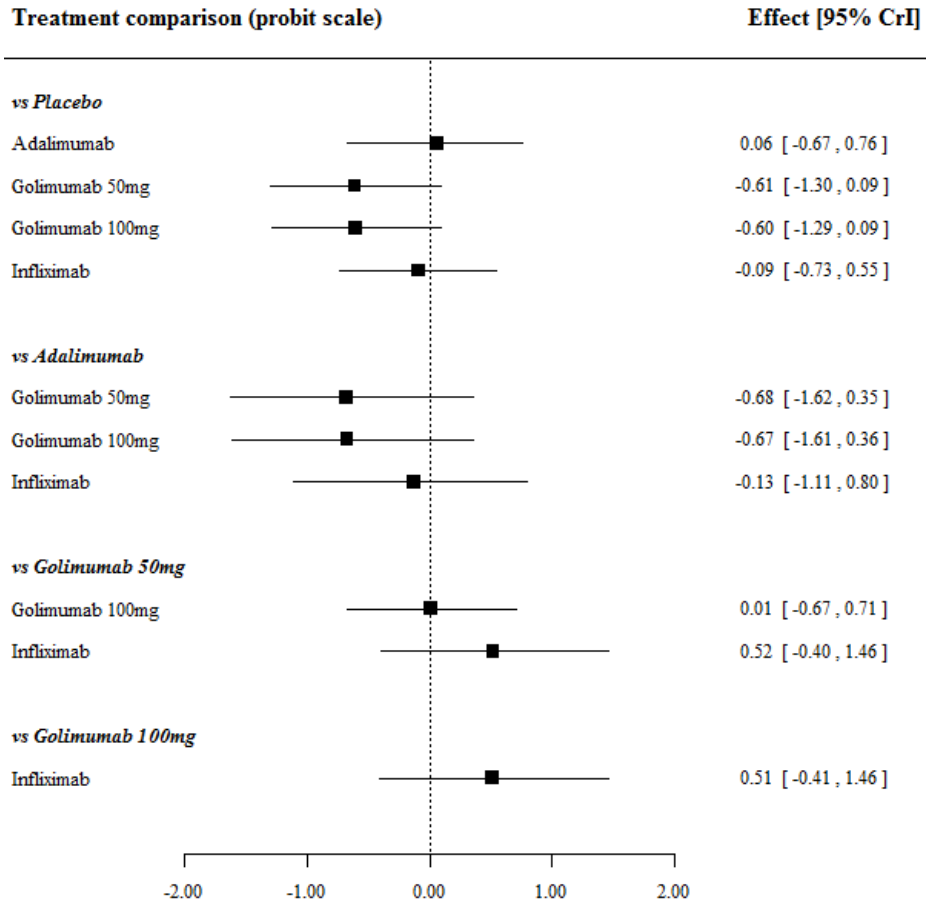
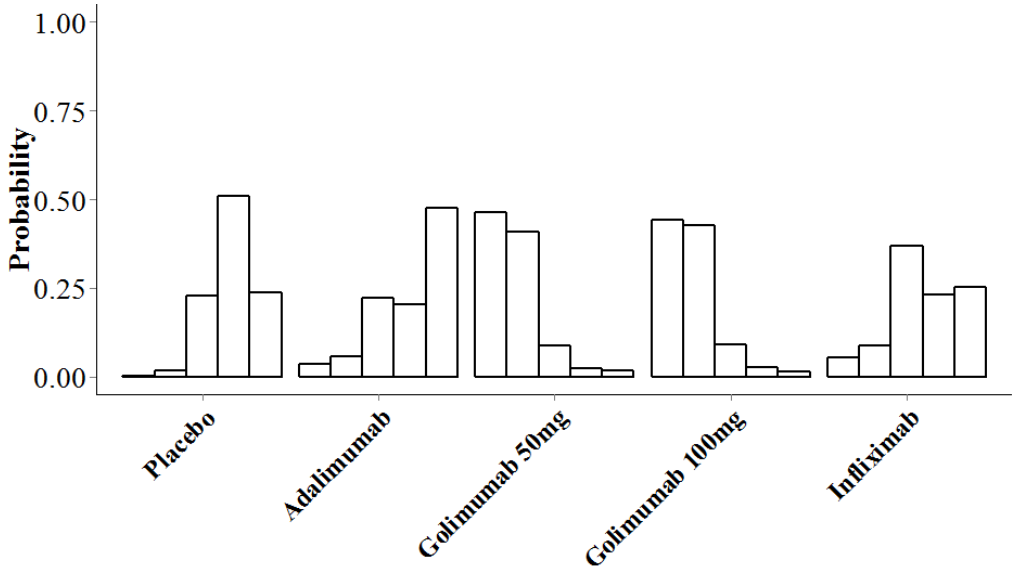


Figure 51: Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 25 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 2 maintenance phase at 8-32 weeks starting in remission. Golimumab 50mg and 100mg were associated with the highest probability of staying in remission, and the smallest probability of moving from remission to no response and from remission to response at 8-32 weeks.

Table 25: Sensitivity analysis 2 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission

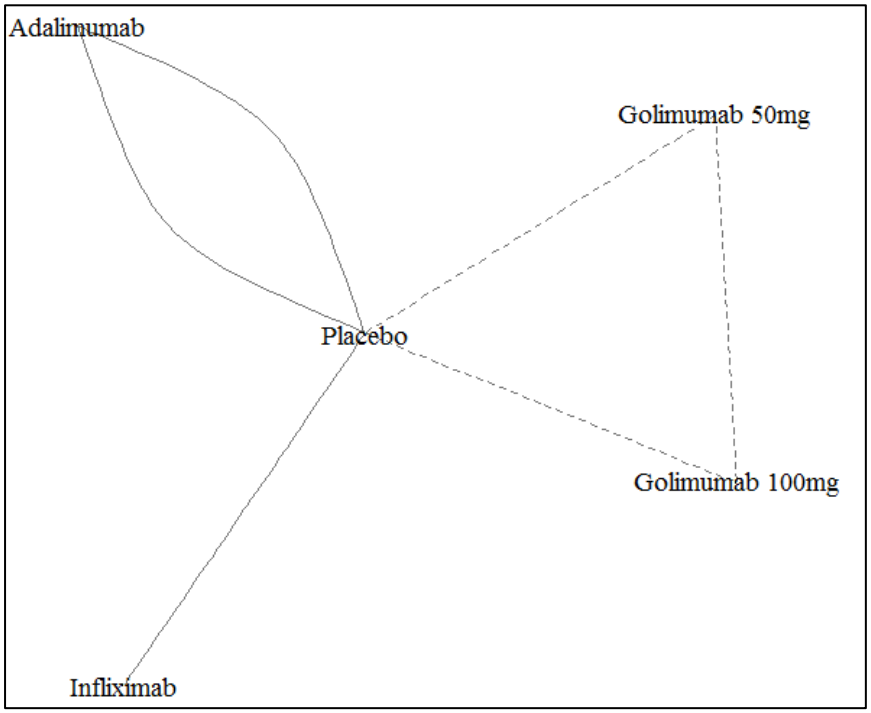
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.354	0.350	0.191, 0.540	0.187	0.183	0.103, 0.294	0.459	0.459	0.250, 0.666
Adalimumab	0.381	0.371	0.114, 0.701	0.178	0.175	0.085, 0.286	0.441	0.434	0.144, 0.769
Golimumab 50mg	0.179	0.159	0.034, 0.442	0.143	0.139	0.043, 0.264	0.678	0.695	0.353, 0.915
Golimumab 100mg	0.181	0.162	0.035, 0.443	0.144	0.141	0.044, 0.264	0.675	0.691	0.349, 0.913
Infliximab	0.331	0.318	0.103, 0.626	0.177	0.174	0.082, 0.286	0.492	0.493	0.195, 0.790

a.9 Sensitivity analysis 2 – maintenance phase 32-52 weeks

a.9.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase for patients starting in response at 32-52 weeks. Data were available from 4 studies comparing two or three treatments. Figure 52 presents the network of evidence for the sensitivity analysis 2 maintenance phase at 32-52 weeks for patients starting in response.

Figure 52: Sensitivity analysis 2 – network of evidence for the maintenance phase for patients starting in response at 32-52 weeks



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 53 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 2 maintenance phase 32-52 weeks for patients starting in response. Figure 54 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 15.92, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.20 (95% CrI: 0.01, 0.67), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except adalimumab and golimumab 100mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab (-0.36; 95% CrI: -1.29, 0.58). However, none of the treatment effects were statistically significant at a conventional 5% level. Infliximab was most likely to be the most effective treatment (probability of being the best = 0.57).

Figure 53: Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))

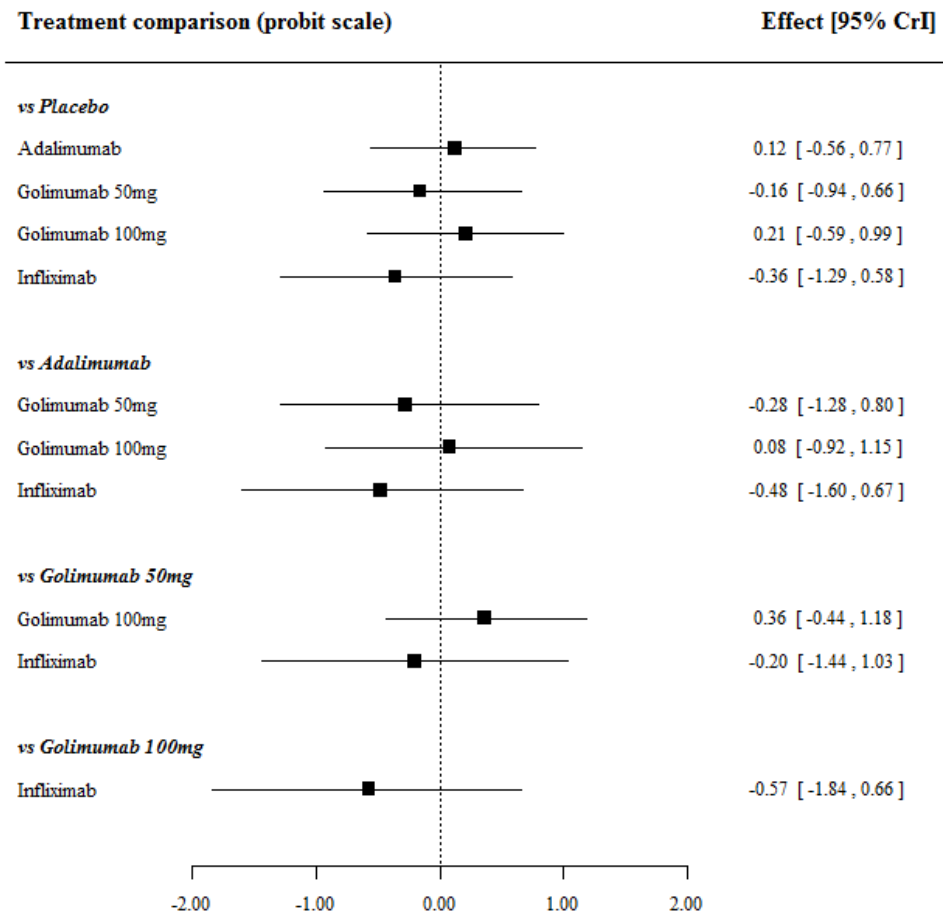
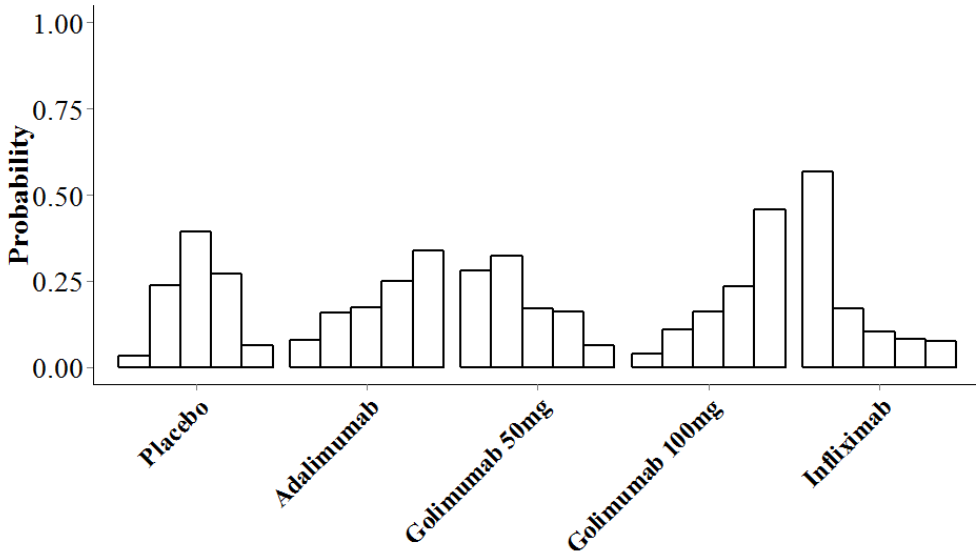


Figure 54: Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 26 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 2 maintenance phase at 32-52 weeks for patients starting in response. Infliximab was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 32-52 weeks. The probabilities of staying in response were comparable among treatments at 32-52 weeks.

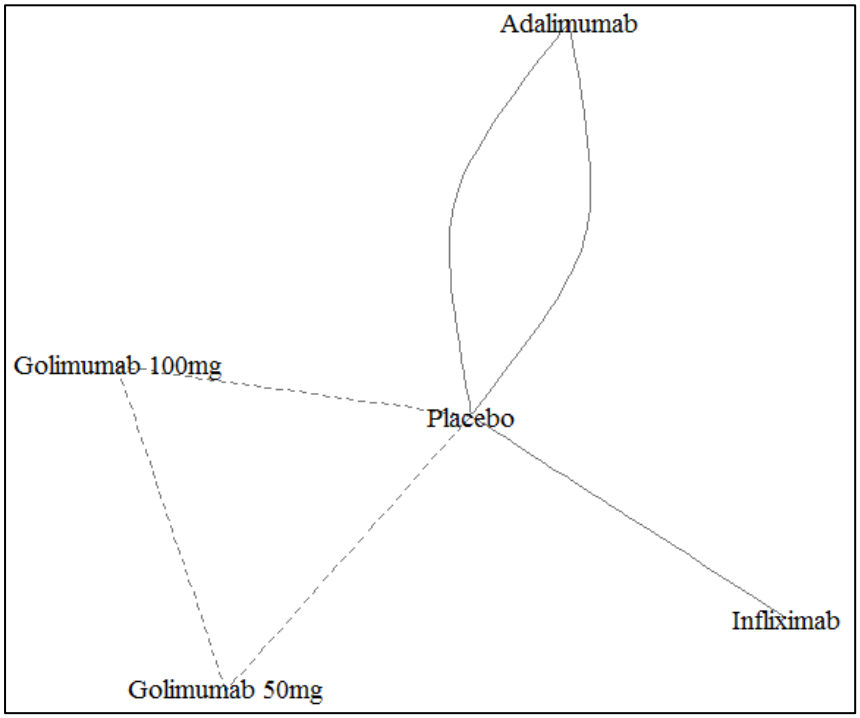
Table 26: Sensitivity analysis 2 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.322	0.318	0.189, 0.472	0.393	0.398	0.200, 0.544	0.286	0.278	0.107, 0.520
Adalimumab	0.371	0.363	0.130, 0.661	0.371	0.378	0.176, 0.524	0.259	0.238	0.05, 0.576
Golimumab 50mg	0.283	0.265	0.066, 0.606	0.370	0.378	0.153, 0.545	0.347	0.332	0.069, 0.713
Golimumab 100mg	0.404	0.395	0.128, 0.730	0.359	0.368	0.153, 0.521	0.237	0.211	0.035, 0.579
Infliximab	0.230	0.202	0.030, 0.575	0.352	0.363	0.115, 0.542	0.418	0.410	0.081, 0.820

a.9.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 32-52 weeks for patients starting in remission. Data were available from 4 studies comparing two or three treatments. Figure 55 presents the network of evidence for the sensitivity analysis 2 maintenance phase at 32-52 weeks for patients starting in remission.

Figure 55: Sensitivity analysis 2 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 56 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 2 maintenance phase at 32-52 weeks for patients starting in remission. Figure 57 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 21.07, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.18 (95% CrI: 0.01, 0.65), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with adalimumab (-0.93; 95% CrI: -1.59, -0.25). However, only the effect of adalimumab was statistically significant at a conventional 5% level. Adalimumab was most likely to be the most effective treatment (probability of being the best = 0.84).

Figure 56: Sensitivity analysis 2 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))

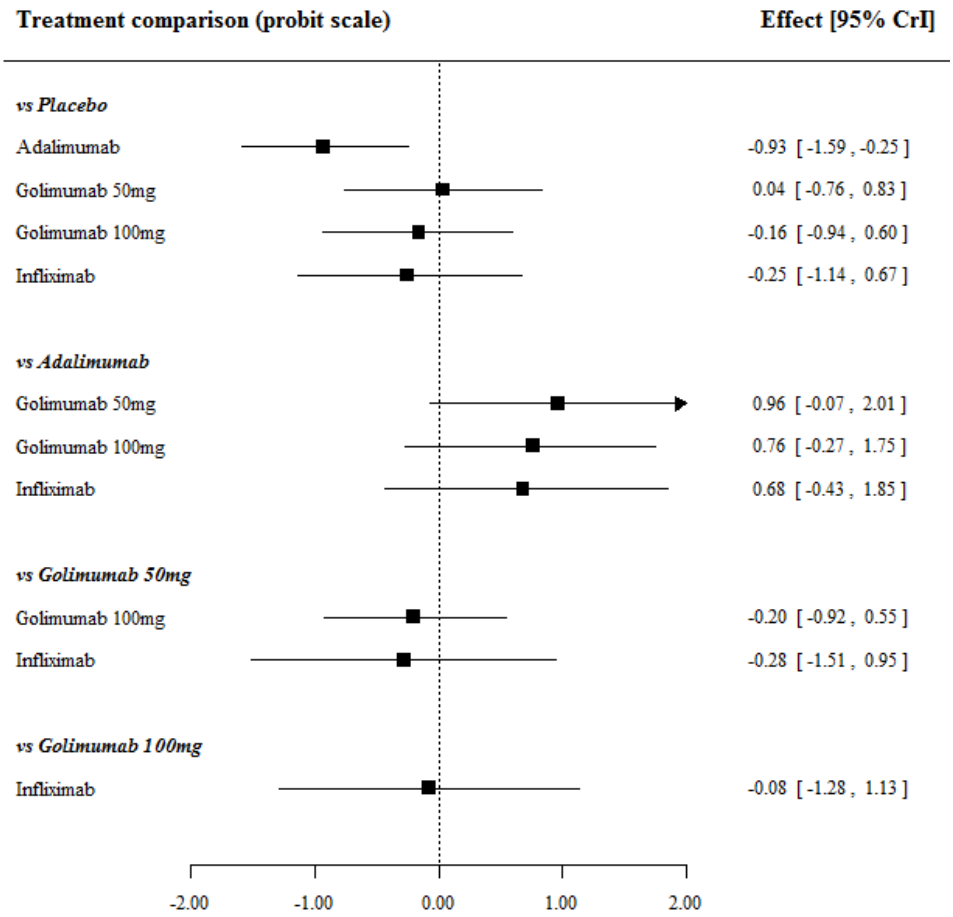
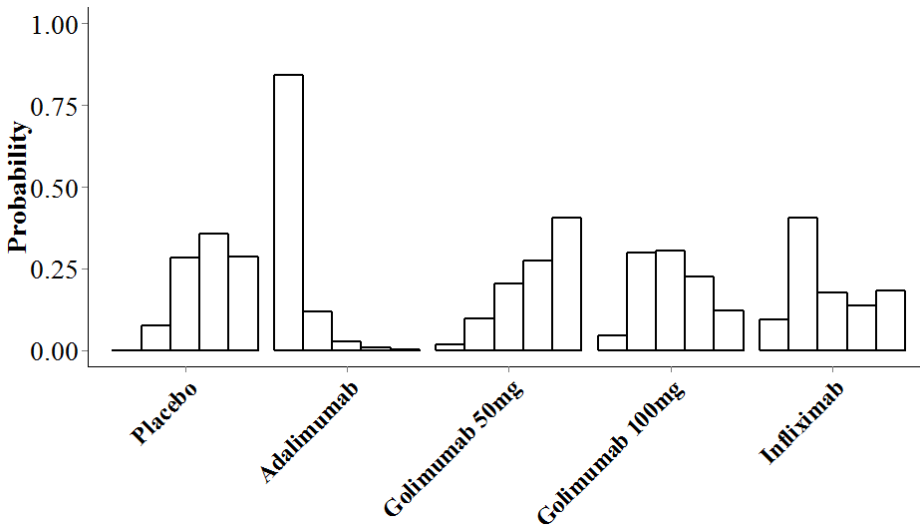


Figure 57: Sensitivity analysis 2 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 27 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 2 maintenance phase at 32-52 weeks starting in remission. Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response or no response at 32-52 weeks.

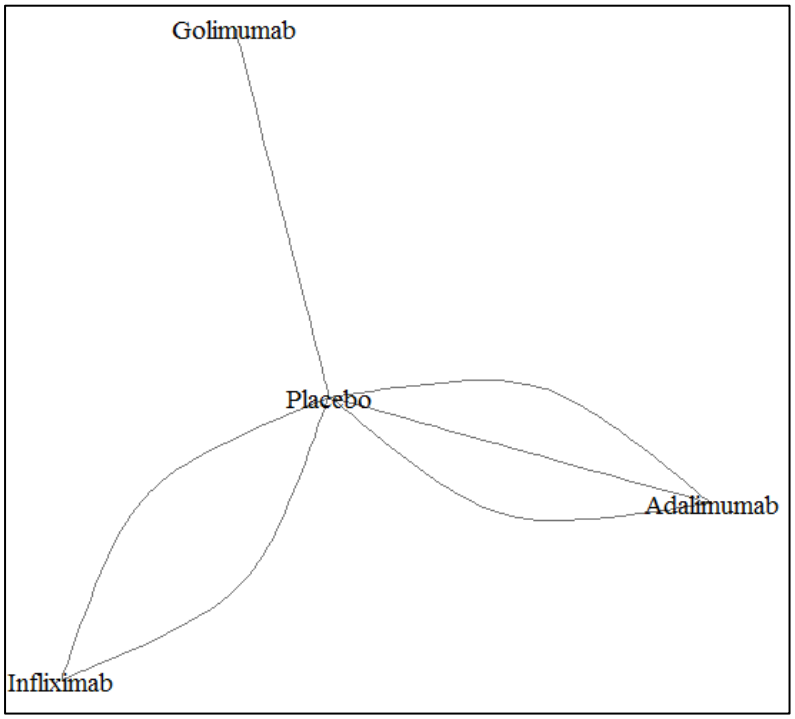
Table 27: Sensitivity Analysis 2 – Probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.296	0.294	0.183, 0.422	0.198	0.184	0.043, 0.445	0.505	0.514	0.233, 0.71
Adalimumab	0.085	0.070	0.013, 0.239	0.111	0.092	0.012, 0.339	0.804	0.830	0.493, 0.963
Golimumab 50mg	0.320	0.307	0.083, 0.633	0.188	0.176	0.035, 0.425	0.492	0.494	0.137, 0.83
Golimumab 100mg	0.258	0.24	0.059, 0.558	0.180	0.165	0.030, 0.425	0.563	0.573	0.186, 0.872
Infliximab	0.239	0.214	0.040, 0.581	0.171	0.157	0.026, 0.416	0.590	0.609	0.176, 0.906

a.10 Sensitivity analysis 3 – induction phase

Sensitivity analysis 3 involved replacing ULTRA2 anti-TNF naïve data with ULTRA2 ITT data and including data from Suzuki *et al*. A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the induction phase. Data were available from 6 studies comparing two treatments. Figure 58 presents the network of evidence for the sensitivity analysis 3 induction phase.

Figure 58: Sensitivity analysis 3 – network of evidence for the induction phase



Note: solid line indicates a 2-arm trial

Figure 59 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 3 induction phase. Figure 60 presents the probabilities of treatment rankings for this analysis. The model fitted the data well, with the total residual deviance, 23.63, being close to the total number of data points included in the analysis, 24. The between-study standard deviation was estimated to be 0.09 (95% CrI: 0.004, 0.38), which implies mild heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab (-0.91; 95% CrI: -1.21, -0.62). All the treatment effects were statistically significant at a conventional 5% level. Infliximab was most likely to be the most effective treatment (probability of being the best = 0.97).

Figure 59: Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the induction phase (SD~HN(0,0.32^2))

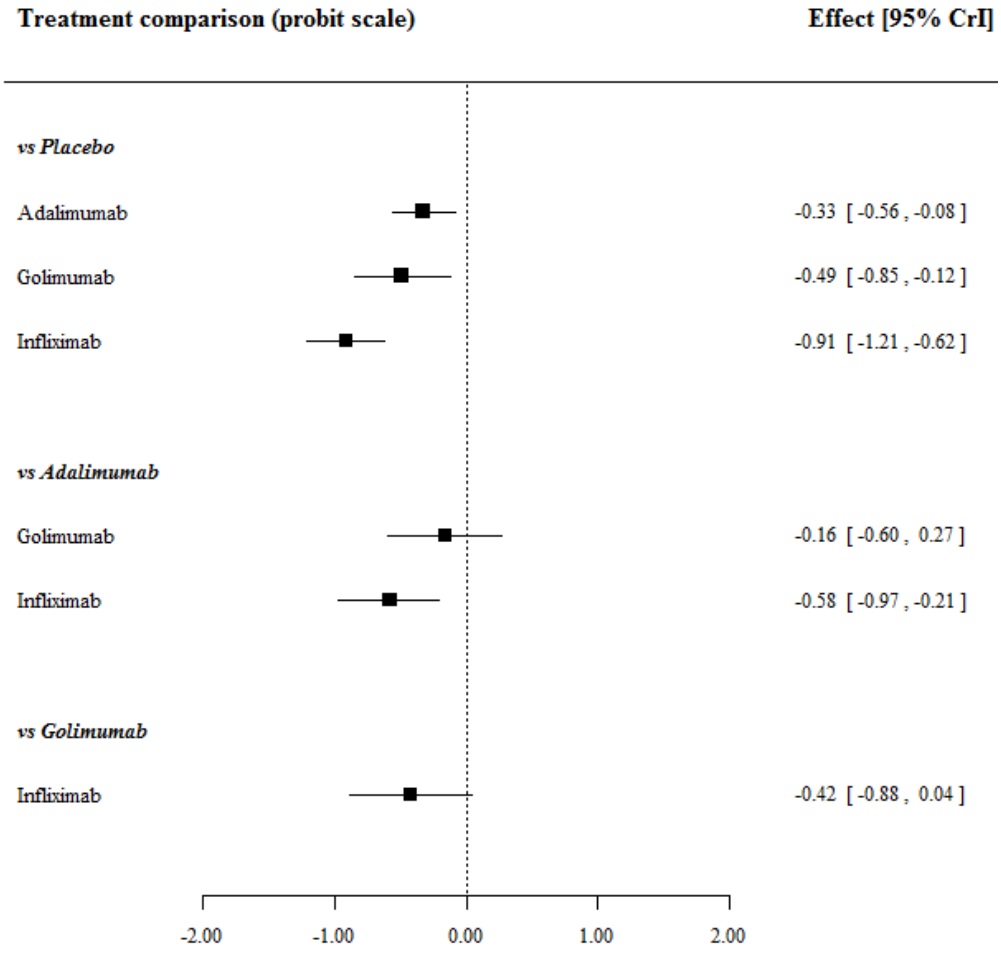
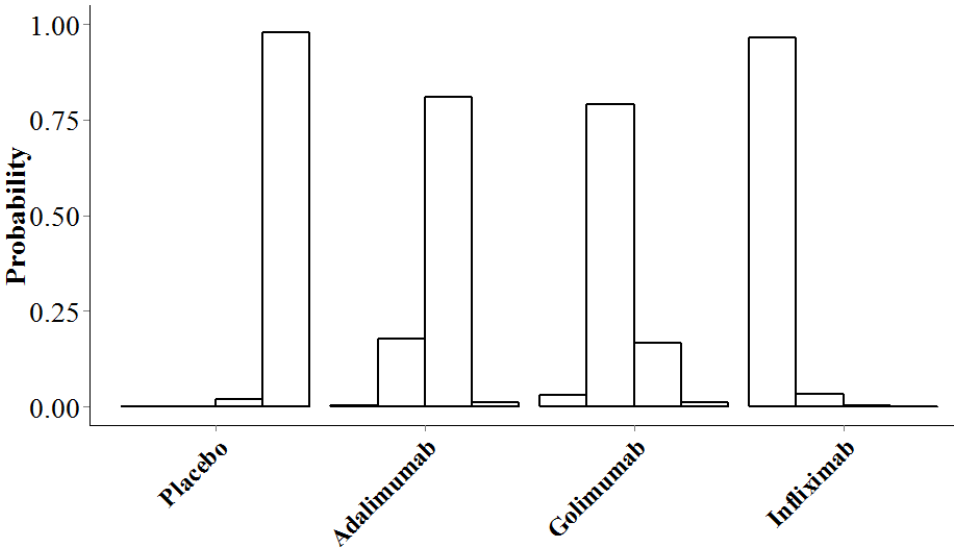


Figure 60: Sensitivity analysis 3 – ranking probability histograms for the induction phase



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 28 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 3 induction phase. Infliximab was associated with the highest probability of moving from no response to response and no response to remission.

Table 28: Sensitivity analysis 3 – probabilities of being in each category for the induction phase

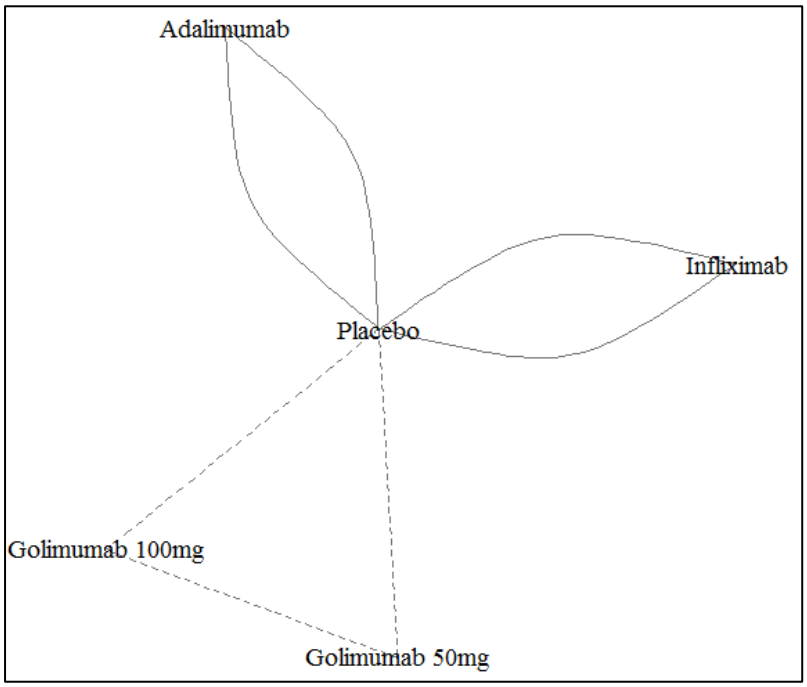
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.649	0.650	0.596, 0.699	0.258	0.258	0.222, 0.294	0.093	0.092	0.064, 0.128
Adalimumab	0.521	0.519	0.414, 0.632	0.317	0.319	0.257, 0.367	0.162	0.160	0.095, 0.241
Golimumab	0.458	0.456	0.315, 0.610	0.336	0.339	0.267, 0.387	0.205	0.201	0.106, 0.331
Infliximab	0.301	0.264	0.196, 0.419	0.358	0.360	0.304, 0.408	0.340	0.338	0.222, 0.477

a.11 Sensitivity analysis 3 – maintenance phase 8-32 weeks

a.11.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase for patients starting in response at 8-32 weeks. Data were available from 5 studies comparing two or three treatments. Figure 61 presents the network of evidence for the sensitivity analysis 3 maintenance phase at 8-32 weeks for patients starting in response.

Figure 61: Sensitivity analysis 3 – network of evidence for the maintenance phase at 8-32 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study.

Figure 62 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 3 maintenance phase at 8-32 weeks for patients starting in response. Figure 63 presents the probabilities of treatment rankings for this analysis. There was some suggestion that model did not represent the data well with the total residual deviance, 14.05, being smaller than would be expected given the total number of data points included in the analysis, 22. The probability of observing a value less than 14.05 is 0.100, which means that this could have occurred by chance. Similar to the base case analysis, all 5 studies had smaller residual deviance than expected. The between-study standard deviation was estimated to be 0.15 (95% CrI: 0.01, 0.55), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 100mg. However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 100mg was most likely to be the most effective treatment (probability of being the best = 0.40).

Figure 62: Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32weeks for patients starting in response (SD~HN(0,0.32^2))

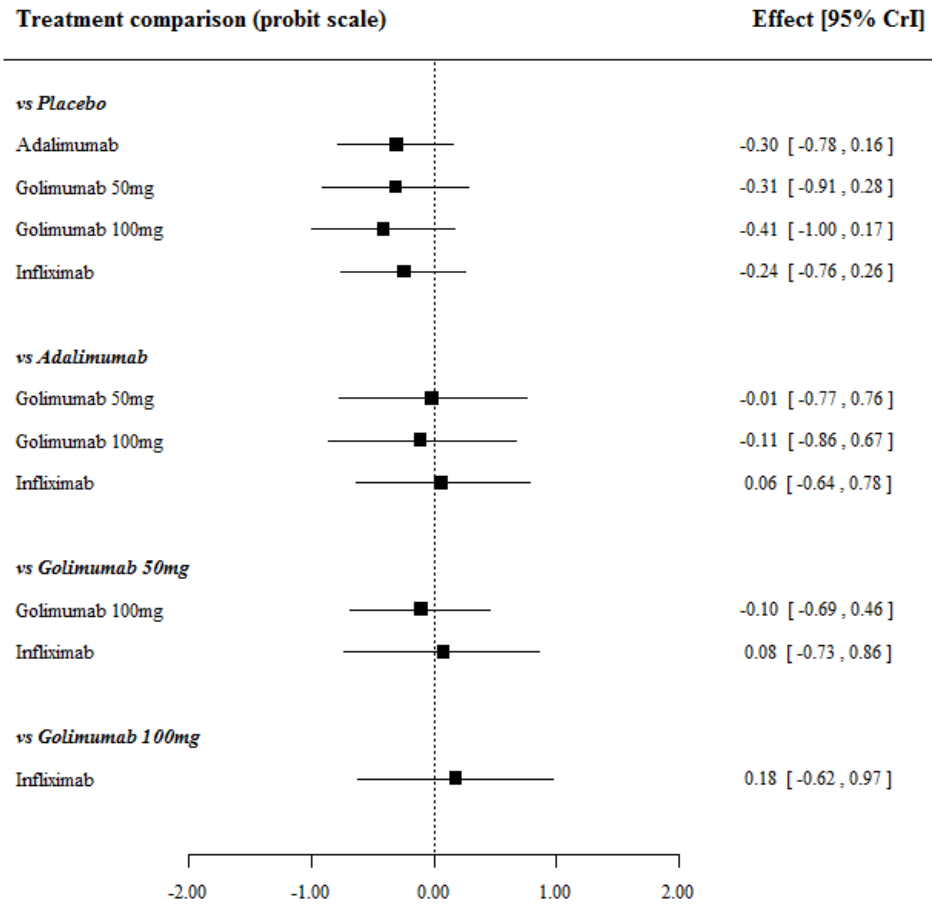
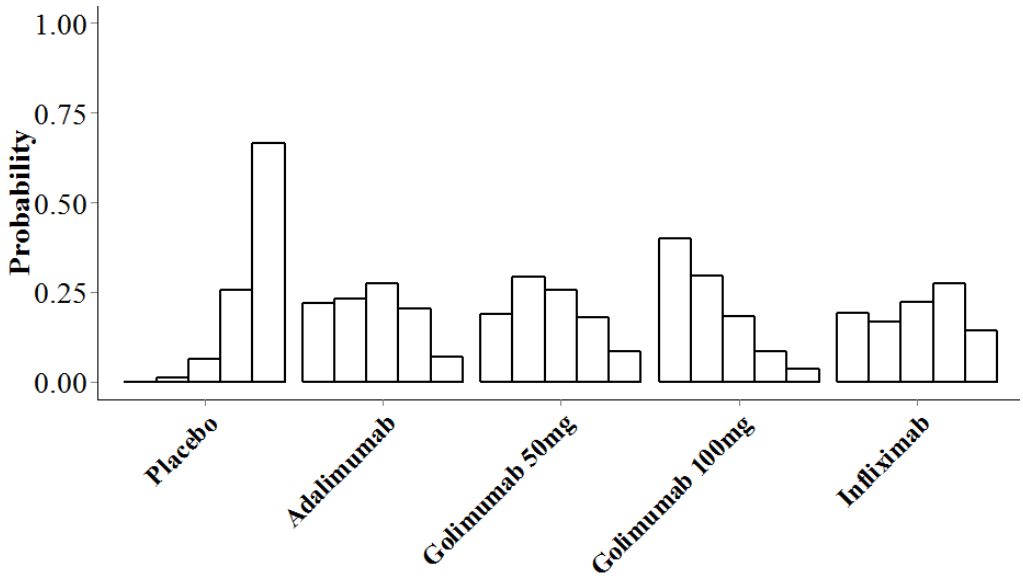


Figure 63: Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 29 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 3 maintenance phase at 8-32 weeks starting in response. Golimumab 100mg was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 8-32 weeks. The probabilities of staying in response were comparable among treatments.

Table 29: Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in response

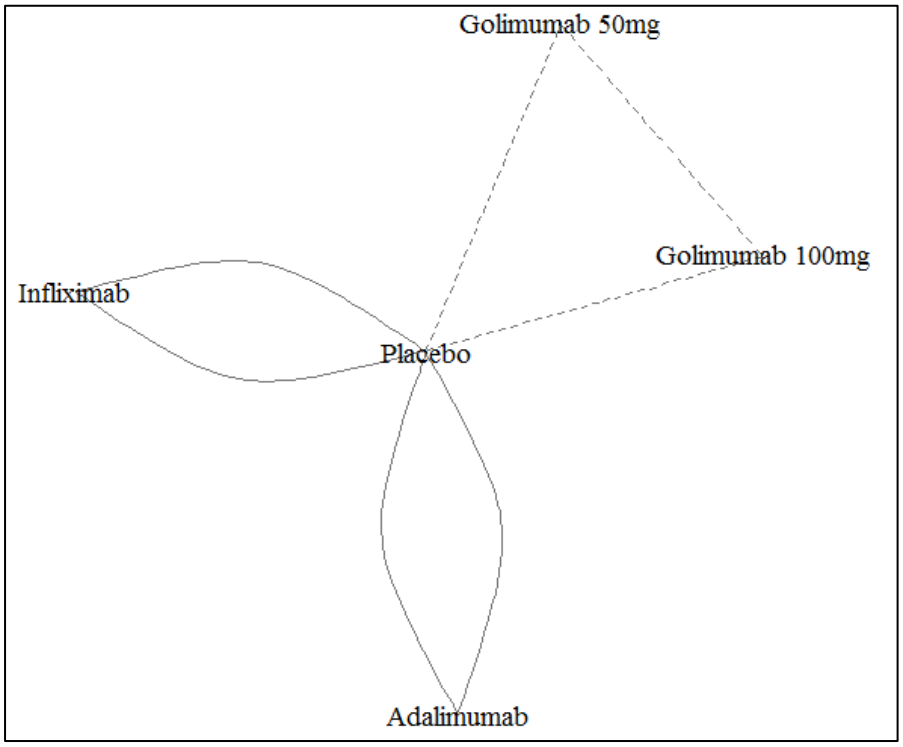
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.545	0.545	0.459, 0.630	0.270	0.270	0.223, 0.32	0.185	0.183	0.120, 0.260
Adalimumab	0.427	0.425	0.243, 0.626	0.292	0.293	0.223, 0.353	0.280	0.274	0.129, 0.470
Golimumab 50mg	0.425	0.422	0.202, 0.669	0.289	0.290	0.204, 0.354	0.287	0.276	0.107, 0.523
Golimumab 100mg	0.386	0.382	0.176, 0.628	0.293	0.294	0.213, 0.356	0.321	0.313	0.129, 0.560
Infliximab	0.451	0.449	0.246, 0.664	0.287	0.289	0.211, 0.349	0.262	0.255	0.109, 0.465

a.11.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 8-32 weeks for patients starting in

remission. Data were available from 5 studies comparing two or three treatments. Figure 64 presents the network of evidence for the sensitivity analysis 3 maintenance phase at 8-32 weeks for patients starting in remission.

Figure 64: Sensitivity analysis 3 – network of evidence for maintenance phase at 8-32 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study

Figure 65 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 3 maintenance phase at 8-32 weeks for patients starting in remission. Figure 66 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 17.77, being close to the total number of data points included in the analysis, 22. The between-study standard deviation was estimated to be 0.16 (95% CrI: 0.01, 0.59), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments were associated with beneficial treatment effects relative to placebo with the greatest effects being associated with golimumab 50mg (-0.62; 95% CrI: -1.33, 0.06) and golimumab 100mg (-0.61; 95% CrI; -1.29, 0.07). However, none of the treatment effects were statistically significant at a conventional 5% level. Golimumab 50mg and golimumab 100mg were most likely to be the most effective treatments (probability of being the best = 0.46 and 0.44, respectively).

Figure 65: Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~HN(0,0.32^2))

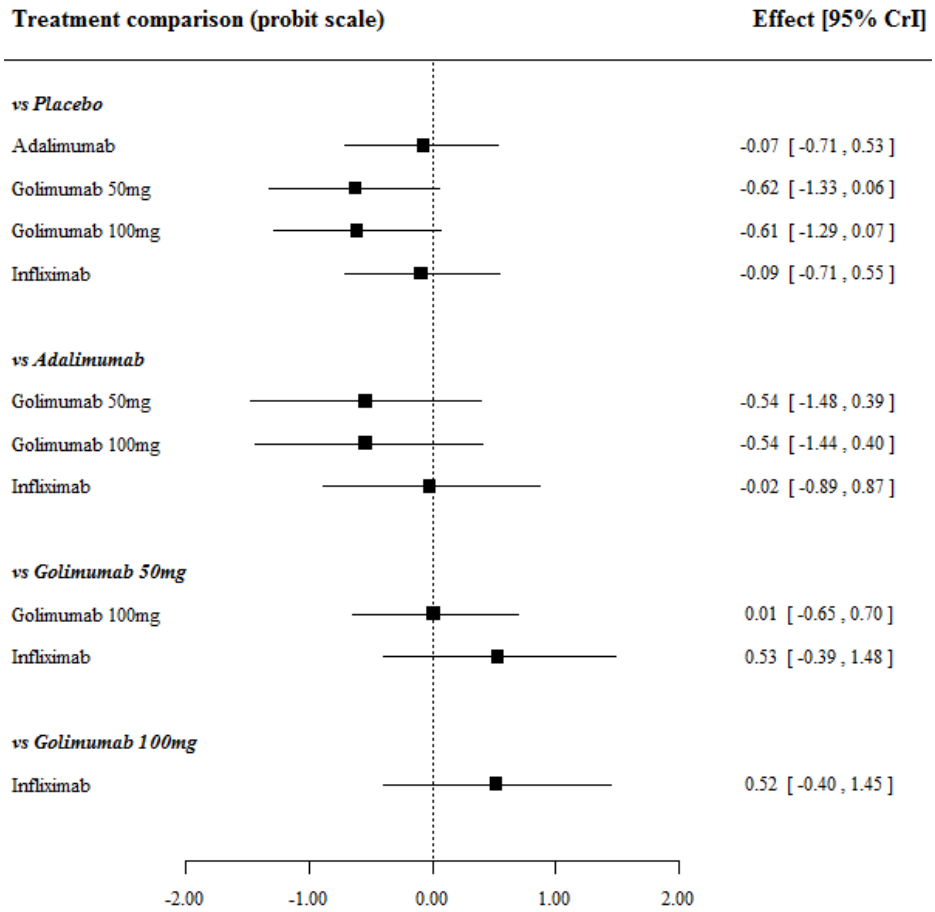
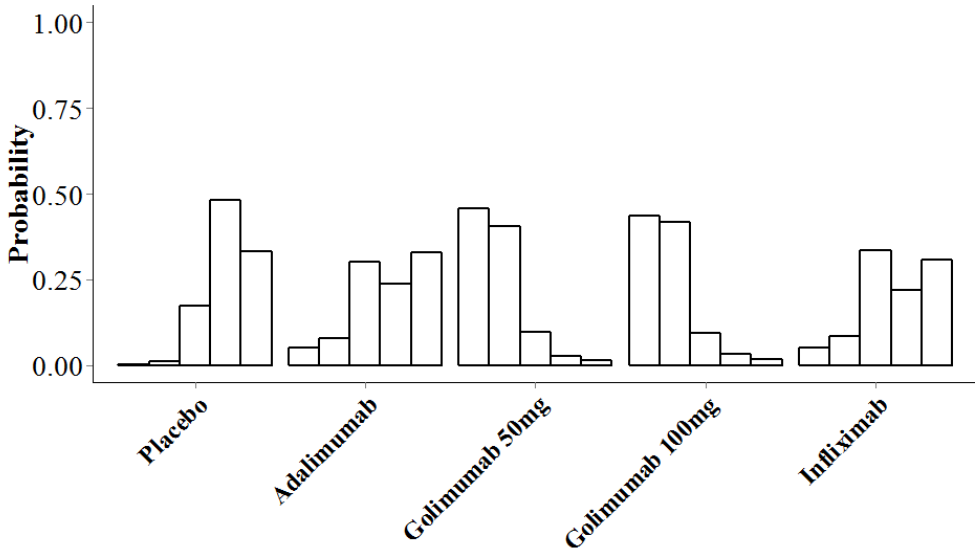


Figure 66: Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 8-32 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 30 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 3 maintenance phase at 8-32 weeks starting in remission. Golimumab 50mg and 100mg were associated with the highest probability of staying in remission, and the smallest probability of moving from remission to no response and from remission to response at 8-32 weeks.

Table 30: Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 8-32 weeks for patients starting in remission

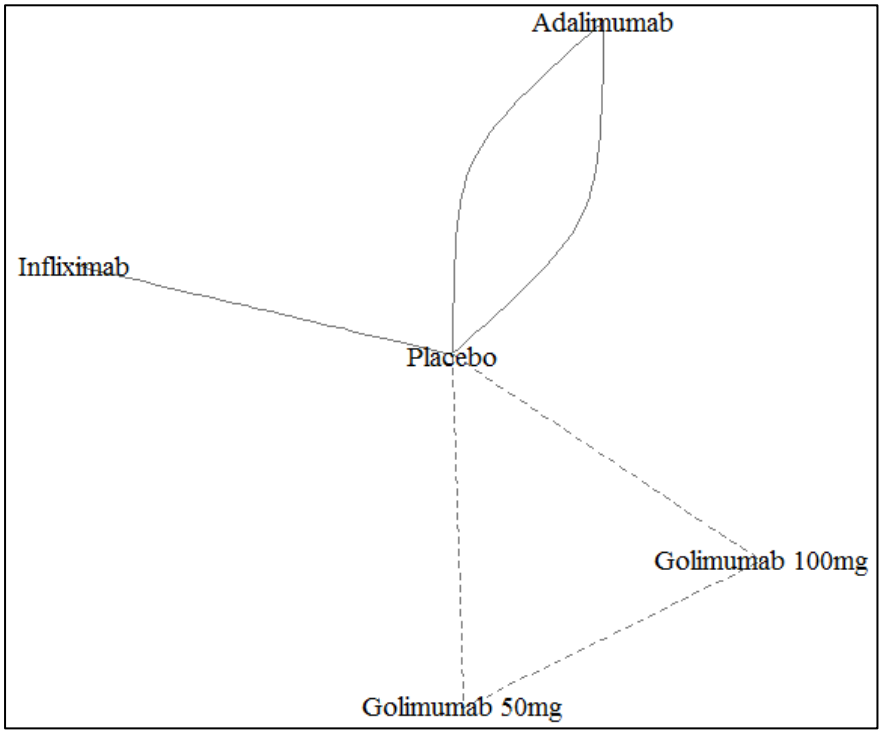
	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.389	0.386	0.240, 0.556	0.185	0.180	0.102, 0.291	0.426	0.425	0.242, 0.613
Adalimumab	0.367	0.358	0.132, 0.649	0.177	0.172	0.086, 0.286	0.456	0.452	0.183, 0.751
Golimumab 50mg	0.199	0.182	0.043, 0.451	0.147	0.144	0.047, 0.264	0.654	0.664	0.342, 0.900
Golimumab 100mg	0.202	0.184	0.046, 0.465	0.148	0.144	0.049, 0.270	0.650	0.663	0.332, 0.890
Infliximab	0.363	0.352	0.130, 0.655	0.176	0.172	0.084, 0.284	0.461	0.461	0.178, 0.753

a.12 Sensitivity analysis 3 – maintenance phase 32-52 weeks

a.12.1 Patients starting in response

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase for patients starting in response at 32-52 weeks. Data were available from 4 studies comparing two or three treatments. Figure 67 presents the network of evidence for the sensitivity analysis 3 maintenance phase at 32-52 weeks for patients starting in response.

Figure 67: Sensitivity analysis 3 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in response



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study

Figure 68 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 3 maintenance phase 32-52 weeks for patients starting in response. Figure 69 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 15.21, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.18 (95% CrI: 0.01, 0.64), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except golimumab 100mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with infliximab (-0.38; 95% CrI: -1.27, 0.55). However, none of the treatment effects were statistically significant at a conventional 5% level. Infliximab was most likely to be the most effective treatment (probability of being the best = 0.55).

Figure 68: Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~HN(0,0.32^2))

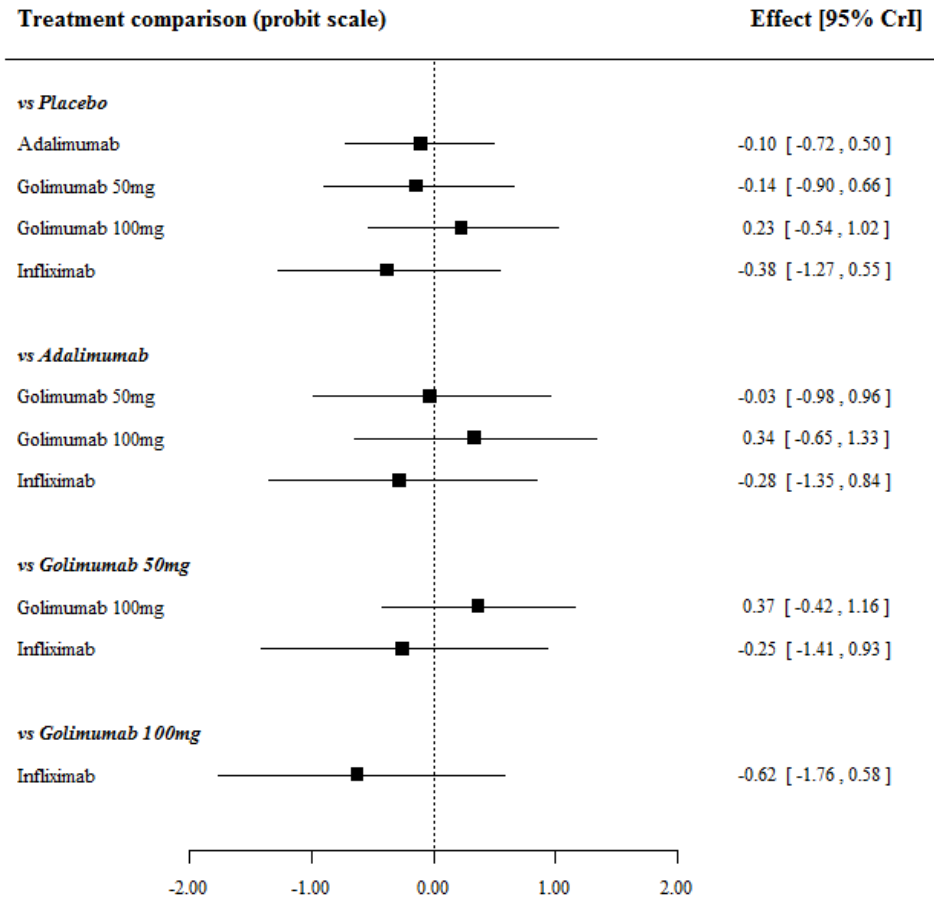
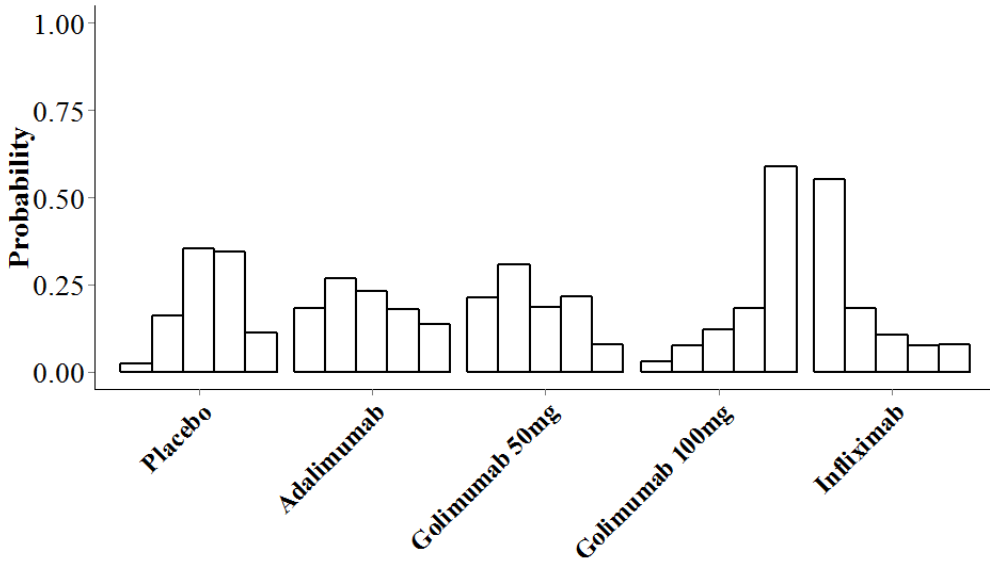


Figure 69: Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in response



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 31 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 3 maintenance phase at 32-52 weeks for patients starting in response. Infliximab was associated with the highest probability of moving from response to remission, and the smallest probability of moving from response to no response at 32-52 weeks. The probabilities of staying in response were comparable among treatments at 32-52 weeks.

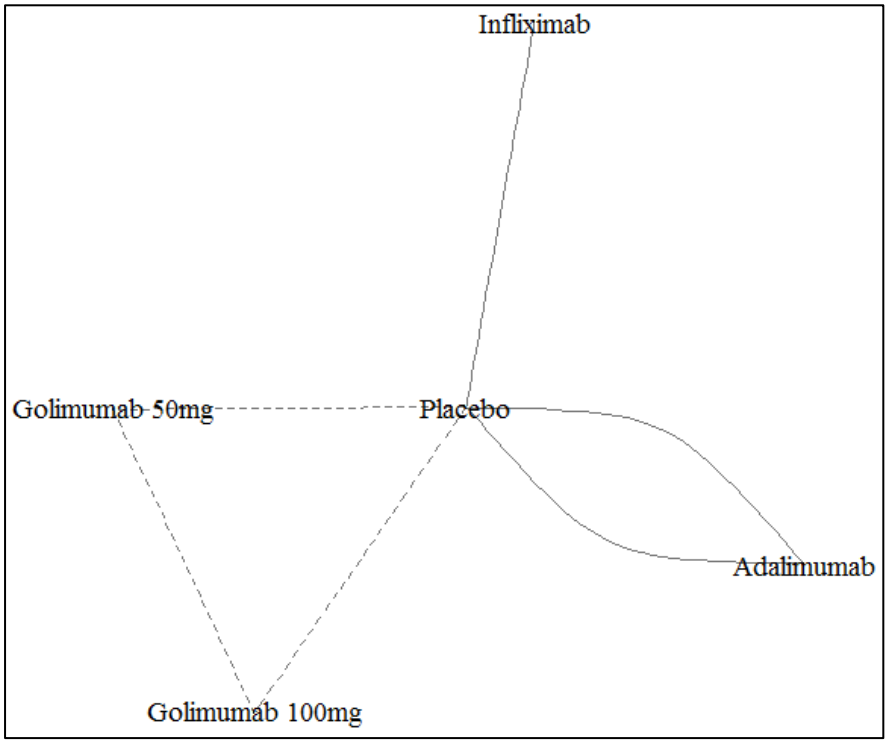
Table 31: Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in response

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.344	0.340	0.218, 0.484	0.393	0.395	0.252, 0.514	0.263	0.257	0.114, 0.449
Adalimumab	0.314	0.304	0.109, 0.580	0.382	0.385	0.219, 0.520	0.305	0.290	0.086, 0.605
Golimumab 50mg	0.309	0.293	0.081, 0.625	0.374	0.380	0.191, 0.514	0.317	0.302	0.067, 0.661
Golimumab 100mg	0.436	0.431	0.149, 0.759	0.354	0.363	0.172, 0.490	0.210	0.187	0.031, 0.520
Infliximab	0.240	0.213	0.041, 0.575	0.358	0.369	0.151, 0.513	0.402	0.392	0.088, 0.771

a.12.2 Patients starting in remission

A network meta-analysis was used to compare the effects of adalimumab, golimumab, and infliximab relative to placebo on clinical response in the maintenance phase at 32-52 weeks for patients starting in remission. Data were available from 4 studies comparing two or three treatments. Figure 70 presents the network of evidence for the Sensitivity Analysis 3 maintenance phase at 32-52 weeks for patients starting in remission.

Figure 70: Sensitivity analysis 3 – network of evidence for the maintenance phase at 32-52 weeks for patients starting in remission



Note: solid line indicates a 2-arm trial; dashed line indicates a 3-arm study

Figure 71 presents the effects of each treatment relative to placebo on the probit scale for the sensitivity analysis 3 maintenance phase at 32-52 weeks for patients starting in remission. Figure 72 presents the probabilities of treatment rankings for this analysis. The model fitted the data reasonably well, with the total residual deviance, 20.55, being close to the total number of data points included in the analysis, 18. The between-study standard deviation was estimated to be 0.18 (95% CrI: 0.01, 0.65), which implies mild to moderate heterogeneity between studies in treatment effects.

All treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with adalimumab (-0.85; 95% CrI: -1.49, -0.16). However, only the effect of adalimumab was statistically significant at a conventional 5% level. Adalimumab was most likely to be the most effective treatment (probability of being the best = 0.80).

Figure 71: Sensitivity analysis 3 – comparative effect of anti-TNF-α treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~HN(0,0.32^2))

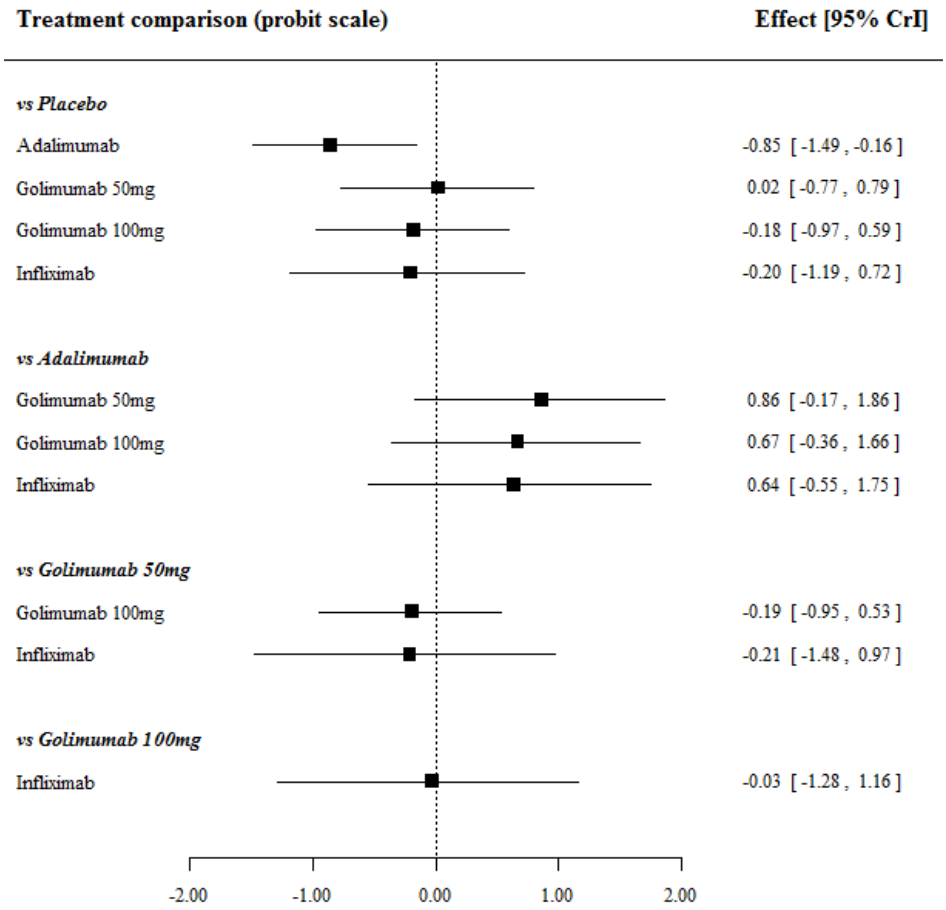
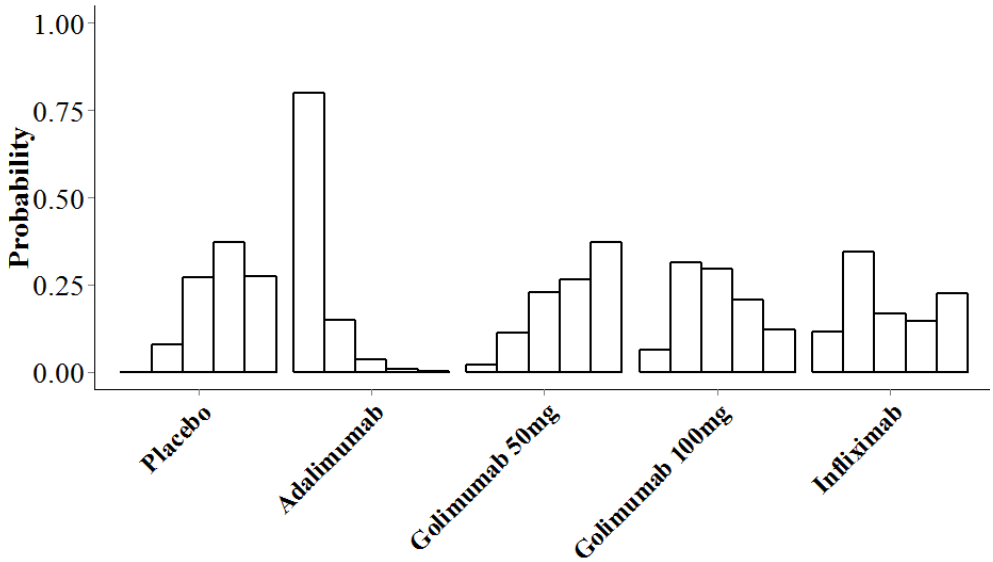


Figure 72: Sensitivity analysis 3 – ranking probability histograms for the maintenance phase at 32-52 weeks for patients starting in remission



Note: the horizontal axis represents the rank of each treatment, i.e., from the best rank (left hand side) to the worst rank (right hand side) within each treatment

Table 32 presents the probabilities of achieving each of the following categories: no response, response and remission for the sensitivity analysis 3 maintenance phase at 32-52 weeks for patients starting in remission. Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response or no response at 32-52 weeks.

Table 32: Sensitivity analysis 3 – probabilities of being in each category for the maintenance phase at 32-52 weeks for patients starting in remission

	No response			Response			Remission		
Treatment	Mean	Median	95% CrI	Mean	Median	95% CrI	Mean	Median	95% CrI
Placebo	0.296	0.293	0.187, 0.422	0.202	0.188	0.042, 0.44	0.502	0.511	0.246, 0.706
Adalimumab	0.097	0.082	0.016, 0.267	0.123	0.104	0.014, 0.343	0.781	0.805	0.464, 0.955
Golimumab 50mg	0.313	0.299	0.083, 0.625	0.191	0.178	0.036, 0.423	0.496	0.501	0.144, 0.821
Golimumab 100mg	0.252	0.235	0.057, 0.544	0.182	0.168	0.03, 0.418	0.566	0.577	0.197, 0.868
Infliximab	0.250	0.229	0.037, 0.588	0.176	0.163	0.025, 0.413	0.574	0.582	0.178, 0.911

5.2.3.6 Biosimilars to infliximab

As defined by the EMA, a biosimilar medicine is "a biological medicine that is developed to be similar to an existing biological medicine (the reference medicine)."[71] In this assessment, the reference medicine is infliximab (Remicade). Two biosimilars to infliximab were also considered within the scope of this assessment: Remsima® (Celltrion) and Inflectra® (Hospira). The EMA has stated that a biosimilar and reference medicine may display differences due to their complex nature and methods of production and that, in the approval process, any differences need to be demonstrated not to affect safety or effectiveness.[71]

A submission[72] was made to NICE for consideration as part of the current assessment by the manufacturers of Remsima (Celltrion Healthcare). However, no sponsor submission was presented by Hospira, the manufacturers of Inflectra. EPAR reports were available for both Remsima[73] and Inflectra.[74]

In June 2013, the EMA Committee for Medicinal Products for Human Use (CHMP) recommended authorisation of Remsima and Inflectra as biosimilars to infliximab, which were reported to be the first authorisation in the European Union for a biosimilar monoclonal antibody. Both Remsima and Inflectra were developed as the product CT-P13.

It was stated in the EPARs for Remsima and Inflectra that an extensive comparability exercise between CT-P13 and Remicade was undertaken, which demonstrated that "*all major physicochemical characteristics and biological activities of Remisima/Inflectra were comparable to those of Remicade*."

The clinical programme to evaluate CT-P13 was based on two main clinical trials:

- a pharmacokinetic equivalence study performed in adult patients with ankylosing spondylitis (Study CT-P13 1.1, PLANET-AS)
- a therapeutic equivalence study of CT-P13 compared with Remicade in adult patients with active rheumatoid arthritis (Study CT-P13 3.1, PLANET-RA)

Both studies were planned with a one-year treatment duration and primary endpoints were evaluated at 30 weeks. Further efficacy and safety data up through 54 weeks were submitted during the EMA assessment.

A third study (CT-P13 1.2) was a small pilot study in RA patients with purpose of facilitating pivotal trial (CT-P13 3.1) conduct.

a) Study CT-P13 1.1, PLANET-AS

PLANET-AS was a prospective Phase I, randomised double-blind multicentre study, in which 250 patients were randomised (CT-P13 N=125, Remicade N=125). Patients received CT-P13 (5mg/kg) or Remicade (5mg/kg) at weeks 0, 2 and 6 and then every 8 weeks to week 54. The primary objective of the study was to demonstrate comparable pharmacokinetics of CT-P13 and Remicade at steady state (between weeks 22 and 30). The primary parameters evaluated were AUC_T and C_{max} after dose 5 (weeks 22-30), with secondary parameters being average concentration at steady state ($C_{av, ss}$) $C_{min\ ss}$, terminal elimination half-life ($T_{1/2}$), total body clearance at steady state (CL_{ss}) and volume of distribution at steady state (V_{ss}). Additional observed parameters were maximum serum concentrations (C_{max}), minimum concentration immediately before next dose (C_{min}) and time to reach C_{max} (T_{max}) after each dose. Efficacy parameters included the proportion of patients achieving clinical response according to the ASAS20 and ASAS40 criteria. The EPARs reported that PLANET-AS demonstrated that (at 5mg/kg) pharmackokinetic behaviour between CT-P13 and Remicade was similar, a view that was also supported by pharmacokinetic data from RA patients in study CT-P13 3.1 (PLANET-RA). Furthemore, the EPARs also stated that the proportions of patients experiencing clinical response according to the ASAS20 and ASAS40 criteria at Weeks 14 and 30 were similar across the CT-P13 and Remicade groups.

b) Study CT-P13 3.1, PLANET-RA

PLANET-RA was a prospective Phase III, randomised, double-blind, multicentre study, in which 302 patients were randomised to CT-P13 and 304 to Remicade (randomisation was stratified by geographical region and baseline CRP level). Patients received CT-P13 or Remicade at 3mg/kg at weeks 0, 2, 6 and then every 8 weeks up to 54 weeks (adminstered in combination with a stable dose of methotrexate and folic acid). The primary objective of the trial was to demonstrate that CT-P13 was equivalent to Remicade up to week 30 in efficacy as measured by ACR20. Secondary objectives were ACR20, ACR50 and ACR70 responses at weeks 14 and 30, DAS28 at week 14 and 30, EULAR response at week 14 and 30, SDAI and CDAI at weeks 14 and 30 and SF-36 at weeks 14 and 30.

Fewer patients randomised to the CT-P13 arm (n=69, 22.8%) discontinued PLANET-RA by week 54 than patients in Remicade arm (n=82, 27.0%). Patients received CT-P13 and Remicade at the RA dose of 3mg/kg. It was stated in the EPARs that a similar proportion of patients at week 30 in the CT-P13 (184/302, 60.9%) and Remicade (178/304, 58.6%) arms achieved ACR20 response (see Table 33).

Table 33: ACR20/50/70 responders at week 30 in PLANET-RA (all randomised population)

Treatment arm	n/N (%)	Estimate of treatment difference	95%CI of treatment difference
ACR20: CT-P13	184/302 (60.9)	0.02	(-0.06, 0.10)
ACR20: Remicade	178/304 (58.6)		
ACR50: CT-P13	105/248 (42.3)	0.02	(-0.07, 0.10)
ACR50: Remicade	102/251 (40.6)		
ACR70: CT-P13	50/248 (20.2)	0.02	(-0.05, 0.09)
ACR70: Remicade	45/251 (17.9)		

Furthermore, at week 30, the findings for the secondary endpoints (including ACR50, ACR70 and decrease in DAS28) were also described as being consistent with the results of the primary endpoint. Efficacy results were reported to be comparable between treatment arms up to week 54. It was concluded in the EPAR that PLANET-RA provided that robust evidence of therapeutic equivalence between CT-P13 and Remicade. ACR responses between CT-P13 and Remicade remained comparable through the 12 month PLANET-RA extension study.[73]

The safety profile of CT-P13 was evaluated in the clinical studies described above. A total of 871 patients were included in the safety population. It was reported in the EPARs that the type and incidence of adverse drug reactions with CT-P13 and Remicade were broadly similar and that no new safety concern was identified. Additionally, it was stated that no marked differences in immunogenicity between CT-P13 and Remicade were observed up to 54 weeks, with comparable effects of antibodies on efficacy and safety. Whilst there was a numerical imbalance described in

serious adverse events observed in study CT-P13 3.1 (PLANET-RA) (with a higher number of serious infections), reported numbers were stated to be low and therefore the CHMP concluded that this observed difference was likely to be due to chance.

In summary, the EMA considered CT-P13 to be biosimilar to the reference product Remicade and judged that the submitted data in the submissions for Remsima and Inflectra allowed for extrapolation to all other indications of Remicade.

An ECCO position statement was presented by Danese and Gordon[75] stating that the use of biosimilars in patients with IBD requires clinical trials in the IBD patient population to allow comparison between the biosimilar and reference products, on the basis of potential differences in manufacturing and structure that could lead to important differences in immunogenicity and efficacy. However, a subsequent statement issued on behalf of the Working Party on Similar Biological (Biosimilar) Medicinal Products of the CHMP argued that no pharmacokinetic and safety issues are known to be particular to IBD, the most responsive known population (RA) was assessed for immunogenicity and that the data submitted allow extrapolation to patients with IBD. [76]

5.3 Discussion

A total of ten RCTs were identified in the clinical effectiveness systematic review, of which nine related to adults and one was conducted in a paediatric population. All of the adult RCTs were performed against placebo (with the exception of UC-SUCCESS) and were a maximum of one year in study duration. No head-to-head RCTs were identified in which interventions of interest were assessed against each other.

The risks of bias associated with included RCTs were assessed using the Cochrane risk of bias instrument. Only three RCTs could be considered as being at overall low risk of bias as allocation concealment, blinded outcome assessment and completeness of outcome data were all judged as low risk. It should be noted that one of the maintenance trials (PURSUIT-Maintenance) re-randomised patients who had previously responded to golimumab induction therapy in two previous trials; the extent of this potential bias on patient outcomes is unclear.

The outcome measures pre-specified in the final NICE scope were all addressed by the included trial evidence, with the exception of rates of relapse. Clinical response and remission data based on complete Mayo scores were well reported across trials. Evidence was identified to demonstrate that patients receiving infliximab, adalimumab or golimumab were more likely than patients receiving placebo to achieve clinical response and remission at induction and maintenance time points. Patients in the UC-SUCCESS trial who received combination treatment with infliximab and azathioprine

experienced the most favourable rates of steroid-free remission compared with infliximab and azathioprine treatment groups. Seven RCTs performed in adult populations contributed data on clinical response and remission at induction or maintenance time points to network meta-analyses.

Based on the NMA, in the induction phase, all treatments were associated with statistically significant beneficial effects relative to placebo with the greatest effect being associated with infliximab. Infliximab was also associated with the highest probability of moving from no response to response and from no response to remission. The effects of adalimumab and golimumab on these two probabilities were broadly comparable.

For patients classified as responders at the end of the induction phase, treatment effects were not statistically significant, although the greatest effect was associated with golimumab 100mg at 8-32 weeks. Golimumab 100mg was associated with the highest probability of moving from response to remission and staying in response, and the smallest probability of moving from response to no response. However, at 32-52 weeks, only infliximab and golimumab 50mg were associated with beneficial effects on clinical response, although the effects were not statistically significant. Infliximab was associated with the highest probability of moving from response to remission and staying in response, and the smallest probability of moving from response to no response at 32-52 weeks. The probabilities of staying in response were comparable among treatments at both 8-32 weeks and 32-52 weeks.

For patients classified as being in remission at the end of the induction phase, all treatments except for adalimumab were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 50mg and 100mg, although the effects were not statistically significant at 8-32 weeks. Golimumab 50mg and 100mg were associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response and from remission to no response. At 32-52 weeks, all treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo, with the greatest effect being associated with adalimumab (the only treatment with statistically significant effect). Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response and from remission to no response.

Sensitivity analyses were conducted by replacing ULTRA2 anti-TNF-naïve data with ULTRA2 ITT data (sensitivity analysis 1), including Suzuki *et al* (sensitivity analysis 2), and replacing ULTRA2 anti-TNF-α naïve data with ULTRA2 ITT data plus including Suzuki (sensitivity analysis 3). The results suggested that when ULTRA2 ITT data were replaced by ULTRA2 anti-TNF-α naïve data for patients starting in remission at 8-32 weeks and in response at 32-52 weeks, the estimate of the effect

of adalimumab on clinical response changed from being slightly worse than placebo to being slightly better than placebo. However, the estimates were associated with considerable uncertainty.

Available data on hospitalisation outcomes were very limited, but suggested that outcomes may be more favourable for adalimumab-treated and infliximab-treated patients compared with placebo (with no data available from golimumab trials). Data on surgical intervention were also very sparse, with a potential inconclusive benefit for intervention-treated patients compared with placebo. No trials reported whether surgical outcomes were elective or emergency in nature. However, more data are required to demonstrate the impact of interventions on hospitalisation and surgical intervention more conclusively. Data were available from a single trial to support the use of infliximab in induction and maintenance treatment in a paediatric population.

The main safety issues highlighted in the RCT evidence appeared to be generally consistent with those previously discussed in the respective SmPCs (including serious infections, malignancies and administration site reactions). Deaths occurring during and after the study period were described in some trials evaluating golimumab (PURSUIT-Maintenance) and infliximab (ACT trials), of which infection or malignancy commonly appeared to be implicated. This underlines the importance of monitoring and treating serious infections and malignancies in patients receiving immunosuppressive treatment.

The trials included in the clinical effectiveness systematic review typically excluded patients with ulcerative proctitis, patients with fulminant/acute severe disease, those with a history of or at imminent risk of bowel surgery, pregnant or lactating women, and patients with diseases of the central nervous system (e.g. demyelinating disease). Furthermore, patients with history of serious infection and immunodeficiency were also typically excluded, as were individuals with a history of malignancy or signs of dysplasia. Therefore, the effects of adalimumab, golimumab or infliximab in these UC populations are unknown.

Two biosimilars (Remsima and Inflectra) to Remicade were considered as part of the evidence base for infliximab as part of this assessment. The sponsor submission received from the manufacturers of Remsima (Celltrion) and the EPAR reports for Remsima and Inflectra indicated that both biosimilars were approved by the EMA on the basis of reported similar pharmacokinetic and efficacy (demonstrated in AS and RA patients) profiles to Remicade. No further trials of Remsima or Inflectra were identified in the course of this assessment.

6. ASSESSMENT OF COST-EFFECTIVENESS

6.1 Systematic review of existing cost-effectiveness evidence

6.1.1 Methods

6.1.1.1 Identification of studies

A comprehensive search was undertaken to systematically identify literature relating to the cost effectiveness of infliximab, adalimumab and golimumab for treating moderate-to-severe ulcerative colitis after the failure of conventional therapy. The search strategy comprised the following main elements:

- Searching of electronic databases
- Contact with experts in the field
- Handsearching of bibliographies of retrieved papers.

The following electronic databases were searched from inception for economic evaluations:

- MEDLINE, MEDLINE in-Process and Other Non-Indexed Citations: Ovid. 1946-present
- EMBASE: Ovid. 1974-2013
- Cochrane Library: Wiley Interscience
    - Cochrane Database of Systematic Reviews (CDSR). 1996- present
    - Database of Abstracts of Reviews of Effects (DARE). 1995- present
    - Cochrane Central Register of Controlled Trials (CCRT). 1995- present
    - Cochrane Methodology Register. 1904- January 2014
    - Health Technology Assessment Database (HTA). 1995- present
- NHS Economic Evaluation Database (NHS EED). 1995- present
- CINAHL: EBSCO. 1982- present
- Web of Science Citation Index: Web of Knowledge. 1900- present
- Conference Proceedings Citation Index: Web of Knowledge. 1990- present
- BIOSIS Previews: Web of Knowledge. 1969- present
- EconLit: Ovid. 1886-present

The MEDLINE search strategy is presented in Appendix 10. The search strategy combined freetext and MeSH (Medical subject headings) or thesaurus terms relating to *ulcerative colitis*, with freetext and MeSH or thesaurus terms relating to *infliximab, adalimumab and golimumab* combined with highly sensitive economic filters to retrieve economic evaluations. The search strategy was translated across all databases. No date or language restrictions were applied. Literature searches were conducted during January 2014. References were collected in a bibliographic management database and duplicates were removed.

6.1.1.2 Inclusion / exclusion criteria

Studies were included in the systematic review if they reported full economic evaluations comparing infliximab, adalimumab and/or golimumab, against each other or against any other intervention, within their licensed indications for the treatment of patients with moderate to severe UC. The inclusion and exclusion criteria applied within the systematic review are presented in Box 1. Studies were included only if they were reported as full papers; conference abstracts were excluded from the review as they present insufficient detail to allow for a rigorous assessment of study quality.

Box 1: Inclusion and exclusion criteria for review of cost-effectiveness studies

Inclusion criteria

- Full economic evaluations comparing infliximab, adalimumab and/or golimumab against each other or any other intervention for the treatment of patients with moderate to severe UC

Exclusion criteria

- Studies assessing biologics in the acute setting (e.g. management of UC exacerbations)
- Studies in which the same biologic is used in all treatment groups within the analysis
- Non-comparative studies and partial economic evaluations (e.g. costing studies)
- Abstracts, letters and commentaries
- Studies not reported in English
- Studies relating to patients with diseases other than UC

6.1.1.3 Review methods

The results of the economic searches were sifted by title and abstract. The full papers of studies which potentially met the inclusion criteria were retrieved for further inspection. Studies included in the systematic review were critically appraised using the Drummond checklist for economic evaluations.[77] In addition, the manufacturers of the products considered within this appraisal submitted economic evidence to NICE; these models were assessed against the NICE Reference Case.[78] The structure and formulae included in the manufacturer's submission models were scrutinised by two members of the Assessment Group (PT and HB). It should be noted that this appraisal includes an update of Technology Appraisal Guidance 140;[79] the economic evaluation reported within the 2007 Schering Plough submission to NICE[80] is not included in this review as it has previously been critiqued for NICE,[81] however one of the studies included in the review[82] reports an analysis of this model.

6.1.1.4 Results

The systematic searches identified a total of 907 potentially relevant citations (see Table 34 and Figure 73). In addition, 4 manufacturer's submissions were received by NICE.[63,65,67,72] Two of the four submissions were submitted by the same manufacturer – one relating to golimumab and one relating

to infliximab; as these relate to virtually identical models, they are considered as a single analysis within this assessment. Three of the included submissions to NICE[63,65,67] included economic analyses; the submission from Celltrion[72] did not include any economic analysis. Fourteen studies were excluded as they were available only in abstract form. A total of three published studies and three manufacturer's submissions reported economic analyses relating to the use of biologics for the treatment of moderate to severe UC (see Table 35).

Table 34: Summary of search results for existing economic evaluations

Database	Date range	Date searched	Number of results
Medline (Ovid)	1946-Present	15/01/14	96
Embase (Ovid)	1974-Present	15/01/14	372
CINAHL (EBSCO)	1982-Present	22/01/14	23
SCI & SSCI (WOK)	1900-Present	22/01/14	243
BIOSIS (WOK)	1969-Present	22/01/14	186
Cochrane HTA (Wiley)	1991-Present	21/01/14	30
Cochrane DARE (Wiley)	1991-Present	21/01/14	28
Cochrane EED (Wiley)	1991-Present	21/01/14	24
EconLit (Ovid)	1886-Present	15/01/14	1

Figure 73: Study selection results for review of economic evaluations

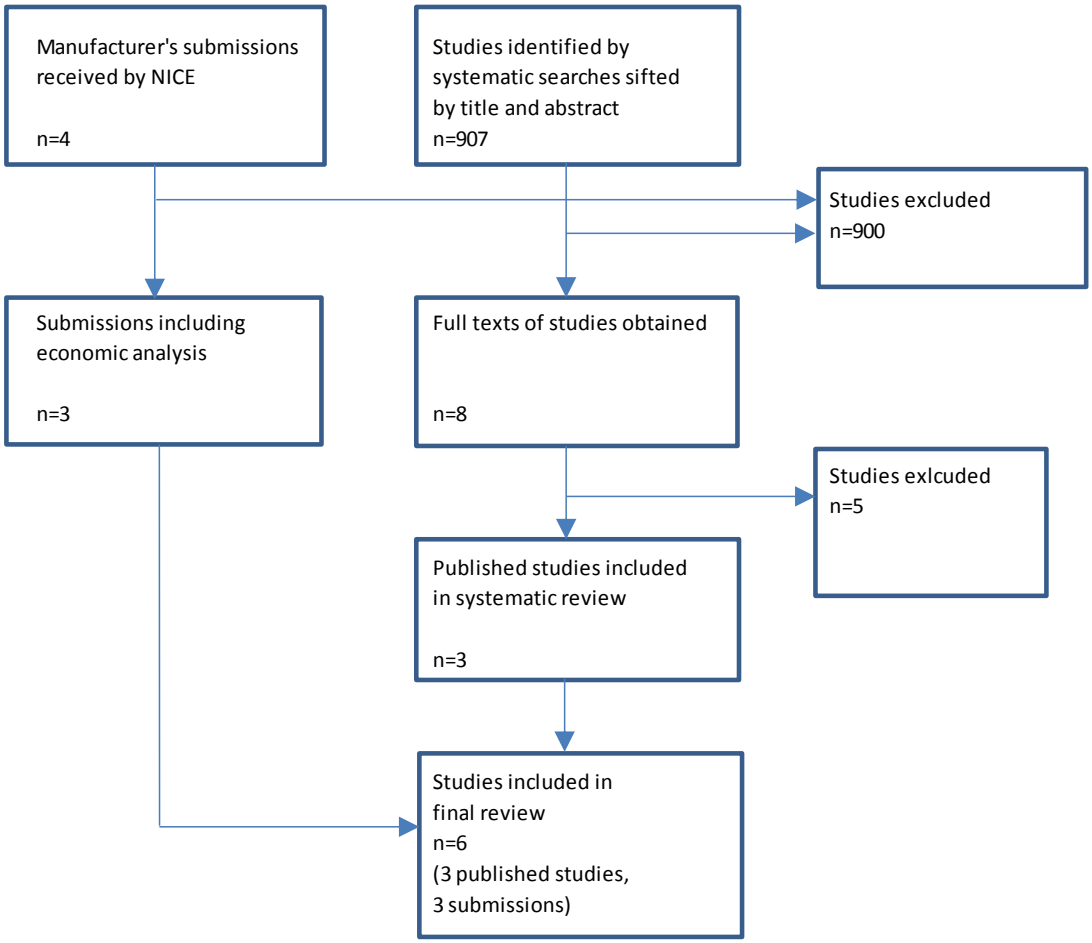


Table 35: Summary table of included published studies

Study	Year of publication	Perspective	Economic comparisons	Outcome measure	Time horizon	Conflicts of interest
Included published economic evaluations						
Park *et al*[83]	2012	US (payer)	Colectomy+IPAA vs standard medical care (including infliximab)	QALYs	Lifetime	Non-commercial
Tsai *et al*[82]	2008	UK NHS	Infliximab vs standard care	QALYs	10-years	Schering-Plough Ltd
Xie *et al*[84]	2009	Canadian (public payer)	Infliximab vs usual care	QALYs	5-years	3 of 6 authors disclosed a conflict of interest
Included manufacturer's submissions						
AbbVie submission[63]	n/a	UK NHS	Adalimumab versus conventional non-biologic treatment	QALYs	10-years	Manufacturer of adalimumab
MSD submission[65,67]	n/a	UK NHS	Pairwise comparisons of infliximab, golimumab, adalimumab and immediate colectomy	QALYs	10-years	Manufacturer of infliximab and golimumab

Section 6.1.1.5 presents a summary and critical appraisal of the three published economic studies included in this review.[82-84] Sections 6.1.1.6 and 6.1.1.7 present critical reviews of the individual manufacturer's submissions from AbbVie and MSD respectively.[63,65,67]

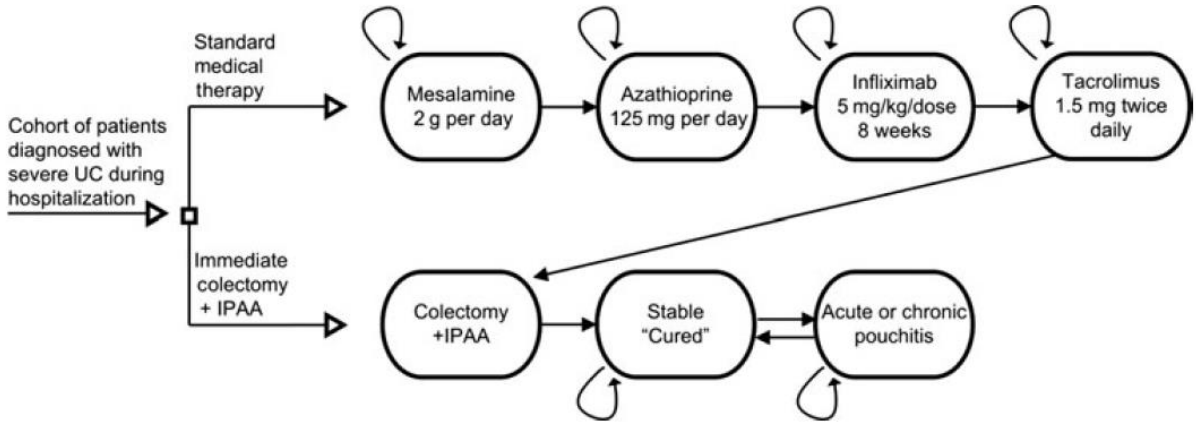
6.1.1.5 Review of published economic evaluations

Park et al - Cost-effectiveness of early colectomy with ileal pouch-anal anastamosis versus standard medical therapy in severe ulcerative colitis[83]

Park *et al*[83] report the methods and results of an economic analysis of early colectomy plus ileal pouch anal anastomosis (IPAA) versus standard medical therapy in patients with severe UC in the US. The model population is intended to reflect 21-year old patients with newly diagnosed pancolitis UC confirmed by colonoscopic biopsies. The economic analysis compares two sequences of treatments: (1) immediate colectomy with IPAA, and; (2) standard medical therapy which is assumed to be comprised of a sequence of (i) mesalamine 2g per day (ii) azathioprine 125mg per day (iii) infliximab 5mg/kg/dose 8 weeks (iv) tacrolimus 1.5mg b.i.d (v) colectomy+IPAA (see Figure 74). The analysis does not consider the comparative cost-effectiveness of alternative sequences of medical treatments. The authors purport to have adopted a societal perspective, however it does not appear that any

indirect costs borne outside of the health sector (e.g. lost productivity or out-of-pocket expenses) have been included in the economic analysis. It would be more accurate to describe the adopted perspective as that of the health care payer. Health economic results are presented in terms of the incremental cost per quality adjusted life year (QALY) gained over a lifetime horizon. In line with the Reference Case set out in Gold *et al*,[85] costs and health outcomes were discounted at a rate of 3% per year. Costs were valued at 2009 prices.

Figure 74: Model diagram presented by Park *et al*[83]



The economic analysis uses a Markov approach to evaluate relevant events, costs and health outcomes. The duration of each Markov cycle is not entirely clear from the paper; the text indicates that the first cycle is 3-months in duration and the cycle length appears to be 8-weekly thereafter, however the table of parameter values presented within the paper suggests that probability parameters are defined according to various time intervals. In addition, the text does not mention whether a half-cycle correction has been applied or not. The precise health states adopted in the model are also not entirely clear from the text; whilst a model diagram is presented in the paper (see Figure 74), this details the sequences of treatments in each group but does not specify the relevant clinical events that patients may experience. It appears that separate states are assigned for patients who are on treatment, in remission, experiencing UC flare, post-colectomy and experiencing death. With respect to medical treatment, the model appears to separate response and remission based on the Simple Colitis Activity Index (SCAI) score.[86] The model also includes the possibility of patients developing colorectal cancer; this is assumed to result in colectomy. It appears that the model does not include an excess risk of death due to colorectal cancer. Patients enter the model at the point of being hospitalised during their initial flare definitively diagnosing them of pancolitis UC through endoscopic biopsies. Whilst this may indicate a more severe population than that stated within the scope of this appraisal, the model uses RCT evidence from studies[50] which relate to a moderate to severe UC population *and* assumes that the flare is resolvable without surgery. Following diagnosis, patients are then assumed to receive intravenous methylprednisolone and subsequently mesalamine as maintenance therapy once

they are able to tolerate oral medicines. Patients progress along the treatment pathway to infliximab, tacrolimus and potentially colectomy+IPAA if remission is not achieved. Patients in the intervention group within the model bypass all medical treatments and immediately undergo surgery. Different cost and health-related quality of life (HRQoL) estimates are applied to each health state.

Treatment benefits are defined differently for surgery and medical treatment. For the standard medical therapy group, treatment benefits are characterised as response, remission and UC flare rates. For the surgery group, treatment benefits are determined according to colectomy success rates, the avoidance of adverse events (pouchitis and infertility), the requirement for antibiotics, and remission rates for antibiotics. Effectiveness data were drawn from a number of sources including RCTs and non-randomised studies.[50,87-99] The effectiveness of infliximab in inducing and maintaining response and remission was based on the results of the ACT1/2 studies.[50] The rate of developing colorectal cancer was derived from an observational study.[100] The approach for determining the effectiveness of medical and surgical treatment options is essentially a naïve indirect comparison; as such the results of the analysis should be interpreted with a degree of caution.

Health utilities were assigned for the following events and states: UC flare (0.48), remission (0.91), post-colectomy (0.87), and infertility following IPAA (0.74). Health utilities were drawn from a variety of sources.[101-104] The elicitation methods from which these estimates were derived were not consistently clear; whilst several studies used the Time Trade Off (TTO) method, the study reportedly used to derive a disutility for female infertility is not a quality of life valuation study and actually relates to the medical costs of epididymitis and orchitis in men;[104] it is unclear how the information contained within this paper has been used to inform the economic analysis.

The model includes resource costs associated with diagnosis of pancolitis UC, drug therapy, colectomy+IPAA, managing pouchitis and the diagnosis of colorectal cancer. The costs associated with hospitalisations, outpatient visits, procedures, and laboratory costs were estimated using national reimbursements from the Centers for Medicare and Medicaid Services and average reimbursement rates from all patient billing records in 2009 at Stanford University Medical Center. The Office of Statewide Health Planning and Development tables were used to validate institutional rates with the intention of reflecting national average cost estimates. Wholesale costs of medical therapies were estimated by prices from online pharmacies and were validated against the drug costs at Lucile Packard Children's Hospital pharmacy.

The analysis includes deterministic sensitivity analysis and probabilistic sensitivity analysis (PSA). Results are presented as mean costs and QALYs, incremental cost-effectiveness ratios (ICER), one-

way deterministic sensitivity analyses and summary results of the probability of achieving the greatest net benefit at a given willingness to pay threshold.

Table 36 presents the headline results of the economic analysis. The analysis indicates that standard medical treatment produces more health (0.06 QALYs) at a considerably greater cost than colectomy+IPAA ($88,607). The incremental cost-effectiveness of standard medical treatment versus colectomy+IPAA is estimated to be $1,476,783 per QALY gained.

Table 36: Headline cost-effectiveness results presented by Park *et al*[83]

Strategy	QALYs	Costs	Inc. QALYs	Inc. costs	ICER
Colectomy +IPAA	20.72 (17.53-22.76)	$147,763 ($137,013 to $158,904)	-	-	$1,476,783 (dominated to $3,281,923)
Standard medical treatment	20.78 (18.45-22.37)	$236,370 ($219,057 to $255,328)	0.06 (-0.72 to 1.03)	$88,607 (73,726 to $105,865)	

Confidence intervals shown in parentheses

Assuming a willingness to pay threshold of $50,000 per QALY gained, the probability that colectomy+IPAA produces the greatest net benefit is approximately 1.0. Assuming a willingness to pay threshold of $100,000 per QALY gained, the probability that colectomy+IPAA produces the greatest net benefit is approximately 0.96. The sensitivity analysis indicates that the utility of the cure state after receiving colectomy+IPAA was the only variable which reduced the ICER to below $100,000/QALY gained. The authors state that the level of HRQoL for patients with UC would need to be very low in order for exhaustive medical therapy in severe UC to be cost-effective.

The Park *et al* study clearly addresses the question as to whether colectomy+IPAA is cost-effective in comparison to medical management over a lifetime horizon. However, the description of the mathematical model is unclear, hence the assumptions underpinning the analysis are not transparent and their credibility is difficult to judge. The population reflected in the economic analysis is only partially relevant to the scope of this appraisal as the patients considered within the model are by definition hospitalised for UC flare. However, the model also appears to assume that the flare can be resolved, hence patients may go on to receive biologic therapy in a non-acute setting. Given the absence of head-to-head trials comparing medical and surgical management options, the need for an indirect comparison is inevitable and may lead to bias in the model results. It is also noteworthy that the study relates to the US setting, therefore its relevance to UK clinical practice may be questionable.

Tsai et al – A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis[82]

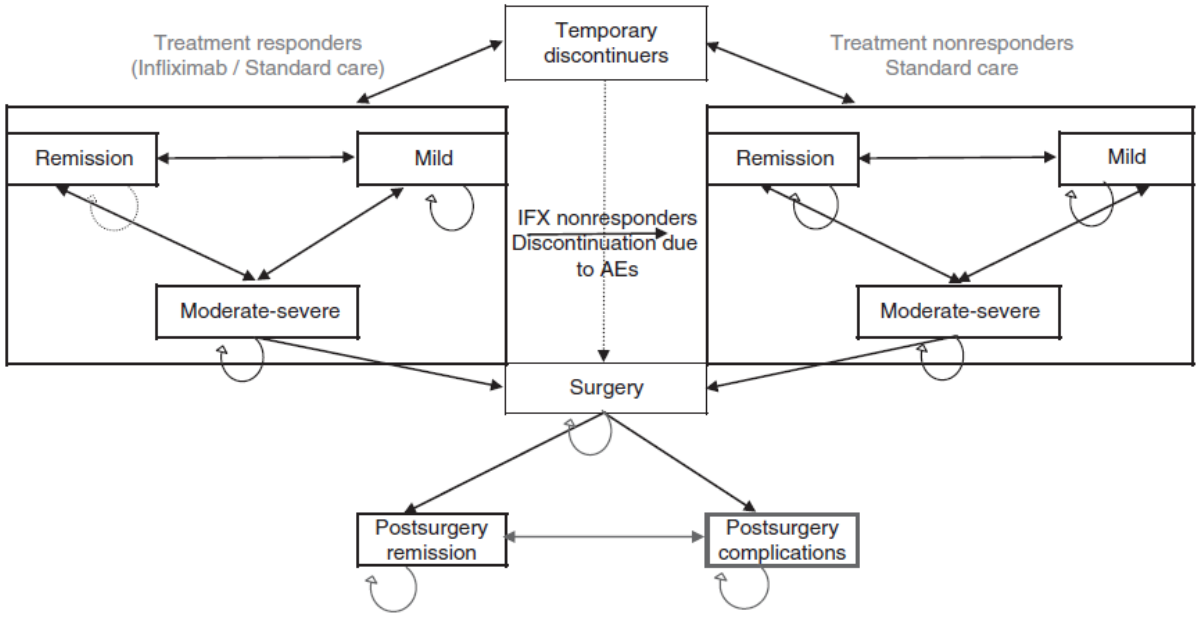
Tsai *et al* report the methods and results of an economic analysis which compares two infliximab-based strategies versus standard care in patients with moderate-severe UC from the perspective of the UK NHS. The model structure and parameter values appear to be very similar to the economic analysis submitted to NICE to inform technology appraisal TA140, although it should be noted that the total cost estimates for each group reported by Tsai *et al*[82] differ to those reported within the manufacturer's submission.[80] Patients within the model were assumed to have a mean body weight of 73.1kg. The base case scenario evaluated a treatment strategy of infliximab 5mg/kg every 8 weeks only for patients achieving response whilst a secondary analysis evaluated infliximab 5mg/kg every 8 weeks for patients achieving and maintaining remission following induction. Standard care was assumed to include colectomy+IPAA and other medications (5-ASAs, corticosteroids and immunosuppressants). Cost-effectiveness was assessed in terms of the incremental cost per QALY gained over a 10-year time horizon for each comparison of infliximab versus standard care. A fully incremental analysis was not undertaken between the two responder/remission stopping rule treatment approaches. The perspective adopted was that of the NHS. Costs borne by Personal Social Services (PSS) were not included in the model. The authors note that whilst productivity costs are substantial for patients with UC, these were not included in the economic analysis. Costs and health outcomes were discounted at a rate of 3.5% per year. Costs were valued at 2006-2007 prices.

The economic evaluation takes the form of a Markov model, as shown in Figure 75. The model appears to include eleven health states: (i) mild [responders]; (ii) moderate-severe [responders]; (iii) remission [responders]; (iv) mild [non-responders]; (v) moderate-severe [non-responders]; (vi) remission [non-responders]; (vii) temporary discontinuers; (viii) surgery [tunnel state]; (ix) post-surgery remission; (x) post-surgery complications, and; (xi) death. The cycle length used within the model was specified according to the time intervals of the assessment visits in the ACT1/2 studies. [50] The first cycle is 8 weeks in duration, followed by 6 weeks in cycle 2, followed by 8 weeks for all subsequent cycles. It should be noted that within the ACT trials, the assessments at these timepoints were based on Partial Mayo scores and may not correspond to full Mayo scores (the latter of which includes endoscopic visualisation).[105] The paper does not mention whether a half-cycle correction was applied to account for the timing of events within the model.

All patients enter the model in the moderate-severe health state. At the end of each cycle, patients achieving a Mayo score of 0-2 and 3-5 transit to the remission state and mild health state respectively and continue to receive the same treatment. Patients who do not achieve remission or response are classified as non-responders. A "temporary discontinuers" state is included for patients experiencing temporary adverse events; this is a tunnel state which is applied for one 8-week cycle. After resolution

of adverse events, these patients return to their prior health state. Non-responders and patients permanently discontinuing active treatment (e.g. due to adverse events) transit to the corresponding non-responder states and cannot restart infliximab treatment. Patients in the moderate-severe states can undergo surgery which may result in complications. Different costs and utilities are applied to each health state. The model does not include any survival difference between the competing treatments hence the differences in QALYs are driven entirely by differences in sojourn time in each health state.

Figure 75: Model diagram presented by Tsai *et al*[82]



Transition probabilities in each group were estimated using data from the ACT1/2 studies.[50] No details are provided within the paper with respect to how these studies were pooled. Transition probabilities for patients in the responder states for infliximab and standard care were drawn from the treatment and placebo arms of these trials respectively. Transition probabilities for non-responders for both arms were drawn from placebo arm of the ACT studies.[50] As ACT1 employed a longer study duration than ACT2, the former trial alone was used to estimate transition probabilities beyond 30-weeks. The ACT1/2 studies were also used to estimate the probabilities of temporary discontinuation based on the observed adverse event rates. Transition probabilities for patients undergoing surgery were derived from the literature.[106-108] None of the transition probabilities applied within the model are reported in the paper.

HRQoL values are assigned for remission (0.88), mild (0.76), moderate-severe (0.42), temporary discontinuers (0.42), surgery (0.61), post-surgery remission (0.61), post-surgery complications (0.55). Utility values for UC states are stated to have been drawn from an EQ-5D survey of UC patients, however this appears to be misreferenced as the publication title relates to resource use in patients

with Crohn's Disease.[109] Health utility values for patients who were temporary discontinuers and for those with post-surgical complications were drawn from Arseneau *et al*.[110] The text states that these utilities were indexed to the Woehl utility set to avoid any implausible results using regression analysis. No further details are provided regarding this regression analysis.

The model includes treatment- and state-specific costs associated with drug acquisition and administration, consultant visits, hospitalisations, blood tests, and endoscopy. The sources used to value the costs of drug acquisition and administration are unclear. The model includes the costs of concomitant medications based on the baseline characteristics of patients in the ACT1/2 studies,[50] and assumes that the use of immunosuppressants and 5-ASAs remain constant whilst corticosteroid use declines linearly over time for patients responding and achieving remission. The costs associated with non-serious adverse events were calculated separately but are not detailed. The costs associated with severe adverse events were assumed to be subsumed within the costs of hospitalisation. The costs of colectomy+IPAA were based on NHS Reference Costs. Hospitalisation rates for the infliximab and standard care groups were based on the ACT1/2 trials[50] and were valued using NHS Reference Costs. In addition, healthcare resource use associated with pre-surgical UC states was estimated from a panel of six UK gastroenterologists; these resource use estimates were valued using national published cost estimates.

The model results are presented as mean costs and QALYs for each treatment group. The economic analysis includes one-way deterministic sensitivity analysis and PSA. Decision uncertainty is represented using cost-effectiveness planes.

Table 37 presents the results of the economic analysis. Within the base case analysis, in which medical treatment is assumed to be continued only for those patients in whom response is achieved, infliximab is estimated to produce an additional 0.75 QALYs at an additional cost of £20,662; this corresponds to an ICER of £27,424 per QALY gained. Within the secondary analysis, in which patients are assumed to continue treatment only if they achieve and maintain remission, infliximab is estimated to produce an additional 0.39 QALYs at an additional cost of £7,615; this corresponds to an ICER of £19,696 per QALY gained. It should be noted that the estimates of absolute costs and absolute QALYs for the standard care group differs between the two analyses; whilst the alternative treatment stopping rules clearly influence which patients continue to receive infliximab and the duration over which patients would receive biologic therapy, it is unclear why this would affect outcomes for the standard care group.

Table 37: Headline cost-effectiveness results presented by Tsai *et al*[82]

Strategy	QALYs	Costs	Inc. QALYs	Inc. costs	ICER
Base case scenario (responders only)					
Infliximab	4.591	£66,460	0.75	£20,662	£27,424
Standard care	3.838	£45,798	-	-	
Secondary analysis (remission only)					
Infliximab	4.154	£53,874	0.387	£7,615	£19,696
Standard care	3.767	£46,259	-	-	

In the responders sensitivity analysis, the ICER ranged from £21,066 per QALY gained (lower patient weight) to £86,320 per QALY gained (1-year time horizon). The results of all other deterministic sensitivity analyses in the responder comparison produced an ICER below £32,000 per QALY gained. In the more stringent remission only analysis, the deterministic sensitivity analyses produced ICERs in the range £14,728 per QALY gained (lower patient weight) to £46,765/QALY (1-year time horizon). The results of all other deterministic sensitivity analyses in the remission only comparison produced an ICER below £23,000 per QALY gained. The results of the probabilistic analysis are presented in the form of cost-effectiveness planes. The authors state that *"The PSA showed that the results were robust with [the] majority of simulations clustered together. In both responder and remission treatment strategies, IFX SMT resulted in additional QALYs at an additional cost compared to standard care."*[82] The probabilistic results for the remission only scenario appear somewhat dubious as the samples appear to be truncated at the y-axis of the cost-effectiveness plane (the estimates of incremental QALYs for infliximab versus standard care cannot drop below zero). The underlying reason for this within the model is unclear.

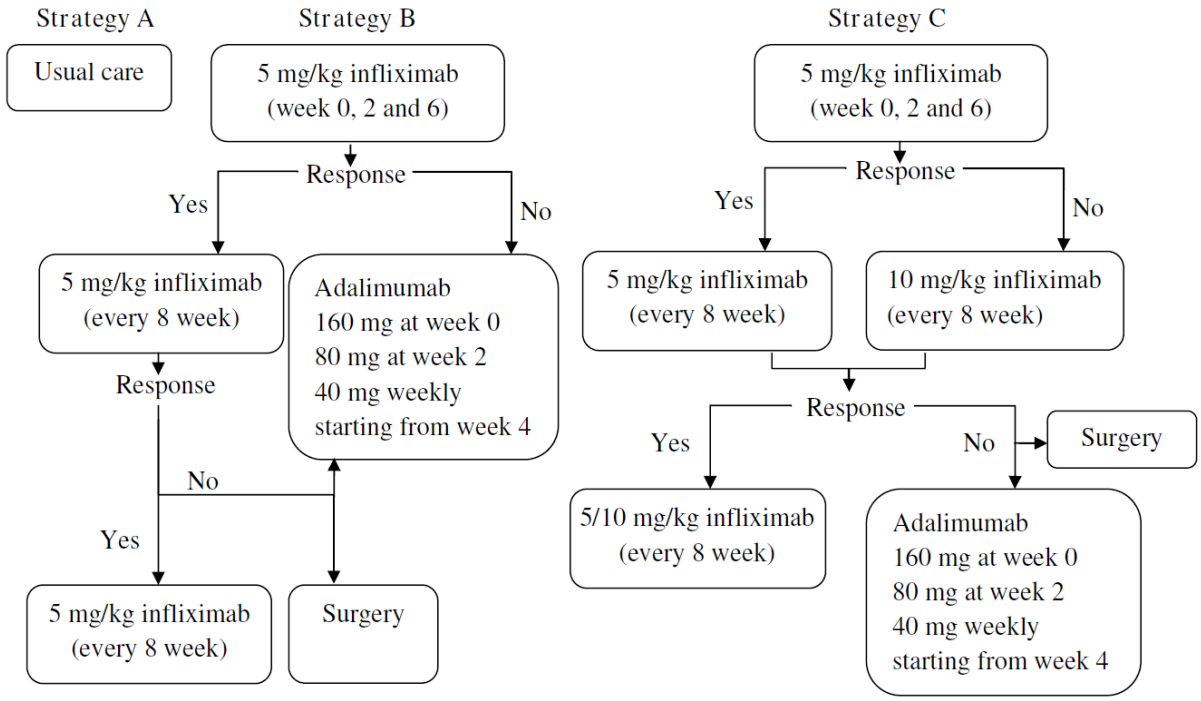
As noted earlier, the Tsai *et al* analysis appears to be based on the same model submitted to NICE as part of TA140 (not reviewed here). Whilst the QALY estimates reported within Tsai *et al*[82] are virtually the same as those reported within the manufacturer's submission to NICE,[80] the incremental costs reported by Tsai *et al*[82] are lower than those contained within the manufacturer's submission. Consequently, the ICERs presented by Tsai *et al* are notably lower than those reported within the NICE submission (responders only analysis ICER = £27,424 per QALY gained[82] versus £33,866 per QALY gained;[80] remission only analysis £19,696 per QALY gained[82] versus £25,044 per QALY gained[80]).

Overall, the Tsai *et al* model appears to follow a plausible model structure and includes the majority of costs and outcomes relevant to the decision problem. It is also noteworthy that this is the only published UK analysis included in this review. In general, the paper performs well against the Drummond checklist. The two notable issues relate to the absence of other biologic therapies (this is reasonable as golimumab and adalimumab did not have a UK marketing authorisation at the time of publication) and immediate colectomy as comparators and the use of a short time horizon.

Xie et al - Cost-utility analysis of infliximab and adalimumab for refractory ulcerative colitis[84]

Xie *et al* report a cost-effectiveness analysis comparing infliximab plus adalimumab versus usual care in patients with moderate to severe refractory UC in Canada. Patients were assumed to be 40 years of age with a mean body weight of 80kg. The model adopts a Markov approach and costs and outcomes are evaluated over a 5-year time horizon. Three options were compared within the economic analysis (1) "Strategy A" – "usual care", which includes conventional medical treatment (5-ASAs plus immunosuppressants) without anti-TNF-α drugs; (2) "Strategy B" – "5mg/kg infliximab plus adalimumab initial and maintenance therapy", which includes 5mg/kg infliximab followed by a switch to adalimumab if there is no response to initial therapy or if response is lost during maintenance therapy (3) "Strategy C" – "5mg/kg and 10mg/kg infliximab+adalimumab", which involves initial therapy using 5mg/kg infliximab, if there is no response, the dose is escalated to 10mg/kg infliximab, then if response is lost during maintenance therapy, switch to adalimumab (see Figure 76). Surgery is included in the pathway for all three treatment groups within the model but is not included as a comparator. Costs and health outcomes were discounted at a rate of 5% per year. All costs were valued at 2008 prices.

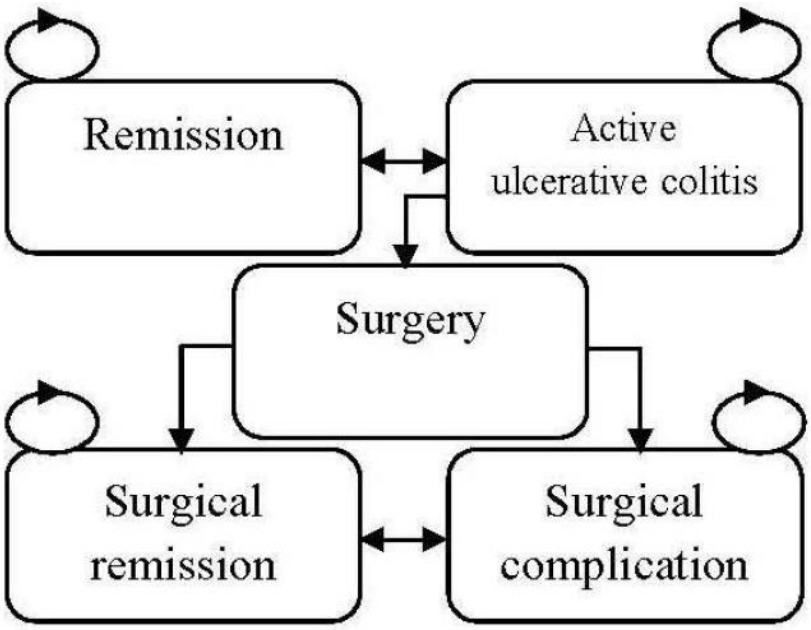
Figure 76: Treatment sequences evaluated by Xie *et al*[84]



The model includes 5 mutually exclusive health states: (i) remission, which is defined as a total Mayo score ≤2 without individual sub-scores exceeding 1 point; (ii) active UC, which includes those patients who do not respond and those who do respond but do not achieve remission; (iii) surgery, which is a tunnel state; (iv) surgical remission, and; (v) surgical complication (see Figure 77). The

model adopts a variable cycle length according to the timing of full Mayo Score assessments adopted in the ACT1/2 studies (0-8 weeks, 9-30 weeks, 31-54 weeks, then 27 weekly thereafter).[50] Patients enter the model in the active moderate-severe state and following initial therapy either achieve remission or not. Those patients who achieve remission may subsequently lose response and transit to active UC, or they may maintain remission. For responders in the active UC state, patients may achieve remission or remain in the active UC state with maintenance. Patients who are non-responders can undergo colectomy+IPAA, switch to adalimumab (strategy B) or receive an increased dose of infliximab (strategy C). Following surgery, patients may experience complications which may or may not be resolved; these complications may arise immediately after surgery or at a later timepoint. Death is not included as an event in the model due to the short time horizon. Different costs and utilities are applied to each model health state.

Figure 77: Model diagram presented by Xie *et al*[84]



The majority of clinical parameters within the model were drawn from the ACT1/2 studies,[50] derived using a fixed-effects meta-analysis. Remission rates in responders and non-responders were estimated as time-independent parameters based on the ACT1/2 studies.[50] The authors note that remission rates drawn from the placebo arms of the ACT trials reflect the use of active concomitant medications. Rates of early and late surgery were drawn from an RCT reported by Janerot *et al*[111] and a non-randomised cohort study reported by Hoie *et al*.[107] Rates of complications and the probability of their resolution were derived from non-randomised studies.[112,113] Remission rates, the probabilities of maintaining remission over time and the proportion of non-responding patients in those with active UC for each treatment group were modelled as time-dependent parameters based on the ACT1/2 studies.[50] Owing to the absence of randomised evidence at the time of the analysis, remission rates for

adalimumab were assumed to be equivalent those for 5mg/kg infliximab. The model did not incorporate the effects of adverse events on health outcomes or costs as the ACT1/2 studies reported that the proportions of patients with any adverse event were similar among the infliximab and standard care groups.

The model includes four utility values: remission (0.79), active UC (0.32), surgical remission (0.68) and surgical complications (0.49). The model makes no distinction between the HRQoL outcomes for patients who achieve response but not remission and patients who do not achieve response; this may be considered as a very pessimistic assumption which could bias against infliximab and adalimumab therapy, particularly given the low valuation of HRQoL for patients with active UC. All health utilities were drawn from a previous economic modelling study reported by Arseneau *et al.*[110] The method of utility elicitation within this study appears to be time trade-off (TTO).

The model includes the costs associated with drug acquisition, drug administration, colectomy and IPAA, medical examination and the management of surgical complications. Drug acquisition costs were drawn from provincial drug benefit lists (including an 8% mark-up). The costs of medical examinations were derived from the Ontario Schedule of Benefits. The costs of surgery were derived from the literature.[114]

The results of the economic analysis are presented as mean costs and QALYs and ICERs for each treatment group based on point estimates of parameters. Pairwise comparisons are presented for the infliximab and adalimumab options versus usual care. A fully incremental analysis between all options in the model is reported in the text. PSA was also conducted with decision uncertainty represented using cost-effectiveness acceptability curves (CEACs).

Table 38 presents the headline cost-effectiveness results reported by Xie *et al.*[84] The model analysis suggests that the Strategy B is expected to produce more health gain than Strategies A and C. Strategy C is dominated by Strategy B. The incremental cost-effectiveness of Strategy B versus usual care was estimated to be approximately $358,823 per QALY gained.

Table 38: Headline cost-effectiveness results reported by Xie *et al*[84]

Strategy	QALYs	Cost	Inc. QALYs	Inc. cost	ICER
Strategy B (infliximab+adalimumab)	2.18	$82,756	0.16	$58,488	$358,823
Strategy C (infliximab [plus dose escalation]+adalimumab)	2.15	$101,272	-	-	Dominated
Strategy A (usual care)	2.02	$24,268	-	-	-

The PSA suggests that assuming a willingness to pay threshold of $150,000 per QALY gained, the probability that usual care is optimal is approximately 1.0. The deterministic sensitivity analysis suggests that the lowest ICER is achieved by increasing the utility for remission (ICER=$273,081 per QALY gained for Strategy B vs Strategy A and $428,676 per QALY gained for Strategy C vs Strategy A), whilst the highest ICER is achieved by lowering the utility for remission (ICER=$527,236 per QALY for Strategy B vs Strategy A, $889,227 per QALY gained for Strategy C vs Strategy A).

Overall, the analysis reported by Xie *et al*[84] appears to adequately address the decision problem using a generally appropriate model. However, the analysis is limited by the use of a short time horizon, the absence of surgery as a comparator and questionable assumptions regarding the health gains associated with achieving response without remission.

Discussion of published economic evaluations

Three published economic analyses met the inclusion criteria for the systematic review. One analysis compared early colectomy+IPAA versus standard medical treatment,[83] one compared infliximab versus usual care,[82] whilst the third compared infliximab plus adalimumab versus usual care.[84] Only one study (Tsai *et al*[82]) was undertaken from the perspective of the UK NHS. The included studies were broadly consistent in terms of the disease-specific factors included in the analyses; all analyses included remission and response, and surgery as a consequence of ineffective medical treatment. Only one study (Park *et al*[83]) included the increased risk of colorectal cancer associated with UC within the analysis. One study (Xie *et al*[84]) did not include mortality for any patient group in the model. Only Park *et al*[83] included surgery as a treatment option; the other options focussed solely on medical treatment strategies. The study reported by Xie *et al*[84] included adalimumab as part of the pathway, however owing to a lack of RCT evidence at the time of the analysis, the authors assumed that adalimumab was equivalent to 5mg/kg infliximab; this assumption may not be appropriate given more recent evidence.[45,50] None of the included studies evaluated the cost-effectiveness of golimumab versus any other treatment. The time horizons considered in the economic analyses differ considerably, ranging from 5-years to the patient's remaining lifetime. It is also noteworthy that whilst the study reported by Tsai *et al* reported favourable results for infliximab (<£30,000 per QALY

gained), Xie *et al* reported considerably less favourable estimates (>CAN $350,000 per QALY gained). This contrasting finding may in part be explained by differences in assumptions regarding the level of HRQoL attributable to patients achieving response but not remission. Overall, none of the included studies present sufficient evidence relating to the cost-effectiveness of infliximab, adalimumab and golimumab versus standard medical or surgical treatment options for moderate to severe UC from the perspective of the UK NHS and PSS.

The next sections present a critique of the economic evidence submitted by the manufacturers of infliximab, adalimumab and golimumab.

6.1.1.6 Cost-effectiveness evaluation of golimumab (with PAS), infliximab and adalimumab relative to colectomy for moderate to severe ulcerative colitis in the UK (MSD submissions[65,67])

The MSD submissions include details of a systematic review of previous models together with the methods and results of a *de novo* model developed to assess the cost-effectiveness of adalimumab, golimumab, infliximab and standard non-biologic treatment for moderate to severe UC. Whilst MSD submitted two *de novo* models and two submission reports,[65,67] these relate to virtually the same overall model and analysis, hence they are detailed and critiqued together within this section.

Summary of manufacturer's review of existing economic analyses

The MSD submissions include a systematic review of economic evaluations of treatments for UC. The manufacturer undertook searches in Embase, PubMed, and the National Health Service Economic Evaluation Database (NHS EED) to identify published economic evaluations in UC to help inform the model structure and relevant parameters. A total of 12 published health economic analyses were included in the MSD review.[82,84,115-124] Several of these studies do not include biologic treatment options and some of the analyses relate to the management of severe UC exacerbations which is beyond the scope of this appraisal. The MSD review highlights the following points with respect to previous health economic analyses:

- Markov models are commonly used to evaluate treatments for UC
- all of the published models report outcomes in terms of QALYs/LYGs
- none of the studies included all relevant therapies
- there is variability in parameter sources and values between economic studies
- resource use estimates used in published models are typically derived from experts in the field rather than empirical research studies.

The MSD submissions also highlight a distinction between two distinct types of models (1) Markov models in which the model structure is based on sequences of therapies, and (2) Markov models in

which the model structure is based on severity. The MSD submissions do not report the results of these previous economic analyses, hence they are not discussed further here.

MSD model scope

The MSD model compares adalimumab 160/80/40mg, infliximab 5mg/kg, golimumab 200/100/50(100mg) and standard non-biologic treatment for patients with moderate to severe UC who have failed previous drug treatment. Standard non-biologic treatment is assumed to be immediate colectomy. The perspective of the analysis is that of the UK NHS. Golimumab is assumed to be given at an initial dose of 200mg, followed by 100mg at week 2, then 50mg every 4 weeks, thereafter for patients with body weight less than 80kg. For those patients with body mass greater than or equal to 80kg, golimumab is assumed to be given as an initial dose of 200mg, followed by 100mg at week 2, then 100mg every 4 weeks, thereafter. Infliximab is assumed to be given at a dose of 5mg/kg followed by additional 5mg/kg infusion doses at 2 and 6 weeks after the first infusion, then every 8 weeks thereafter. Adalimumab is assumed to be given as 160mg at week 0 (the dose can be administered as four injections in 1 day or as two injections per day for 2 consecutive days) and 80mg at week 2. After induction treatment, 50% of patients are assumed to receive the recommended dose of 40mg every other week (EOW) whilst the remainder are assumed to receive 40mg every week (EW). The MSD submission states that 22.9% patients in the ULTRA2 trial require dose escalation but also states that experts advising on the submission suggested that the actual proportion of patients in clinical practice may be as high as 80%. The manufacturer argues that the assumption that 50% patients dose escalate is conservative.

Patients receiving biologic treatments who achieve a response or remission at induction are assumed to continue maintenance therapy with the same biologic treatment. The model does not include sequences in which alternative biologics are used. Golimumab and adalimumab are assumed to be given as subcutaneous (s.c.) injections whilst infliximab is given as an intravenous (i.v.) infusion. For all treatment options, a proportion of patients are also assumed to receive ongoing "background" non-biologic therapies including 5-ASAs, corticosteroids and immunosuppressants. Standard clinical management, defined in the NICE scope as *"a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), and thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors and surgical intervention"*[37] is not included as a treatment option in the MSD economic analysis. Upon model entry, patients are assumed to be 40 years of age and 56% are assumed to be male. The model uses a 2-monthly cycle length. Costs and health outcomes are discounted at a rate of 3.5% each year and are evaluated over a 10-year time horizon.

MSD model structure

The MSD model structure is shown in Figures 78 and 79. The model adopts a hybrid approach whereby an initial decision tree is used to determine the probabilities of induction response or remission for biologic drug treatments, together with the probabilities of survival and the incidence of complications for patients undergoing immediate colectomy, whilst a Markov component is used to estimate long-term outcomes for maintenance drug therapy and surgery. The decision tree structure is identical for all biologic drug treatments and includes initial outcomes defined in terms of no response, response and remission. For the standard care (colectomy) option, the decision tree outcomes are different and instead relate to the probabilities of surviving surgery and experiencing early complications resulting from that surgery. The Markov model is comprised of 8 mutually exclusive health states: (1) response [pre-colectomy; maintenance]; (2) remission [pre-colectomy; maintenance]; (3) response [relapse management]; (4) relapse [relapse management]; (5) colectomy; (6) remission [post-colectomy]; (7) late complications [post-colectomy] and (8) death.

For the biologic treatment groups, patients are initially allocated to no response, response or remission based on the results of a *de novo* network meta-analysis (NMA) of induction therapy trials[45,48-50,125,126] undertaken by the manufacturer. Patients in whom response or remission is achieved at induction are assumed to remain on maintenance treatment using the same biologic treatment. Subsequent model transitions are informed by a separate NMA based on the results of the trials of biologic maintenance therapies.[45,48,50,125] Patients who do not respond to induction therapy, or those who lose response during maintenance treatment, are assumed to enter the relapse management state and receive i.v. steroids. Patients who respond to i.v. steroids then transit to the "Response (relapse management)" state where they either continue responding or relapse. Patients who do not respond are assumed to undergo immediate colectomy. Colectomy is dealt with as a tunnel state; following surgery a small proportion of patients are assumed to die whilst the remainder are assumed to be in post-colectomy remission. Patients who survive their surgery are assumed to be at ongoing risk of post-colectomy complications (anal fistula, bowel obstruction and pouchitis). A small proportion of patients receiving drug treatment are assumed to be at risk of serious infection and hospitalisation.

Figure 78: MSD model structure (re-drawn by the Assessment Group)

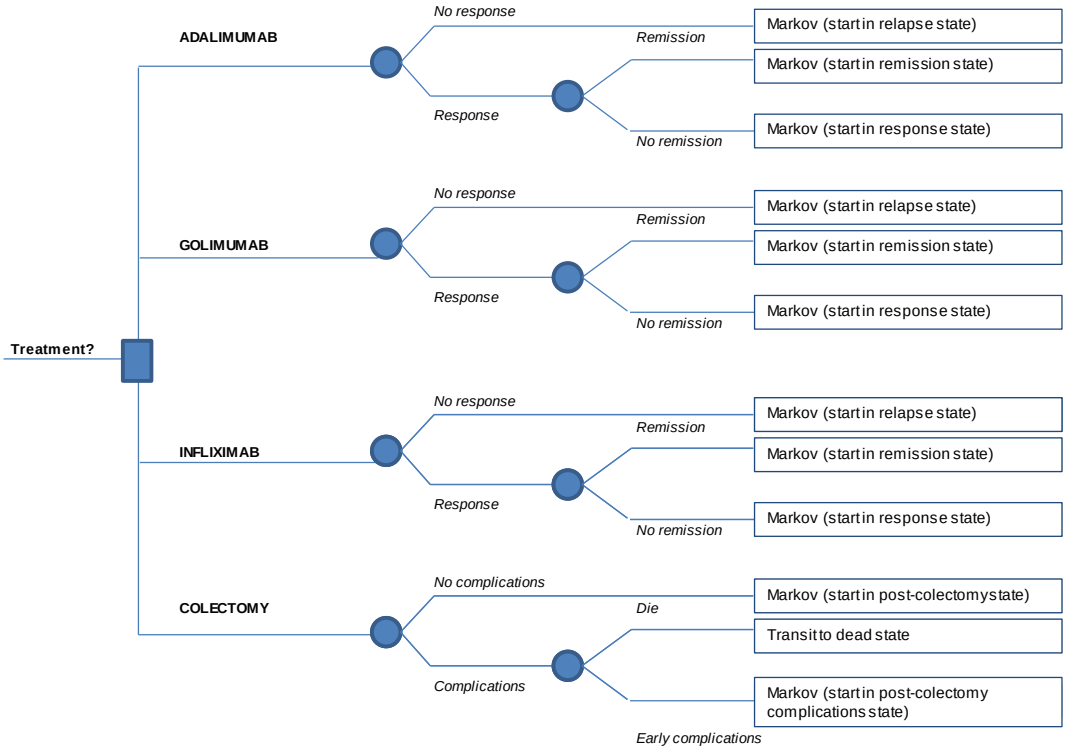
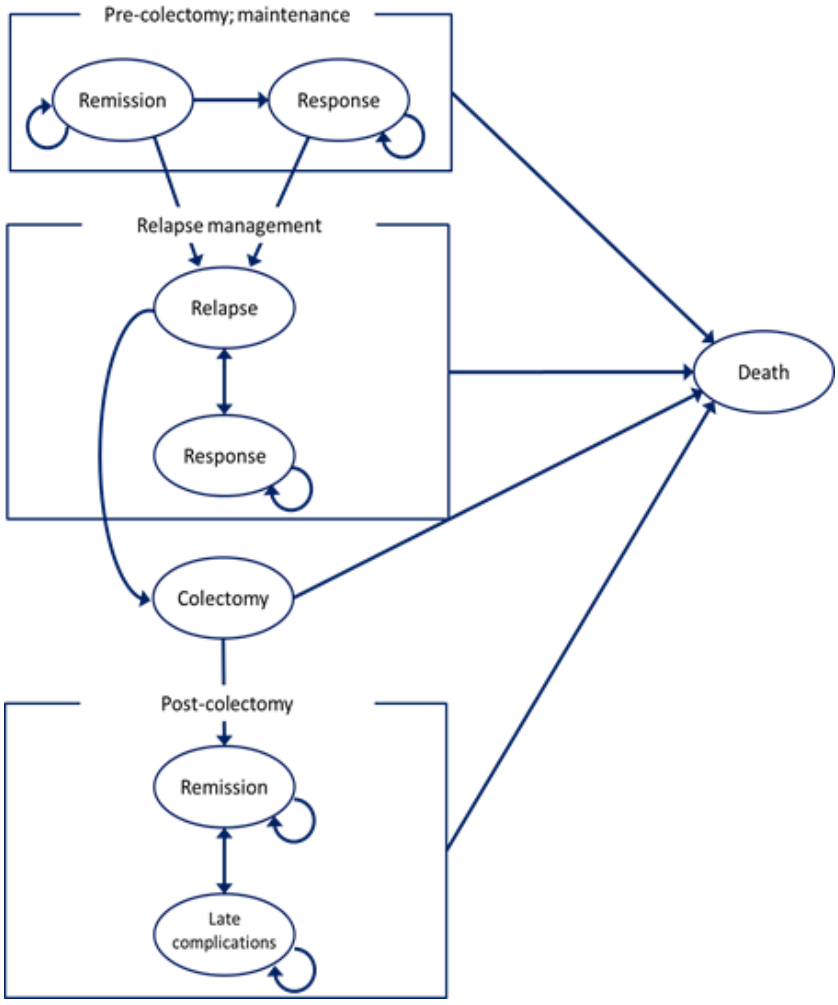


Figure 79: MSD model Markov component[65,67]



The model uses simple matrix multiplication to determine health state populations during each model cycle based on the state population in the previous Markov cycle and a single time-independent transition matrix over the entire time horizon. Costs and utilities are attached to each health state. Total QALYs are modelled as a function of sojourn time in each health state, mortality associated with colectomy and other-cause (general population risk) mortality.

Evidence sources used to inform the MSD model parameters

A summary of evidence sources used to inform the model's parameters is presented in Table 39.

Table 39: Summary table of evidence sources used to inform the MSD model parameters

Parameter group	Source
Pre-colectomy transition probabilities excluding death – standard non-biologic treatment	Manufacturer's NMA - average study effect of placebo controlled trials in random effects NMA of induction response[65,67]
Odds ratios for biologic treatment effects	Odds ratios derived from manufacturer's NMA[65,67]
Effectiveness of i.v. steroids following failure of biologic treatment	Model fitted to ensure 27% relapsers require colectomy based on Turner *et al*[127]
Probability of serious infection	Grijalva *et al*[128]
Health utilities for pre-colectomy response / remission	EQ-5D estimates from the PURSUIT trial[48,49] in the golimumab model, EQ-5D estimates from the ACT1/2[50] trials in the infliximab model
Health utilities for post-colectomy states*	Woehl *et al*,[109] Tsai *et al*,[82] HODaR, Punekar and Hawkins,[119] Chaudhary and Fan,[115] Arseneau *et al*[110]
Resource use	PURSUIT trial,[48] ACT1/2 trials[50] and interviews with 9 gastroenterologists[65,67]
Unit costs*	Curtis *et al*,[129] NHS Reference Costs[130]

** Sources for some HRQoL parameters are not clear from the MSD submissions*

Methods for modelling effectiveness

Estimates of relative effectiveness of biologic treatments versus conventional non-biologic non-surgical treatment were derived from NMA models of induction and maintenance therapy undertaken by the manufacturer.[65,67]

The baseline model employed within the MSD NMA model is not discussed within the submissions.[65,67] The MSD economic model includes a worksheet named *"Input Efficacy and Trans Prob"* in which the probabilities of response and remission for non-biologic therapy are inputted as 0.36 and 0.09 for induction treatment, and 0.83 and 0.86 per 2-month cycle of maintenance therapy respectively. The source is stated in the model as *"Average study effect of placebo controlled trials in random effects NMA of induction response."* No additional detail on the baseline model is provided within the MSD submissions, thus it is not possible to determine whether these estimates are appropriate.

Relative treatment effects were drawn from *de novo* NMAs undertaken by the manufacturer, based on the results of a systematic literature review. Separate analyses were undertaken for induction and maintenance therapy. For induction, a NMA was undertaken using data from 6 RCTs.[45,125,46,48-50] For maintenance treatment, relative treatment effects were based on a NMA of 3 RCTs.[46,48,50] The evidence networks employed in the manufacturer's NMAs are presented in Figures 80 and 81 respectively. It should be noted that the manufacturer's NMA includes non-licensed indications of infliximab, although these are not included in the health economic analysis.

Figure 80: Evidence network for induction therapy[65,67]

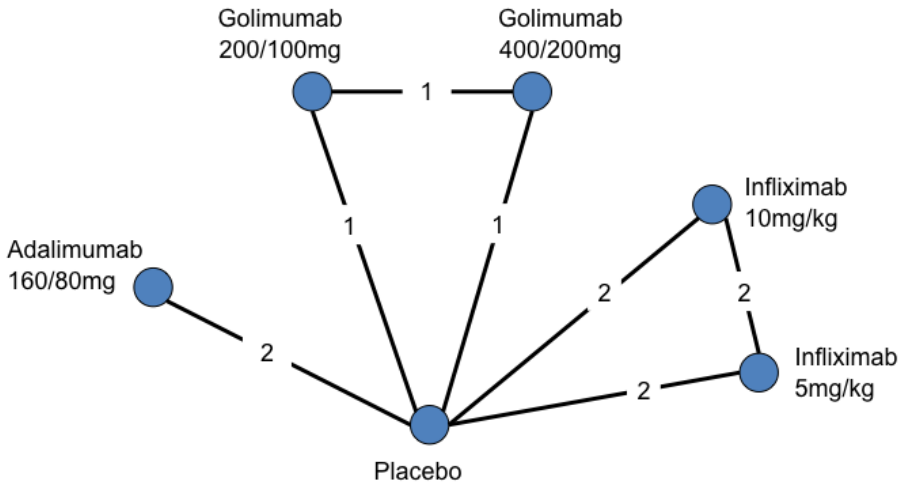
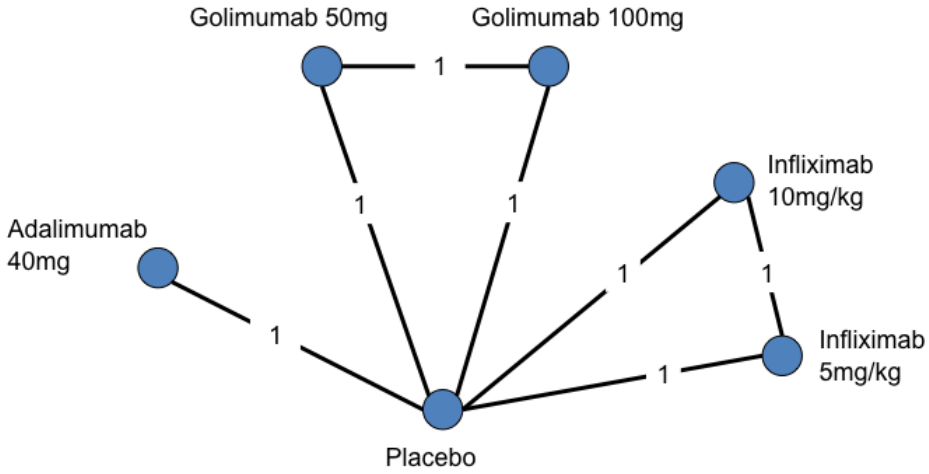


Figure 81: Evidence network for maintenance therapy[65,67]



The NMAs use logistic regression models to estimate treatment effects, given an assumption that the data are binomial (separate models are used to estimate the odds of sustained response and the odds of sustained remission respectively). Relative treatment effects were parameterised in terms of odds ratios and are converted to relative risks in the health economic model. In instances whereby only one RCT informed each treatment (which was predominantly the case for the maintenance outcomes),

heterogeneity could not be estimated and therefore a fixed-effects model was employed. Where multiple studies were available, a random-effects approach was used.[65,67] The NMA model for maintenance therapy includes a complex "novel" imputation of estimates of sustained response and sustained remission for golimumab from the PURSUIT trial using data from the non-randomised placebo group (see MSD golimumab submission page 55[67]). The results of the manufacturer's NMA are presented in Tables 40 and 41 respectively.

Table 40: Manufacturer's NMA results – induction treatment

Intervention	Odds ratio	Lower 95% CrI	Upper 95% CrI
Induction response (reference placebo)			
Golimumab 200/100mg	2.12	1.01	3.95
Golimumab 400/200mg	2.47	1.19	4.65
Infliximab 5mg/kg	4.12	2.08	8.14
Infliximab 10mg/kg	3.81	1.95	7.59
Adalimumab 160mg/kg	1.87	0.96	3.65
Induction remission (reference placebo)			
Golimumab 200/100mg	2.99	1.32	6.28
Golimumab 400/200mg	3.32	1.56	7.23
Infliximab 5mg/kg	5.27	2.60	11.64
Infliximab 10mg/kg	3.90	1.88	8.56
Adalimumab 160mg/kg	2.25	1.08	4.72

Table 41: Manufacturer's NMA results – maintenance treatment

Intervention	Odds ratio	Lower 95% CrI	Upper 95% CrI
Sustained response (reference placebo)			
Adalimumab 40mg	1.31	0.67	2.59
Infliximab 5mg/kg	2.12	1.02	4.54
Infliximab 10mg/kg	2.51	1.17	5.51
Golimumab 50mg	1.51	0.94	2.47
Golimumab 100mg	1.75	1.08	2.84
Golimumab 50mg-100mg	1.62	1.07	2.50
Placebo following golimumab	0.78	0.47	1.28
Sustained remission (reference placebo)			
Adalimumab 40mg	0.76	0.22	2.56
Infliximab 5mg/kg	1.30	0.44	4.05
Infliximab 10mg/kg	2.26	0.73	7.49
Golimumab 50mg	0.83	0.29	2.40
Golimumab 100mg	0.98	0.36	2.78
Golimumab 50mg-100mg	0.92	0.36	2.45
Placebo following golimumab	0.45	0.15	1.34

Table 42 illustrates how the ORs are applied within the transition matrix for maintenance therapy within the health economic model, using the adalimumab treatment group as an example.

Table 42: Transition matrix for adalimumab

Health state	Response (pre-colectomy; maintenance)	Remission (pre-colectomy; maintenance)	Response (relapse management)	Relapse (relapse management)	Colectomy	Remission (post-colectomy)	Late Complications (post-colectomy)	Death related to UC
Response (pre-colectomy; maintenance)	0.86*	0.00	0.00	0.14†	0.00	0.00	0.00	0.00
Remission (pre-colectomy; maintenance)	0.17‡	0.83§	0.00	0.00‖	0.00	0.00	0.00	0.00
Response (relapse management)	0.00	0.00	0.83	0.17	0.00	0.00	0.00	0.00
Relapse (relapse management)	0.00	0.00	0.73	0.00	0.27	0.00	0.00	0.00
Colectomy	0.00	0.00	0.00	0.00	0.00	0.97	0.00	0.03
Remission (post-colectomy)	0.00	0.00	0.00	0.00	0.00	0.99	0.01	0.00
Late Complications (post-colectomy)	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00

**probability of maintaining response on non-biologic x relative risk of maintaining response on biologic treatment.*

†probability of relapse calculated as one minus the sum of the row of probabilities

‡One minus probability of sustained remission on non-biologic x relative risk of sustained remission on biologic treatment.

§ Probability of maintaining remission calculated as one minus the sum of the row of probabilities

‖ Transition probability set to zero for all biologic treatments

It should be noted that the matrix employed within the MSD models does not match the transitions implied by the diagram within the MSD submissions[65,67] (see Figure 79). In the executable model, patients who have previously achieved a response can either maintain or lose that response, but they cannot improve (i.e. they cannot subsequently transit to the remission state). Patients who have previously achieved remission can either maintain or lose that remission. However, upon losing remission, the patient cannot transit directly to relapse – they transit to the response state first. This means that no additional patients can achieve remission after induction and no patients with remission can completely lose response during any given model cycle. It is also noteworthy that patients who discontinue treatment with a biologic treatment transit very quickly to colectomy (27% of all non-responding relapsers during each 2-month cycle). This latter value was based on a meta-regression of studies describing the short-term outcomes for adult and paediatric patients treated with i.v. corticosteroids, with or without ciclosporin, for exacerbations of UC.[127]

The model also includes a small risk of experiencing serious infection due to the use of immunosuppressants and biologic therapies based on Grijalva *et al*. The model assumes a hazard ratio of 1.10 for all biologic therapies and a baseline risk of 0.16 for non-biologic therapy.

Health-related quality of life (HRQoL)

The health utility values used in the MSD model are presented in Table 41. Health utilities associated with failure, response, and remission as a result of induction and maintenance treatment, and utility values assigned to the health states "Response (pre-colectomy; maintenance)" and "Remission (pre-colectomy; maintenance)" are assumed to be the same for all biologics, based on EQ-5D valuations derived from the PURSUIT trial[48] within the golimumab model and the ACT1 trial[50] within the infliximab models. The MSD submission suggests that disutilities for adverse events associated with biologics are likely to be captured within these estimates. Different utilities are assumed for the achievement of the same outcome at induction and maintenance (i.e. the utility for response at induction is not the same as the utility for response at maintenance). Utility values for colectomy, post-colectomy and early and late complications of colectomy were based on estimates reported within the literature, although the precise sources are not clear from the MSD submissions[65,67] (see Table 43).

Table 43: Health utility values used in the MSD models

Health state	Utility value (golimumab model; infliximab model)	Valuation method and source (golimumab model; infliximab model)
Response (pre-colectomy; induction)	0.80; 0.79	EQ-5D PURSUIT trial;[48] ACT1 trial
Remission (pre-colectomy; induction)	0.86; 0.84	EQ-5D PURSUIT trial;[48] ACT1 trial
No response (pre-colectomy; induction)	0.70; 0.70	EQ-5D PURSUIT trial;[48] ACT1 trial
Response (pre-colectomy; maintenance)	0.80; 0.82	EQ-5D PURSUIT trial;[48] ACT1 trial
Remission (pre-colectomy; maintenance)	0.89; 0.88	EQ-5D PURSUIT trial;[48] ACT1 trial
Response (relapse management)	0.76; 0.76	EQ-5D PURSUIT trial;[48] ACT1 trial
Relapse (relapse management)	0.42; 0.42	EQ-5D estimates from Tsai *et al*[82] – however the primary source of these estimates (Woehl *et al*[109]) appears to be misreferenced as the cited reference is not a health valuation study and does not report utilities.
Colectomy	0.56; 056*	Unclear. Appears to be based on data from HoDAR reported within Punekar *et al*[119] Disutility for early complications based on time-trade-off (TTO) study reported by Arseneau *et al*.[110]
Remission (post-colectomy)	0.60; 0.60	
Late complications (post-colectomy)	0.60; 0.60	

** includes disutility for proportion of patients experiencing early complications of surgery*

Resource use and costs

The model includes direct costs of drug acquisition, consultant visits, endoscopy, inpatient hospital admissions, colectomy, management of surgery-related complications, adverse events and other UC costs. A Patient Access Scheme (PAS), in which the price of 100mg golimumab is assumed to be equal to that of 50mg golimumab, is applied within the model. It should be noted that at the time of writing, this had not been approved by the Department of Health.

Drug acquisition costs

Cost of non-biologic "background" therapies

The usage and per-cycle costs of non-biologic background therapies assumed within the model are presented in Tables 44 and 45. Patients in the standard non-biologic treatment group are assumed to receive mesalazine 4g daily (acute), 2g daily (chronic), azathioprine 2-2.5mg/kg daily, 6-mercaptopurine 1-1.5mg/kg daily, ciprofloxacin 500mg twice daily and prednisolone. The same use of background therapies is assumed for all biologic treatment arms. It should be noted however that in the colectomy group, patients are assumed to undergo immediate colectomy, so the actual drug acquisition cost for the colectomy group is zero within the model. Patients in the biologic treatment groups are assumed to receive the same "background therapies" with the exception of ciprofloxacin

(although as described above, this is applied as a zero cost in the colectomy group). Resource use estimates for these therapies appear to be based on the placebo arm of the PURSUIT trial[48] within the golimumab model and from the placebo arm of the ACT1/2 trials[50] within the infliximab model. These are similar, but not the same (see Tables 44 and 45); as with the utility values, the justification for using different assumptions concerning resource use in each model is not clear. The source of the unit costs is not reported within the submission, however, estimates appear to be drawn from the British National Formulary (BNF[38]).

Table 44: Background therapies resource use and costs used in MSD golimumab model

Treatment group	Background therapies included (proportion of patients)	Cost per cycle
Induction treatment		
Standard non-biologic treatment	Mesalazine (0.83), azathioprine (0.15), 6-MP (0.15), ciprofloxacin (0.00), prednisolone (1.00)	£251.43*
Adalimumab	Mesalazine (0.81), azathioprine (0.16), 6-MP (0.16), prednisolone (0.44)	£200.03
Golimumab	Mesalazine (0.81), azathioprine (0.16), 6-MP (0.16), prednisolone (0.44)	£200.03
Infliximab	Mesalazine (0.81), azathioprine (0.16), 6-MP (0.16), prednisolone (0.44)	£200.03
Maintenance treatment		
Standard non-biologic treatment	Mesalazine (0.80), azathioprine (0.16), 6-MP (0.16), ciprofloxacin (0.00), prednisolone (0.49)	£121.15*
Adalimumab	Mesalazine (0.80), azathioprine (0.15), 6-MP (0.15), prednisolone (0.51)	£120.98
Golimumab	Mesalazine (0.80), azathioprine (0.15), 6-MP (0.15), prednisolone (0.51)	£120.98
Infliximab	Mesalazine (0.80), azathioprine (0.15), 6-MP (0.15), prednisolone (0.51)	£120.98
Relapse management (following prior treatment failure)		
Relapse management	Mesalazine (0.80), azathioprine (0.16), 6-MP (0.16), ciprofloxacin (0.00), prednisolone (0.49)	£121.15
Relapse management (i.v. steroids)	Mesalazine (0.83), azathioprine (0.15), 6-MP (0.15), ciprofloxacin (0.00), prednisolone (1.00), i.v. prednisolone (1.00)	£405.43

** acquisition costs not included in model results for standard non-biologic treatment*

Table 45: Background therapies resource use and costs used in MSD infliximab model

Treatment group	Background therapies included (proportion of patients)	Cost per cycle
Induction treatment		
Standard non-biologic treatment	Mesalazine (0.71), azathioprine (0.15), 6-MP (0.15), ciprofloxacin (0.00), prednisolone (1.00)	£233.57*
Adalimumab	Mesalazine (0.72), azathioprine (0.36), 6-MP (0.13), prednisolone (0.54)	£191.11
Golimumab	Mesalazine (0.72), azathioprine (0.36), 6-MP (0.13), prednisolone (0.54)	£191.11
Infliximab	Mesalazine (0.72), azathioprine (0.36), 6-MP (0.13), prednisolone (0.54)	£191.11
Maintenance treatment		
Standard non-biologic treatment	Mesalazine (0.71), azathioprine (0.15), 6-MP (0.15), ciprofloxacin (0.00), prednisolone (0.57)	£118.10*
Adalimumab	Mesalazine (0.72), azathioprine (0.36), 6-MP (0.13), prednisolone (0.54)	£113.99
Golimumab	Mesalazine (0.72), azathioprine (0.36), 6-MP (0.13), prednisolone (0.54)	£113.99
Infliximab	Mesalazine (0.72), azathioprine (0.36), 6-MP (0.13), prednisolone (0.54)	£113.99
Relapse management (following prior treatment failure)		
Relapse management	Mesalazine (0.71), azathioprine (0.29), 6-MP (0.15), ciprofloxacin (0.00), prednisolone (0.57)	£118.10
Relapse management (i.v. steroids)	Mesalazine (0.71), azathioprine (0.29), 6-MP (0.15), ciprofloxacin (0.00), prednisolone (1.00), i.v. prednisolone (1.00)	£387.57

** acquisition costs not included in model results for colectomy group*

Biologic therapies

Table 46 shows the biologic acquisition costs per cycle for each treatment group. The table indicates that the estimated costs of induction using infliximab is markedly higher than that for adalimumab and golimumab, however, the costs of maintenance therapy per cycle are broadly similar for all biologics.

Table 46: Biologic treatment resource use and costs used in MSD models

Treatment group	Assumed regimen	Cost per cycle
Induction treatment (8 week cycle)		
Adalimumab	All patients receive 1 x 160mg adalimumab + 1 x 80mg adalimumab + 3 x 40mg adalimumab	£3,169.26
Golimumab	All patients receive 2 x 200mg golimumab + 2 x 100mg golimumab	£3,051.88
Infliximab	All patients receive 12 x 100mg infliximab over 3 administrations	£5,497.44
Maintenance treatment (2-month cycles)		
Adalimumab	50% patients receive 40mg adalimumab EW (4.33 doses/cycle); 50% patients receive 40mg adalimumab EW (8.67 doses/cycle)	£2,288.91
Golimumab	31.6% patients receive 100mg golimumab every 4 weeks; 68.4% patients receive 50mg golimumab every 4 weeks	£1,653.10
Infliximab	All patients receive 5mg/kg infliximab EOW	£1,985.19

Health state resource costs

Tables 47 and 48 present the health state costs (excluding drug acquisition) for the biologic and colectomy groups, respectively. The resource use estimates underpinning these cost estimates were reported to be based on interviews with nine expert gastroenterologists. Resource use was costed using standard costing sources.[129,131]

Table 47: Other health state costs per 2-month cycle – biologic treatments

State and treatment phase	Consultant visit cost	Endoscopy cost	Inpatient cost	Colectomy cost	Late complications cost	Other UC cost	AE cost	Total cost per cycle
Response; induction phase	£91.58	£18.32	£0.00	£0.00	£0.00	£1.80	£29.37	£141.07
Remission; induction phase	£43.70	£4.97	£0.00	£0.00	£0.00	£1.80	£29.37	£79.84
Failure; induction phase	£162.76	£40.30	£0.00	£0.00	£0.00	£1.80	£29.37	£234.22
Response (pre-colectomy; maintenance)	£91.58	£18.32	£0.00	£0.00	£0.00	£1.80	£29.37	£141.07
Remission (pre-colectomy; maintenance)	£43.70	£4.97	£0.00	£0.00	£0.00	£1.80	£29.37	£79.84
Response (relapse management)	£91.58	£18.32	£0.00	£0.00	£0.00	£1.80	£26.74	£138.44
Relapse (relapse management)	£162.76	£40.30	£350.86	£0.00	£0.00	£3.44	£26.74	£584.09
Colectomy	£162.76	£40.30	£0.00	£8,967.94	£0.00	£3.44	£0.00	£9,174.42
Remission (post-colectomy)	£53.12	£45.27	£0.00	£0.00	£0.00	£0.92	£0.00	£99.30
Late Complications (post-colectomy)	£67.77	£26.17	£0.00	£0.00	£2,446.85	£1.85	£0.00	£2,542.64

Table 48: Other health state costs 2-month cycle – colectomy group*

State and treatment phase	Consultant visit cost	Endoscopy cost	Colectomy cost w/o early complications	Cost early complicatons of colectomy	Colectomy cost	Late complications cost	Other UC cost	Total cost per cycle
Death due to colectomy Remission (post-colectomy) due to colectomy	£162.76	£40.30	£7,619.25	£4,029.61	£8,967.94	£0.00	£3.44	£20,823.28
Remission (post-colectomy) due to colectomy	£162.76	£40.30	£7,619.25	£4,029.61	£8,967.94	£0.00	£3.44	£20,823.28
Colectomy	£162.76	£40.30	£7,619.25	£4,029.61	£8,967.94	£0.00	£3.44	£20,823.28
Remission (post-colectomy)	£53.12	£45.27	£0.00	£0.00	£0.00	£0.00	£0.92	£99.30
Late Complications (post-colectomy)	£67.77	£26.17	£0.00	£0.00	£0.00	£2,446.85	£1.85	£2,542.64

**Note – the model includes ten further rows of costs by state for the colectomy group however none of these influence the model results*

Model evaluation and uncertainty analysis

The results of the economic analysis are presented as pairwise incremental cost-effectiveness ratios (ICERs) and are interpreted as net monetary benefits (NMB) assuming a willingness to pay threshold of £30,000 per QALY gained. Incremental CEACs are also presented within the submission (see MSD golimumab submission[67] page 122). Uncertainty surrounding estimates of incremental costs and health outcomes was examined using deterministic sensitivity analyses and PSA. The results of the deterministic analyses are presented as tornado diagrams whilst the results of the PSA are presented as cost-effectiveness planes and CEACs.

MSD model results

Tables 49 and 50 present the results within the golimumab and infliximab submissions, respectively[65,67] (note - the fully incremental analysis presented here has been undertaken by the Assessment Group rather than by the manufacturer).

Table 49: Model results from golimumab submission[67] (including PAS)

Treatment	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Probabilistic model results					
Infliximab	5.70	£44,382.28	0.16	£13,003.60	£80,318
Golimumab	5.54	£31,378.68	0.56	£15,610.91	£27,994
Adalimumab	5.49	£32,096.50	-	-	Dominated
Colectomy	4.98	£15,767.78	-	-	-
Results based on point estimates of parameters					
Infliximab	5.65	£43,091.60	0.15	£12,196.82	£80,866
Golimumab	5.50	£30,894.78	0.55	£15,100.53	£27,322
Adalimumab	5.45	£31,370.28	-	-	Dominated
Colectomy	4.95	£15,794.26	-	-	-

Table 50: Model results from infliximab submission[65] (including PAS)

Treatment	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Probabilistic model results					
Infliximab	5.71	£44,189.50	0.17	£12,841.74	£75,998
Golimumab	5.54	£31,347.76	0.57	£15,522.79	£27,163
Adalimumab	5.48	£32,123.34	-	-	Dominated
Colectomy	4.97	£15,824.96	-	-	-
Results based on point estimates of parameters					
Infliximab	5.66	£42,919.73	0.16	£12,166.45	£77,599
Golimumab	5.51	£30,753.28	0.56	£14,963.69	£26,569
Adalimumab	5.45	£31,237.38	-	-	Dominated
Colectomy	4.94	£15,789.59	-	-	-

The model results suggest that infliximab is expected to produce the greatest QALY gain, followed by golimumab and adalimumab. Adalimumab is expected to be less effective and more expensive than

golimumab hence it is ruled out due to simple dominance. The ICER for golimumab versus colectomy is expected to be approximately £27,000 to £28,000 per QALY gained. The ICER for infliximab versus golimumab is expected to be approximately £76,000 to £80,000 per QALY gained. The probabilistic results are slightly different to those derived from point estimates of parameters, however the ICERs appear stable.

Figure 82 presents the results of the PSA in the form of a cost-effectiveness plane for each biologic treatment relative to immediate colectomy. It can be seen that the results overlap considerably for adalimumab and golimumab, however, the plane indicates a generally higher overall cost for infliximab. The dispersion of sampled incremental QALY gains for infliximab versus colectomy is greater than that for adalimumab and golimumab versus colectomy.

Figure 82: Cost-effectiveness plane from MSD model[65,67]

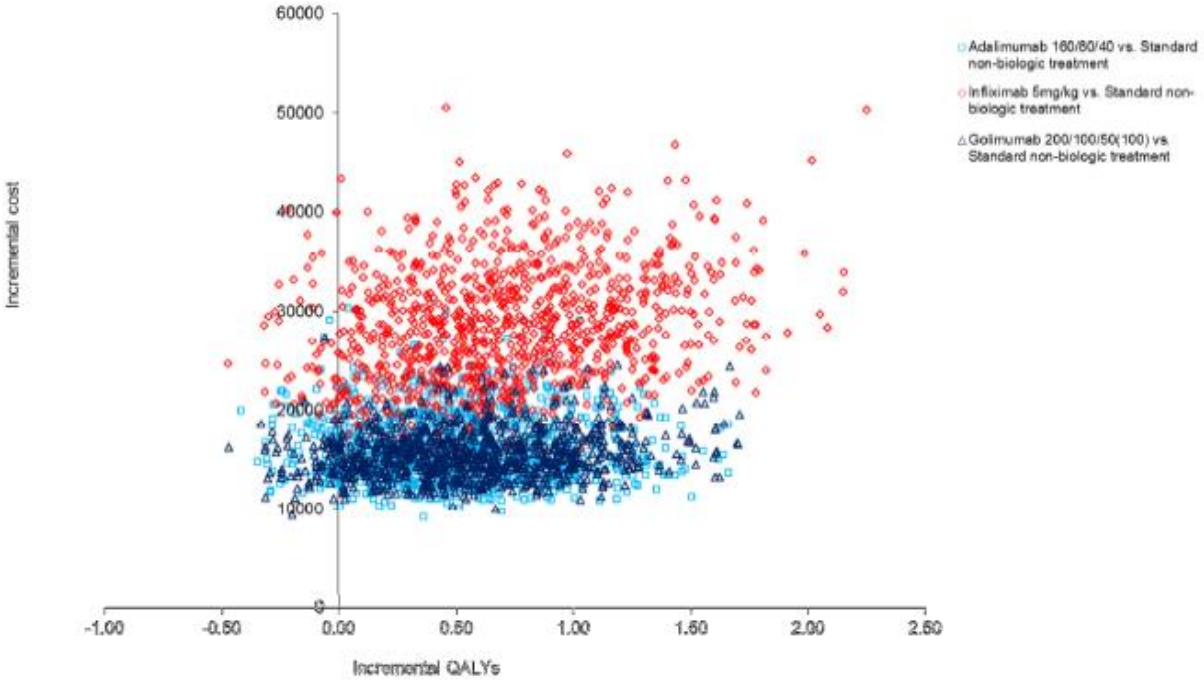
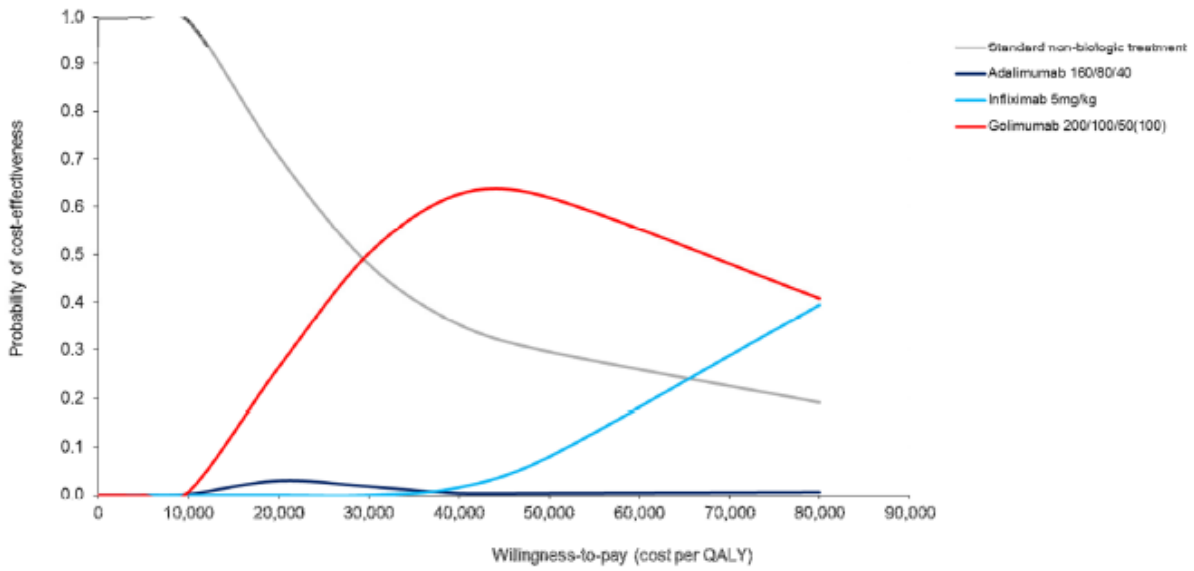


Figure 83 presents incremental CEACs for all options in the model. The CEACs suggest that at willingness to pay thresholds of £25,000 per QALY gained or lower, immediate colectomy has the highest probability of producing the greatest net benefit. At a willingness to pay threshold of £30,000 per QALY gained, golimumab has the highest probability of producing the greatest net benefit, although this is only very slightly higher than 0.50.

Figure 83: Incremental CEACs from MSD model[65,67]



A number of deterministic sensitivity analyses are also presented, however these are difficult to interpret as both infliximab and golimumab are compared in a pairwise manner against colectomy using incremental QALYs, incremental costs and incremental NMB. Deterministic sensitivity analyses are not presented between competing biologic therapies. The DSAs indicate that the cost-effectiveness results for golimumab versus colectomy are sensitive to the utility values for pre-colectomy remission and pre-colectomy relapse,[67] whilst the cost-effectiveness results for infliximab versus colectomy are sensitive to the utility value for post-colectomy remission.[65]

Critical appraisal of the MSD model

The main issues identified by the Assessment Group are presented in Box 2.

Box 2: Main problems and concerns relating to the MSD model

1. Deviations from the NICE Reference Case and final NICE scope, particularly with respect to omission of conventional non-surgical management as a comparator
2. Assumption that treatment failure is equivalent to severe exacerbation
3. Questionable use of evidence concerning improving and worsening of inflammation
4. Questionable validity of use of novel methods for including non-randomised data from PURSUIT
5. Lack of clarity regarding the NMA model
6. Inconsistencies between results of the MSD golimumab and infliximab models
7. Lack of clarity regarding the identification, selection and use of certain model parameters
8. Complex implementation of the model
9. Failure to undertake an incremental analysis
10. Inclusion of a PAS for golimumab which has not yet been agreed by the Department of Health

(1) Deviations from NICE Reference Case and final NICE scope

The extent to which the economic analyses reported in the MSD submissions adhere to the NICE Reference Case is presented in Table 51.

Table 51: Adherence of the MSD model to the NICE Reference Case[78]

Element of health technology assessment	Reference case	Assessment Group comments
Defining the decision problem	The scope developed by the Institute	The scope of the analysis deviates from the final scope from NICE. Non-biologic treatment is assumed to be immediate colectomy. Standard clinical management, defined in the NICE scope as *"a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), and thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors and surgical intervention",*[37] is not included as a treatment option within the model.
Comparator(s)	As listed in the scope developed by NICE	
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health outcomes reflect those of patients with UC
Perspective on costs	NHS and PSS	The economic analysis was undertaken from the perspective of the UK NHS. PSS costs are not mentioned in the submission.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The model is a cost-effectiveness analysis. Analyses are presented as pairwise comparisons rather than a fully incremental economic analysis of all options.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Costs and outcomes are evaluated over a 10-year time horizon. Analyses over a lifetime horizon are not presented in the manufacturer's submission.
Synthesis of evidence on health effects	Based on systematic review	Outcomes are synthesised using NMA models using studies identified through a systematic review.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults	Health outcomes are reported in terms of life years gained LYGs and QALYs gained.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	All utilities except the disutility for surgery-related complications are based on EQ-5D measurements from UC patients and are valued by the general public.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight	No equity weighting is applied.

Element of health technology assessment	Reference case	Assessment Group comments
	regardless of the other characteristics of the individuals receiving the health benefit	
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The economic analysis was undertaken from the perspective of the UK NHS. The sources for prices are not entirely clear.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	The model uses a discount rate of 3.5% for costs and health outcomes.

Overall, the MSD economic analyses are generally in line with the NICE Reference Case. The analysis does however make one important deviation from the final NICE scope with respect to the options included in the economic analysis; non-biologic treatment is assumed to be immediate colectomy. Standard clinical management, which is defined in the NICE scope as *"a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), and thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors and surgical intervention",*[37] is not included in the MSD model. The omission of non-biologic non-surgical treatment options from the MSD model is neither discussed nor justified in the MSD submissions.[65,67]

It is also noteworthy that the MSD model adopts a 10-year time horizon; at this point around 96% of patients in the model are still alive in each treatment group. The MSD submissions state that the *"…time horizon of 10 years can be considered sufficiently long to capture differences in the distribution of health states between the compared biologics; after 10 years of follow-up all patients are expected to have discontinued biologic treatment."* The modelled profiles of incremental costs and benefits are slightly different when a longer time horizon is adopted. It is reasonable to suggest that the manufacturer should have examined the impact of using different time horizons within their economic analysis.

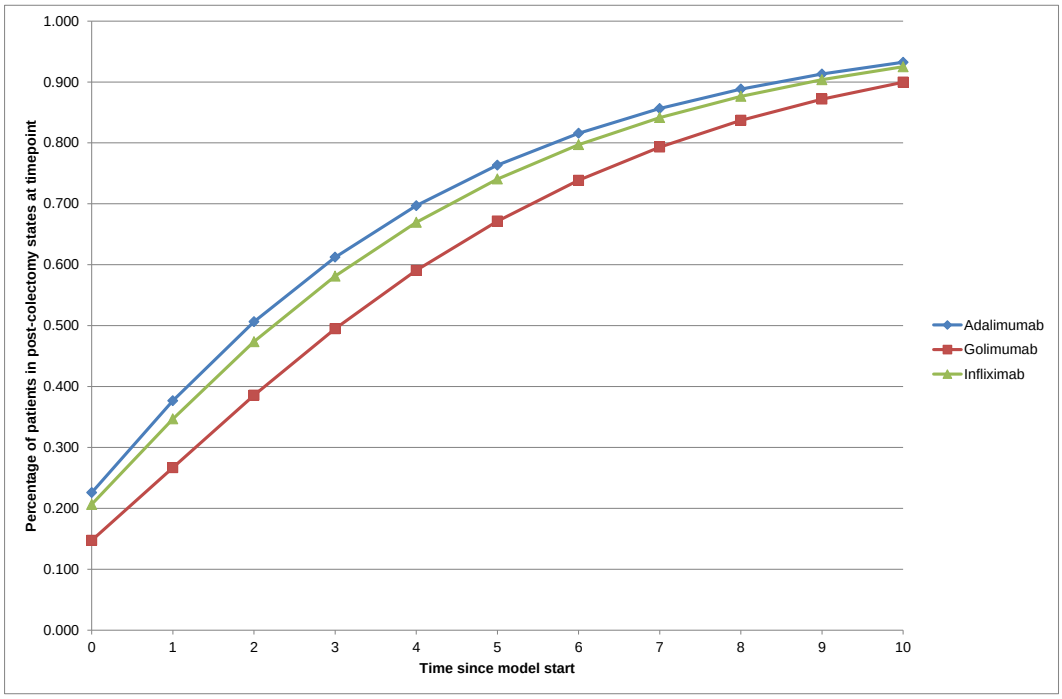
(2) Assumption that treatment failure is equivalent to severe exacerbation

Related to the issue regarding the lack of conventional drug therapies as the comparator for the economic analysis (see previous point above), the MSD economic models appear to also confuse the severity of the patient populations and the associated treatment pathway included in the model. Whilst the scope of the appraisal relates to patients with moderate to severe UC who have failed conventional treatment, the modelled pathway after failure of biologic therapy, and the choice of non-biologic comparators included in the analysis, appear to relate to a population with more severe disease and no

further medical treatment options are considered. The pathway represented by the model after failure of biologic therapy (or in its absence) appears to assume that failure to achieve an induction response, or that the loss of response during maintenance therapy, is synonymous with an acute UC exacerbation. As noted by the MSD submissions, surgery for UC is typically indicated for (i) patients with life-threatening complications (e.g. toxic megacolon or colonic perforation); (ii) dysplasia or proven cancer, or; (iii) severe disease characterised by treatment refractoriness, frequent flare-ups, extra-colonic manifestations, chronic corticosteroid dependence, side effects/intolerance/complications from medications (in particular corticosteroids), or according to clinical judgment.[65,67] After failing biologic treatment, the MSD model assumes that all patients who have failed biologic therapy will receive i.v. steroids and rapidly progress to colectomy (27% of all relapsing patients during each 2-month cycle). This fails to reflect the possibility that patients may continue to receive, and may still obtain clinical benefit from, non-biologic medical treatment options as defined in the model scope (5-ASAs, immunotherapies and/or steroids).

After removing mortality, the MSD model suggests that within 1-year (the approximate duration of the maintenance trials[46,48,50]), 15-20% patients are in the colectomy/post-colectomy health states (see Figure 84).

Figure 84: Proportion of patients in post-colectomy states over time (excluding mortality)



This contrasts with the colectomy rates observed within the RCTs included in the systematic review (see Chapter 5, 0.7% to 5.8% in in individual trial arms at ~1 year). The manufacturer's model also suggests that for patients receiving biologic treatments, 59%-70% will have undergone surgery within

5-years, and 89%-93% will have undergone surgery within 10-years (note - the precise values differ by biologic treatment group). These rates are very high and fail to reflect both the possibility of benefit from further medical therapies and the element of patient choice in deciding whether or not to undergo colectomy. If surgery really was the only remaining treatment option for these patients, it would not have been possible (or ethical) to undertake any of the trials included in this assessment.

Further to this point, the study reported by Turner *et al*,[127] which is used to inform the probability of requiring surgery for active UC, is a systematic review of studies describing the short-term outcome of adult and paediatric patients treated with i.v. corticosteroids, with or without ciclosporin, for *exacerbations* of UC. Within this published analysis, retrospective and prospective studies evaluating adult or paediatric UC patients admitted for first or subsequent exacerbation, who were severe enough to require i.v. corticosteroid therapy, were included if the short-term outcome and/or analysis of predictors of response were reported. This appears to confuse treatment failure with acute exacerbation of UC. The Assessment Group do not believe that either the narrow choice of remaining viable comparators or the treatment pathways assumed within the MSD models are representative of the clinical management of patients with moderate to severe UC in England and Wales.

(3) Questionable use of evidence concerning improving and worsening of inflammation

The NMA model uses separate models to produce information on the probability of sustained remission and the probability of sustained response. Within the health economic models, the odds ratios estimated using the NMA models are applied to the probability of remaining in the states of remission and response, respectively. The NMA logistic regression models treat these data as binomial – a patient either stays in their existing state or they do not. However, the data are multinomial – the observed data from the trials indicates that some patients who lost remission transited to response whilst others transited to no response, whilst some patients in response subsequently achieved remission, some achieved sustained response and some lost response. The structural assumptions employed within the transition matrix (see Table 42) do not reflect this, with some plausible transitions being assigned probabilities of zero. Whilst this problem is a likely consequence of the limitations of the published data from the ULTRA2 trial,[46] it poorly reflects the characteristics of the actual observed data.

(4) Questionable validity of use of methods for including non-randomised data from PURSUIT

The MSD submissions state that *"PURSUIT used a non-conventional trial design, and thus, conventional NMA techniques would not have sufficed for producing comparative effect estimates between golimumab, infliximab, and adalimumab. This NMA employed novel techniques of optimising the use of all available data."*[65,67] This approach was used to "downgrade" the available evidence for placebo within the PURSUIT maintenance trial as patients randomised to placebo were prior

golimumab induction responders. Based on the information provided in the manufacturer's submission (see MSD golimumab submission[67] Table 13 footnotes and text on pages 55-57), the Assessment Group was unable to logically follow or replicate the calculations used to generate hypothetical values for the placebo group. The Assessment Group does however believe that the manufacturer's "novel" method involves omitting the randomised data and instead uses a manipulation of the non-randomised placebo arm data as an input into the NMA. Such manipulation of observed trial data should be viewed with considerable caution. The Assessment Group believe that it would have been more appropriate to use more established methods of bias adjustment (for example, the methods adopted by Turner and Spiegelhalter[132]) and/or to use the published ITT data and examine the likely impact of the bias using sensitivity analyses.

(5) Lack of clarity regarding the NMA model

The NMA model is not reported in detail within either of the MSD submissions and the WinBUGS code was not reported (although this was provided to the Assessment Group during the clarification process). In addition, the baseline model is not described, although the health economic model indicates that baseline probabilities of achieving induction response/remission and maintaining response/remission were derived from *"Average study effect of placebo controlled trials in random effects NMA of induction response."* The appropriateness of these values is unclear.

(6) Inconsistencies between results of the MSD golimumab and infliximab models

The infliximab model and golimumab model are based on the same structure and the same decision problem. The results are however different between the models. In response to a request for clarification on the cause of this discrepancy, the manufacturer stated that the two models use different inputs for health utilities; the infliximab model uses utility data from the ACT1/2 trials whilst the golimumab model uses utility data from the PURSUIT trial. The infliximab model also uses different assumptions about the use of conventional non-biologic therapies than the golimumab model. The justification for using different utility and resource use assumptions in two models which are attempting to reflect exactly the same decision problem is inappropriate. It should also be noted that when the PURSUIT utility vector and resource use assumptions were inserted into the infliximab model, the results still did not coincide.

(7) Lack of clarity regarding the identification, selection and use of certain model parameters

In several instances, the justification for selecting particular parameter sources is unclear. In particular, the justification of the dosing and frequency of background therapies, and the justification for unit costs is not described within the MSD submissions.[65,67]

(8) Complex implementation of the model

Conceptually, the submitted MSD models are simple Markov models employing eleven health states and four treatment groups. However, the implementation of these models is complex; the model employs 30 worksheets, many of which were locked as read-only. This limited the ability of the Assessment Group to verify the inputs and formulae used in the model.

(9) Failure to undertake an incremental analysis

The MSD submissions do not include an incremental analysis in which each treatment option is compared against its next best non-dominated alternative. Instead, pairwise comparisons are made using NMB given a willingness to pay threshold of £30,000 per QALY gained. The infliximab submission states that: *"The ICER for infliximab versus standard non-biologic treatment (colectomy) is £37,682. The positive impact of infliximab in terms of reducing the Burden of Illness and mitigating the Wider Societal Impact of the condition represents additional value for consideration by the committee. Taking into account the shortfall in quality of life, and in the ability of people to contribute to society as a result of their experience with moderately to severely active UC, it is likely that infliximab represents a cost-effective treatment in first-line biologic treatment of UC."*[65] The golimumab submission states that *"At £27,322, the ICER for golimumab falls under a £30,000 threshold, and thus golimumab can be considered a cost-effective treatment option for patients with moderately to severely active UC."*[67]

Importantly, both of these economic conclusions are based on a comparison of biologic therapy versus immediate colectomy. A fully incremental analysis is presented in Tables 49 and 50. Given the ordering of QALY gains across all treatment options, infliximab should be compared against golimumab, thus resulting in a considerably higher ICER of approximately £75,000 to £80,000 per QALY gained (note the discussion around the discrepancy between model results above).

(10) Inclusion of a PAS for golimumab which has not yet been agreed by the Department of Health

Both MSD submissions include a PAS in which 100mg golimumab will be made available at the same price as 50mg golimumab (see MSD golimumab submission[67] page 8). However, at the time of this assessment, the proposed PAS had not been agreed with the Department of Health. Whilst the MSD submissions include a secondary analysis in which the PAS is not included, the absence of fully incremental comparisons by the manufacturer (as described in the previous point) clouds the correct interpretation of the economic analysis. The amended results of this fully incremental analysis, which excludes the PAS are shown in Tables 52 and 53.

Table 52: Model results from golimumab submission[67] (excluding PAS)

Treatment	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Probabilistic model results					
Infliximab	5.67	£44,122.45	0.21	£11,911.17	£56,268
Golimumab	5.50	£37,306.74	-	-	ext dom
Adalimumab	5.45	£32,211.28	0.53	£16,409.68	£30,724
Colectomy	4.92	£15,801.60	-	-	-
Results based on point estimates of parameters					
Infliximab	5.65	£43,091.60	0.20	£11,721.32	£57,980
Golimumab	5.50	£36,805.33	-	-	ext dom
Adalimumab	5.45	£31,370.28	0.50	£15,576.02	£31,069
Colectomy	4.95	£15,794.26	-	-	-

** ext dom – extendedly dominated*

Table 53: Model results from infliximab submission[65] (excluding PAS)

Treatment	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Probabilistic model results					
Infliximab	5.68	£44,126.44	0.22	£11,920.92	£53,258
Golimumab	5.51	£37,198.73	-	-	ext dom
Adalimumab	5.46	£32,205.53	0.54	£16,445.78	£30,428
Colectomy	4.92	£15,759.75	-	-	-
Results based on point estimates of parameters					
Infliximab	5.66	£42,919.73	0.21	£11,682.36	£55,5077
Golimumab	5.51	£36,663.51	-	-	ext dom
Adalimumab	5.45	£31,237.38	0.51	£15,447.78	£30,319
Colectomy	4.94	£15,789.59	-	-	-

** ext dom – extendedly dominated*

The exclusion of the PAS discount for 100mg golimumab results in a situation whereby adalimumab is no longer dominated and golimumab is ruled out due to extended dominance. Based on this version of the model, the ICER for adalimumab versus colectomy is approximately £30,000 per QALY gained. The ICER for infliximab versus adalimumab is at best £53,258 per QALY gained.

6.1.1.7 Adalimumab, golimumab and infliximab, for the treatment of ulcerative colitis (subacute) – AbbVie submission[63]

The AbbVie submission details the methods and results of a *de novo* health economic model developed to assess the cost-effectiveness of adalimumab versus "standard of care" (conventional non-biologic therapies) for the treatment of moderate to severe UC.

AbbVie model scope

The Abbvie model includes a comparison of two options: (1) adalimumab and (2) "standard of care" (standard non-biological therapies) for the treatment of moderate to severe UC from the perspective of

the UK NHS. The intervention arm (adalimumab plus standard non-biological therapies) begins with an induction dose of 160mg adalimumab at Week 0, followed by 80mg adalimumab at Week 2, and a maintenance dose of 40mg adalimumab every other week (EOW) starting from Week 4. At Week 8, those patients who achieve remission or response continue to receive adalimumab, whilst those patients who have lost response to the initial treatment can dose-escalate to 40mg adalimumab every week (EW). At Week 104, patients in the moderate-to-severe health state are assumed to discontinue adalimumab and subsequently receive conventional non-biologic treatment only. The comparator group within the model is comprised of conventional non-biologic drug treatments (anti-inflammatory drugs or immunosuppressants). Patients without response or remission in either treatment group can progress to colectomy at any time. Surgery is assumed to be reserved for patients who have failed both biologic and non-biologic drug treatments but is not evaluated as a treatment comparator in the model. Other biologic agents used for the treatment of UC (golimumab and infliximab) are not included in the AbbVie economic analysis.

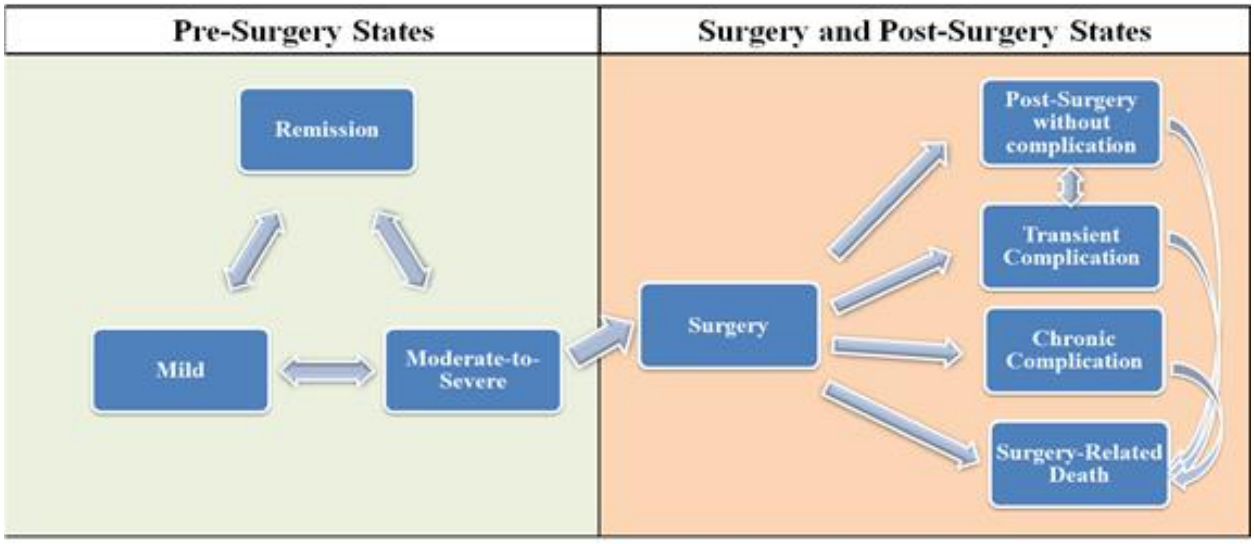
The model is evaluated as a cost-utility analysis whereby the primary health economic outcome is the incremental cost per QALY gained over a 10-year time horizon. The base case population considered relates to patients with moderate to severe UC who are have not previously been exposed to anti-TNF-α therapy and those who have previously been exposed to anti-TNF-α therapy (excluding adalimumab). Patients who are naïve to anti-TNF-α agents were evaluated as a secondary sensitivity analysis. Patients are assumed to have a mean body mass of 75kg. The starting age of patients entering the model is unclear in both the submission and the model. Costs and health outcomes are discounted at 3.5%. Costs were valued at 2013 prices.

AbbVie model structure

The model adopts a Markov approach using a 2-week cycle length (see Figure 85). The model includes a total of eleven health states: 3 pre-surgery states for adalimumab, 3 pre-surgery states for conventional treatments, one surgery state and 4 post-surgery states. These states are: (1) Mild [adalimumab]; (2) Remission [adalimumab]; (3) Moderate-to-severe [adalimumab]; (4) Mild [conventional treatment]; (5) Remission [conventional treatment]; (6) Moderate-to-severe [conventional treatment]; (7) Surgery; (8) Post-surgery without complications; (9) Transient complications; (10) Chronic complications, and; (11) Surgery-related death.

The three pre-surgery health states (remission, mild, and moderate-to-severe disease states) were defined using the Mayo Scoring system (or partial Mayo Scores if full Mayo Scores were not available).

Figure 85: AbbVie model structure[63]



The model is comprised of two treatment phases: (i) an induction phase and (ii) a maintenance phase. The induction phase relates to the first 8-weeks of treatment, in line with recommendations from the EMA.[133] For the adalimumab group, patients who are in the remission or mild disease states at this timepoint are assumed to continue to receive adalimumab into the maintenance period (8 weeks to 52 weeks). At the end of week 8, patients in the moderate to severe disease state are assumed to be non-responders to adalimumab; these patients discontinue treatment with adalimumab and subsequently receive conventional non-biologic therapy. Between week 8 and week 104, patients who have previously achieved remission or response but subsequently lost that response or remission are assumed to either discontinue adalimumab treatment or to dose-escalate to 40mg adalimumab EW. Within the conventional management group, patients transit between the conventional management health states without entering the biologic states.

For both the adalimumab and the conventional management groups, only patients in the moderate-to-severe health state are allowed to transit to surgery. Surgery is treated as a tunnel state, whereby patients can remain in that state for one cycle only. Patients can transit between the "transient complication" state and "post-surgery without complication" state during any cycle. Patients experiencing chronic complications are assumed to remain in that state until the time horizon has been exhausted. Patients undergoing surgery are assumed to be at an increased risk of death. Other-cause mortality is not included in the model. All patients enter the model in the pre-surgery moderate-to-severe state, in line with the inclusion criteria for the ULTRA-2 trial.[46] A half-cycle correction was applied to costs and QALYs. The main driver of health benefits within the model relates to HRQoL benefits associated with increased sojourn time in the pre-surgical health states.

Serious and severe adverse events were not considered in the AbbVie model; the manufacturer's submission notes that most adverse events experienced by patients in the ULTRA2 trial[46] were non-serious and considered to be unrelated to the study drugs.[63] In addition, the manufacturer highlights that the ULTRA2 trial reported slightly higher incidences of serious and severe adverse events in the placebo arm than in the adalimumab arm of the trial, therefore considering serious and severe adverse events in the model would have increased medical costs and reduced health gains within the conventional management group.[63] The exclusion of these events therefore represents a conservative assumption.

The model includes the costs associated with drug acquisition, medical costs related to disease states, hospitalisation, surgery, surgery-related complications, and costs associated with surgery-related death.

The model uses simple matrix multiplication to determine health state populations during each model cycle based on the state population in the previous Markov cycle and a series of time-dependent transition matrices. Costs and utilities are attached to each health state. Total QALYs are modelled as a function of sojourn time in each health state, together with an indirect survival benefit for adalimumab as a consequence of reduced rates of surgery (and hence surgical-related mortality) for this group.

Evidence used to inform the model parameters

A summary of evidence sources used to inform the main groups of parameters within the model is presented in Table 54.

Table 54: Summary table of evidence sources used to inform the AbbVie model parameters

Parameter group	Source
Transition probabilities – pre-surgical states	ULTRA2 trial[46,126] and ULTRA1/2 extension study[63] and other literature[120,134] with the cycle length of matrix probabilities adjusted using Eigen decomposition
Transition probabilities – rate of surgery	Hillson *et al*[135]
Transition probabilities – post-surgery complications and surgery related-mortality	Transition complications rates estimated from Swenson *et al.*[114] Chronic complication rates estimated using studies by Johnson *et al*[136] (fertility), Kruasz *et al*[137] (male impotence) and Abdelrazep *et al*[138] (chronic pouchitis). Peri-operative and post-operative mortality risks were estimated using a study reported by Roberts *et al*.[139]
Health utilities for pre-colectomy response / remission	EQ-5D study published as a poster by Swinburn *et al*[140]
Health utilities for post-colectomy states	Utility values for post-surgery without complication and transient complication based on estimates reported by Tsai *et al.*[82] Utility values for chronic complications based on Arseneau *et al,*[110] Hu *et al*[141] and Smith *et al.*[142]
Resource use	Adalimumab dosing and dose escalation based on SPC and experience within the ULTRA2 trial.[46,126] Use of conventional non-biologic treatments was based on UC-related medication usage rates for all subjects at baseline, as observed in the ULTRA 2trial.[46,126] The disease state resource use were based on the estimates reported by Tsai *et al.*[82] Rates of hospitalisation were based on a mixed effects regression analysis of ULTRA1 and ULTRA2 trial data.[63]
Unit costs	Drug acquisition costs (biologics and conventional treatments) were taken from the Monthly Index of Medical Specialities (MIMS). Hospitalisation costs were based on NHS Reference Costs[131] Other unit costs derived from literature[82,114,143]

Methods for modelling effectiveness

In the main, estimates of baseline and relative effectiveness were taken from the ULTRA2 study and the ULTRA1/2 extension study,[46,63,126] although other literature was used to inform transitions that were not observed within these studies.[120,134] Efficacy data on response/remission from the ULTRA1 trial were not used in the AbbVie model. Transition probabilities between pre-surgery health states were calculated using trial data from ULTRA2[46,126] for weeks 8-104 whilst transitions between states for cycles between weeks 104 to 520 were based on data from and then the ULTRA1/2 extension study[63] for adalimumab and ULTRA2[46,126] for conventional management. Discontinuations due to other reasons, such as adverse events (AEs), were also considered based on trial data.

Four matrices of time-dependent transition probabilities are used within the AbbVie model, according to four time intervals; these are described below.

Transition probabilities – adalimumab group

Period 1 (weeks 0-8): In the induction period, transitions from the moderate-to-severe state were based on the adalimumab group of the ULTRA2 trial.[46,126] As ULTRA2 did not recruit patients with prior response or remission (because it was an induction trial), this study cannot provide information relating to transitions from these states to other states within the first 8-week period. Instead, the probabilities of maintaining remission and response were based on studies reported by Kane *et al*[134] (assuming the probability of maintaining remission reflects that of "adherent patients") and Odes *et al.*[120] A constant hazard was assumed to obtain the 8-week probability in both cases.

Period 2 (weeks 8-52): Transition probabilities were based on a cross tabulation of data on the number of patients in each health state from the adalimumab arm of the ULTRA2 trial.[46,126]

Period 3 (weeks 52-104): Data from the ULTRA1/2 extension study[63] were used to derive transition probabilities for the three pre-surgery health states. As patients in the moderate-to-severe state within the adalimumab group of the model are assumed to discontinue biologic treatment, only those patients who were randomised to the adalimumab arm and who had remission or mild disease at Week 8 in the ULTRA1/2 extension study[63] were included in the analysis.

Period 4 (weeks 104-260): Data from Week 48 to Week 144 of the ULTRA1/2 extension study[63] were used to generate the transition matrix. A multinomial logit regression model was constructed to estimate the transition matrix during each 48-week interval. The dependent variables were the three pre-surgery health states and the independent variables were the health states in the previous visit. The logit model estimates mean predicted probabilities of being in one of the health states given a specific health state at the previous visit.

These four transition matrices were then converted to 2-week probabilities using Eigen matrix decomposition methods reported by Craig and Sendi.[144] The resulting matrices are shown in Table 55.

Table 55: Adalimumab group - 2-week transition probabilities for pre-surgery states by time interval

From state	To state			
	Remission	Mild	Moderate-to-severe	Surgery
From Week 0 to 8				
Remission	0.9974	0.0007	0.0019	-
Mild	0.0003	0.9981	0.0016	-
Moderate-to-severe	0.0551	0.0986	0.8432	0.0031
From Week 8 to 52				
Remission	0.9700	0.0164	0.0136	-
Mild	0.0349	0.9400	0.0251	-
Moderate-to-severe	0.0001	0.0215	0.9753	0.0031
From Week 52 to 104				
Remission	0.9889	0.0000	0.0111	-
Mild	0.0178	0.9436	0.0385	-
Moderate-to-severe	0.0275	0.0217	0.9477	0.0031
Week 104 onward				
Remission	0.9949	0.0047	0.0004[1]	-
Mild	0.0113	0.9869	0.0018[1]	-
Moderate-to-severe	0.0037	0.0019	0.9463[1]	0.0031

[1]4.54% of patients reaching the moderate-to-severe disease state after week 104 discontinue ADA treatment and subsequently receive conventional treatment

Transition probabilities – conventional management group

Period 1 (weeks 0-8): In the induction period, transitions from the moderate to severe state were based on the placebo group outcomes within the ULTRA2 trial.[46] As with the adalimumab matrix for the induction period, estimates of maintaining remission and response were based on studies reported by Kane *et al*[134] (assuming the probability of maintaining remission reflects that of "non-adherent patients") and Odes *et al.*[120] A constant hazard was assumed to obtain the 8-week probability in both cases.

Period 2 (weeks 8-52): Transition probabilities were based on a cross tabulation of data on the number of patients in each health state from the placebo arm of the ULTRA2 trial.[46,126]

Periods 3 and 4 (weeks 52-260): Transition probabilities for each cycle were assumed to reflect those estimated for Period 2 (weeks 8-52).

As with the adalimumab group, these four transition matrices were then converted to 2-week probabilities using Eigen matrix decomposition methods reported by Craig and Sendi.[22] The resulting matrices are shown in Table 56.

Table 56: Conventional management group - 2-week transition probabilities for pre-surgery states by time interval

From state	To state			
	Remission	Mild	Moderate-to-severe	Surgery
From Week 0 to 8				
Remission	0.9799	0.0044	0.0157	-
Mild	0.0013	0.9844	0.0143	-
Moderate-to-severe	0.0291	0.0882	0.8796	0.0031
From Week 8 to 52 (and subsequent cycles)				
Remission	0.9696	0.0028	0.0276	-
Mild	0.0170	0.9217	0.0613	-
Moderate-to-severe	0.0017	0.0074	0.9878	0.0031

Transitions between surgery and post-surgical states

Transitions between surgery and post-surgical health states were based on the literature rather than the clinical studies of adalimumab. Transition complications rates were estimated from a study reported by Swenson *et al.*[114] Chronic complication rates were estimated using studies by Johnson *et al*[136] (fertility), Kruasz *et al*[137] (male impotence) and Abdelrazep *et al*[138] (chronic pouchitis). Peri-operative and post-operative mortality risks were estimated using a study reported by Roberts *et al*[139] taking account of background mortality rates. The underlying transition rates are assumed to be time-independent. The resulting matrix is shown in Table 57.

Table 57: Transition matrix for surgery and post-surgical states (all time periods, both treatment groups)

From state / To state	Surgery	Post-surgery without complication	Transient complication	Chronic complication	Death
Surgery	0.0000	0.7708	0.0101	0.1919	0.0272
Post-surgery without complication	0.0000	0.9893	0.0101	0.0000	0.0006
Transient complication	0.0000	0.9994	0.0000	0.0000	0.0006
Chronic complication	0.0000	0.0000	0.0000	0.9994	0.0006
Death	0.0000	0.0000	0.0000	0.0000	1.0000

Health-related quality of life (HRQoL)

Table 58 reports the HRQoL values used in the AbbVie model and their sources.

Table 58: Health utilities assumed in the AbbVie model[63]

Disease State	Utility	Source
Remission	0.91	Swinburn *et al*[140]
Mild	0.80	Swinburn *et al*[140]
Moderate-to-severe	0.55	Swinburn *et al*[140]
Surgery	0.55	Assumed to same as moderate-to-severe state
Post-surgery without complication	0.61	Tsai *et al*[82]
Transient complication	0.55	Tsai *et al*[82]
Chronic complication	0.43	Weighted mean of Arseneau *et al,*[110] Hu *et al*[141] and Smith[142]

The AbbVie submission argues that whilst it would have been possible to map SF-6D utility estimates from the ULTRA2 trial onto the EQ-5D, this is likely to overestimate the level of HRQoL of patients with more severe disease. Health utilities for the pre-surgery states were instead sourced from an EQ-5D study of 230 patients with UC reported by Swinburn *et al.*[140] This study has only been published in abstract form only, however further details are provided in Appendix 3 of the AbbVie submission.[63] Utility values for the states of post-surgery without complications and post-surgery with transient complications were taken from Tsai *et al.*[82] The utility for the chronic complication state was estimated by using a weighted value of rates and HRQoL impacts of chronic pouchitis (Arseneau *et al*[110]), infertility (Hu *et al*[141] and male sexual dysfunction (Smith and Roberts[142]).

Resource use and costs

Table 59 summarises the values of the resource use and cost parameters used in the AbbVie model.

Table 59: Drug resource cost parameters used in the AbbVie model

Parameters*	Values	Sources
Adalimumab dose escalation - relative dose intensity compared to 40mg EOW		
Week 8 - 52 maintenance phase	7.40%	Primary analysis of ULTRA2[46] and ULTRA1/2 extension study[63]
Week 52 – 104 maintenance phase	24.06%	
Beyond Week 104 maintenance phase	21.49%	
Use of conventional therapies		
Mesalazine	47.0%	Based on baseline usage in ULTRA2[46,126]
Sulfasalazine	7.3%	
Balsalazide	5.9%	
Olsalazine	0.2%	
Azathioprine	28.3%	
Mercaptopurine	6.7%	
Drug acquisition costs		
Adalimumab unit price (40mg)	£352.14	MIMS (March 2014)
Mesalazine	£20.59	MIMS (March 2014)
Sulfasalazine	£2.93	
Balsalazide	£13.10	
Olsalazine	£9.88	
Azathioprine	£2.89	
Mercaptopurine	£105.99	
Total weighted conventional therapy cost per 2-week cycle	£18.60	

** Assumptions regarding specific products, doses, frequency and price are not clear from the AbbVie submission*

Drug acquisition costs (adalimumab and conventional management)

Usage of adalimumab was based on its licensed indication[133] together with estimates of relative dose intensity for dose escalating patients based on the primary analysis of data from the ULTRA2 trial[46] and the ULTRA1/2 extension study.[63] The use of conventional non-biologic therapies was assumed to reflect the baseline usage of these therapies within the ULTRA2 trial.[46,126] The drug acquisition costs for adalimumab and conventional non-biologic therapies were obtained from the MIMS database (accessed March 2014).

Health state resource costs

Other UC health state costs assumed in the AbbVie model are summarised in Table 60.

Table 60: Other health state costs used in the AbbVie model

Parameters	Values	Sources
Hospitalisations per 2-week cycle		
Remission – adalimumab	0.0008	Mixed effects regression analysis of ULTRA1 and ULTRA2 trial data[45,125,126]
Mild – adalimumab	0.0013	
Moderate-to-severe – adalimumab	0.0042	
Remission – conventional management	0.0017	
Mild – conventional management	0.0029	
Moderate-to-severe – conventional management	0.0094	
Hospitalisation costs		
Cost per hospitalisation	£3,533	NHS Reference Costs 2012/13[131] (Major Gastrointestinal Disorders with CC Score 0, elective inpatient, PA25B)
Pre-surgery disease state costs (excluding hospitalisation) per 2-week cycle		
Remission	£20.31	Derived using Tsai *et al.*[82] Includes blood tests, consultation visits, and endoscopies.
Mild	£67.87	
Moderate-to-severe	£203.27	
Post-surgery disease state costs per 2-week cycle		
Surgery	£13,071	Based on Buchanan *et al*[143] and inflated to 2013 prices.
Post-surgery without complication	£118.63	Derived using Tsai *et al*[82] including blood tests, consultation visits, and endoscopies.
		No hospitalisations were considered for post-surgery without complication.
Transient complication	£8,826.05	Based on Swenson *et al*[114], inflated and exchange-rate adjusted to 2013 prices.
Chronic complication	£118.63	Assumed to be the same as post-surgery without complication
Terminal care	£3,533	NHS reference costs 2012/13[131] (Major Gastrointestinal Disorders with CC Score 0, elective inpatient, PA25B)

The frequency of hospitalisations per 2 week cycle was estimated according to treatment arm and disease severity using mixed effects regression on pooled data from the ULTRA1 and ULTRA2 trials.[45,125,126] Other disease state resource use (consultant visits, blood tests and emergency/elective endoscopies) were taken from Tsai *et al*[82] and uplifted to current prices.[129] Hospitalisation and post-surgery terminal care costs were obtained from NHS Reference Costs 2012/13.[131] The costs of surgery and managing complications were taken from Buchanan *et al*[143] and Swenson *et al.*[114]

Model evaluation and uncertainty analysis

The results of the AbbVie economic analysis are presented as an incremental cost-effectiveness ratio; this is based on the point estimates of parameters rather than the expectation of the mean. Uncertainty surrounding incremental costs and outcomes was examined using deterministic sensitivity analyses and PSA. The PSA was undertaken over 1,000 Monte Carlo samples. The results of the deterministic analyses are presented as tornado diagrams whilst the results of the PSA are presented as cost-effectiveness planes and CEACs.

AbbVie model results

Tables 61 and 62 present the results of the AbbVie model for the base case analysis and the secondary analysis of the subgroup of patients who are anti-TNF naïve. Note that the probabilistic ICERs presented in these tables have been generated by the Assessment Group.

Table 61: Model results obtained from the AbbVie model – base case analysis

Treatment	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
*Probabilistic model results**					
Adalimumab	*Treatment-specific costs and QALYs not stored in PSA sub-routine*		0.73	£25,335	£34,590
Conventional management			-	-	
Results based on point estimates of parameters					
Adalimumab	5.73	£76,392	0.74	£25,446	£34,417
Conventional management	4.99	£50,946	-	-	

** generated by the Assessment Group*

Table 62: Model results obtained from the AbbVie model – anti-TNF-α naïve subgroup

Treatment	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
*Probabilistic model results**					
Adalimumab	*Subgroup model does not allow for PSA*				
Conventional management					
Results based on point estimates of parameters					
Adalimumab	6.00	£79,799	0.87	£31,140	£35,970
Conventional management	5.140	£48,659	-	-	

** generated by the Assessment Group*

The base case analysis of the model indicates that over a 10-year time horizon adalimumab is expected to generate an additional 0.73 QALYs at an incremental cost of £25,335 per patient. This leads to an ICER of £34,590 per QALY gained. The results of the model based on the point estimates of parameters are very similar to those produced using the probabilistic model.

The deterministic subgroup analysis of anti-TNF naïve patients indicates that over a 10-year time horizon adalimumab is expected to generate an additional 0.87 QALYs at an incremental cost of £31,140 per patient. This leads to an ICER of £35,970 per QALY gained. It was not possible to generate probabilistic estimates for the subgroup analysis as the subgroup model does not include a PSA sub-routine.

Figures 86 and 87 presents the cost-effectiveness plane and CEACs for the base case analysis.

Figure 86: Cost-effectiveness plane reported by AbbVie – base case analysis[63]

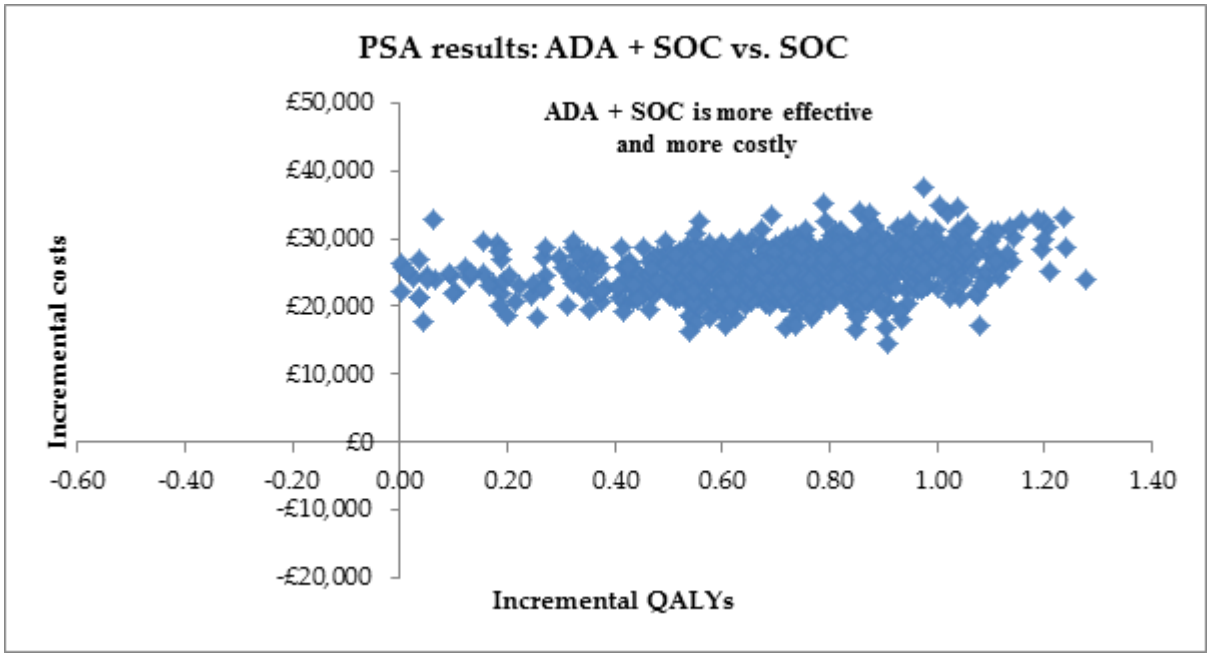
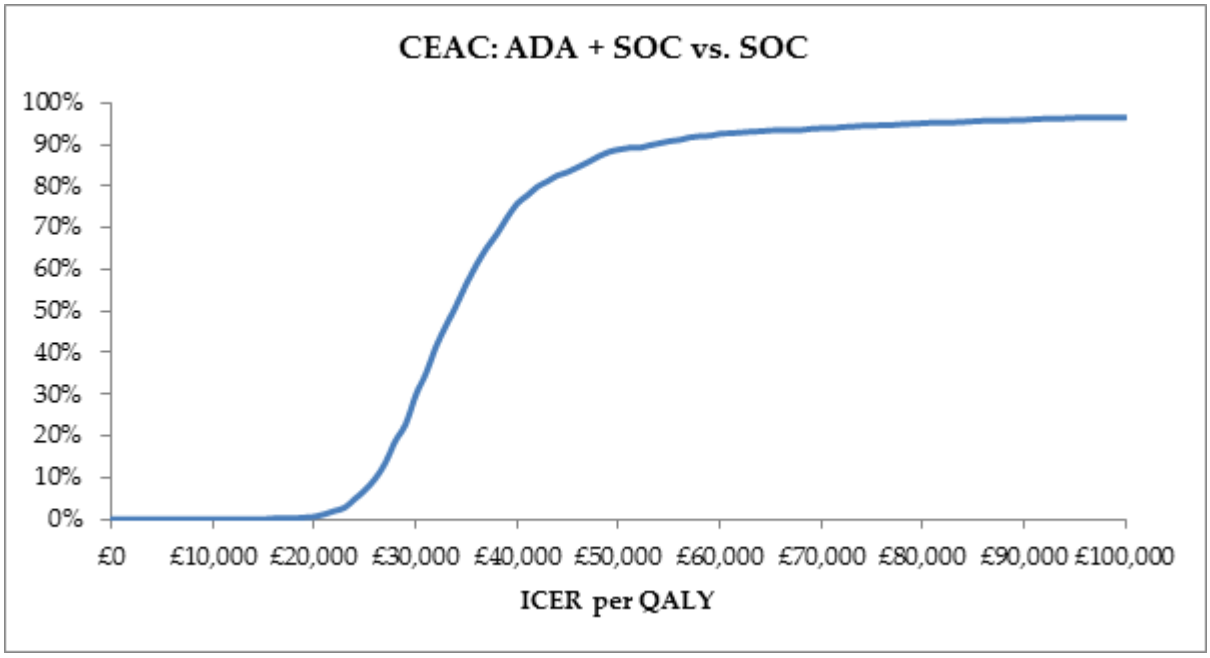


Figure 87: CEAC reported by AbbVie – base case analysis[63]



The cost-effectiveness plane indicates that adalimumab is consistently expected to be more effective and more expensive than conventional management. Assuming a willingness to pay threshold of £20,000 per QALY gained, the probability that adalimumab produces more net benefit than conventional management is approximately 0.01. Assuming a willingness to pay threshold of £30,000 per QALY gained, the probability that adalimumab produces more net benefit than conventional management is approximately 0.30.

The DSA undertaken by the manufacturer indicate that the model is most sensitive to assumptions concerning disease state costs and the health state utilities. Given the narrow scope of the AbbVie economic analysis, the cost-effectiveness of adalimumab compared against other biologic therapies or surgery is unknown.

Critical appraisal of the AbbVie model

The main issues identified by the Assessment Group are presented in Box 3 and discussed below.

Box 3: Main problems and concerns relating to the AbbVie model

1. Deviations from the NICE Reference Case and final NICE scope, particularly with respect to omission of other biologics and surgery as comparators
2. Questionable choice of cycle length necessitating the use of other external evidence on transition probabilities which should not be required
3. Concerns regarding the selection and use of evidence to inform HRQoL parameters
4. Questionable source of surgery rate

(1) Deviations from NICE Reference Case and final NICE scope

The extent to which the economic analyses reported in the AbbVie submissions adheres to the NICE Reference Case is presented in Table 63.

Table 63: Adherence of the AbbVie model to the NICE Reference Case[78]

Element of health technology assessment	Reference case	Assessment Group comments
Defining the decision problem	The scope developed by the Institute	The scope of the analysis deviates from the final scope from NICE.
Comparator(s)	As listed in the scope developed by NICE	The comparator is limited to "standard of care" (conventional non-biologic therapies) only. The model does not include other biologic agents (infliximab and golimumab) included in the final NICE scope. The model does not include surgery as a comparator.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health outcomes reflect those of patients with UC
Perspective on costs	NHS and PSS	The economic analysis was undertaken from the perspective of the UK NHS. PSS costs are not mentioned in the submission.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The economic analysis takes the form of a cost-utility analysis of the two included options.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Costs and outcomes are evaluated over a 10-year time horizon. Analyses over a lifetime horizon are not presented in the manufacturer's submission nor are they possible within the implemented model structure.
Synthesis of evidence on health effects	Based on systematic review	Whilst the submission mentions other relevant trials of infliximab and golimumab, the manufacturer opted to undertake a "within-trial" analysis of adalimumab versus conventional management using efficacy data from the ULTRA2 trial[46,126] only.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults	Health effects are assessed in terms of QALYs. The EQ-5D has been used to assign specific utility values for health states, but weighted averages from other instruments (i.e. TTO) have also been used to value the post-surgery chronic complications health state.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	It would have been possible to map from the SF-6D in the ULTRA2 trial[46,126] to the EQ-5D. Instead, the manufacturer used data from Swinburn *et al*[140] to value pre-surgery states.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of	No equity weighting is applied.

Element of health technology assessment	Reference case	Assessment Group comments
	the other characteristics of the individuals receiving the health benefit	
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The economic analysis was undertaken from the perspective of the UK NHS.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	The model uses a discount rate of 3.5% for costs and health outcomes.

Overall, the economic analysis undertaken by AbbVie is generally in line with the NICE Reference Case. However, similar to the MSD submissions,[65,67] the two most notable concerns relate to the choice of comparators and the adoption of a short model time horizon (10-years).

The AbbVie economic analysis includes only two treatment options: (1) adalimumab and (2) conventional non-biologic treatments. The analysis excludes other relevant biologic therapies for the treatment of UC (infliximab and golimumab) and elective surgery. The appendix to the main submission states that: *"Other anti-TNF therapies which are being appraised as part of this NICE MTA, namely infliximab and golimumab, were not considered as comparators in the present evaluation as they are not NICE recommended for this patient population and therefore would not form routine standard of care at present."*[63] However, infliximab and golimumab were listed in the final NICE scope, hence they should have been included in the economic analysis. As a consequence of their omission, the AbbVie model adopts a very narrow scope and provides no information regarding the comparative cost-effectiveness of the full range of biologic treatment options within this appraisal.

The main submission from AbbVie states a number of arguments regarding why it would not be appropriate to undertake a formal NMA (see AbbVie submission[63] pages 67-68). The main arguments stated are:

(a) Differences in Mayo score estimation between the relevant trials;
(b) Placebo responses have been shown to differ markedly depending on the severity of the trial population, study design and country or region in which the trial was conducted;
(c) Other differences in trial design i.e. the use of adaptive design in the PURSUIT trial,[48] differences in timepoints for the assessment of induction response, eligibility criteria relating to prior treatment failures, prior use of biologics, steroid tapering, open label escape allowance, timing of efficacy assessments and study durations.

However the Assessment Group do not agree that a NMA is inappropriate and AbbVie's justifications for not undertaking such an analysis appear to be flawed. Notably, UC is a chronic disease characterised by ongoing inflammation over time; fluctuations in Mayo score evaluations over the course of 3 days are likely to be minor, hence the use of alternative scoring systems between trials are unlikely to produce any substantial bias. Furthermore, no two trials are identical; whilst it is useful to highlight potential sources of heterogeneity between studies (and this is done well by AbbVie), the Assessment Group does not believe that the presence of this heterogeneity provides a sufficient basis for ignoring treatment options relevant to the decision problem.

In addition, the AbbVie model explicitly excludes elective colectomy as a comparator from the analysis. The Appendix to the main AbbVie submission states that *"Surgery is an important treatment option in UC clinical management and is reserved for patients who have an inadequate response with, are contraindicated to or intolerant of conventional standard of care. Surgery is unlikely to be a first line option for moderately to severely active UC patients. Consistent with this approach, surgery is included in the model as the treatment option for a proportion of patients who failed SOC or ADA+SOC treatment, but not as a comparator to ADA+SOC.*"[63] Since this option was specified in the final agreed NICE scope, and because the appraisal does not relate to first-line treatment, it should have been included in the economic analysis.

It should also be noted that the AbbVie model time horizon is constrained to 10-years (260 2-week cycles). This shorter time horizon is used as a justification for excluding other cause-mortality from the model. The model does not include the functionality to consider longer time horizons; it is unclear whether the profiles of incremental costs and health outcomes for adalimumab versus conventional management would be similar over longer time horizons.

(2) Questionable choice of cycle length necessitating the use of other evidence on transition probabilities

The cycle length adopted within the AbbVie model is 2 weeks. This short cycle length was selected *"to accommodate the ADA dosing schedule."*[63] Given that all patients enter the model in the "moderate-to-severe" health state in line with the ULTRA2 trial, this choice of cycle length leads to a necessity to incorporate other literature[120,134] to populate the transition probabilities from the "Mild" and "Remission" health states to other health states. As a consequence, there is some discrepancy between the observed pre-surgery health state distribution following induction in the ULTRA2 trial[46,126] and the pre-surgery health state distribution following induction estimated by the model (see Table 64). Given a longer cycle length for induction i.e. the 6 weeks used in the trial, it would have been unnecessary to include other data on transition probabilities and the predictions of the model would have likely been more accurate.

Table 64: Comparison of observed and predicted induction outcomes

Treatment group	No response	Response	Remission
Adalimumab group (observed)	0.52	0.33	0.16
Adalimumab group (predicted)	0.51	0.31	0.17
Discrepancy (observed – predicted)	0.01	0.02	-0.01
Placebo group (observed)	0.67	0.24	0.09
Placebo group (predicted)	0.61	0.29	0.09
Discrepancy (observed – predicted)	0.07	-0.05	-0.01

(3) Concerns regarding the selection and use of evidence to inform HRQoL parameters

The ULTRA2 trial[46,126] did not collect HRQoL data from patients using the EQ-5D; however, the SF-36 instrument was included and could be used to derive SF-6D utility values. The manufacturer explored mapping the SF-6D values to the EQ-5D but noted that this would likely overestimate the level of HRQoL of patients with more severe disease. Instead, the manufacturer used data from Swinburn *et al*[140] to value the pre-surgery health states in the model. Whilst the Swinburn *et al*[140] study has been published only in abstract and poster form, more detail is provided in Appendix 3 of the main AbbVie submission.[63] It is noteworthy that the difference in utility for the post-surgery state and the active UC state in the selected utility values within the AbbVie model (0.61-0.55=0.06) is smaller than that observed within other EQ-5D UC valuation studies (e.g. Woehl *et al*[109] estimated this difference to be ~0.71– 0.41= 0.30). The AbbVie model therefore does not assume that surgery results in a substantial increase in HRQoL in patients with active disease.

It is also noteworthy that the choices made with respect to the HRQoL values for other post-surgery health states are not clear from the submission. In particular, the methods for identifying and selecting studies to value the chronic complications,[110,141,142] and the weightings given to each, are unclear from the AbbVie submission.[63] What is clear is that the three valuation studies used to inform the chronic complications utility values used different health instruments; Hu *et al*[141] is based on Committee valuations using the Health Utility Index, Arseneau *et al*[110] reported TTO valuations by UC patients and Smith *et al*[142] report TTO and VAS valuations. Producing a weighted mean utility from studies which use different elicitation methods may produce conceptually inconsistent rankings of identical health states. This parameter does not however have a material impact upon the ICER.

(4) Questionable source of surgery rate

The AbbVie model estimates the 2-week probability of undergoing surgery from a 1-year study reported by Hillson *et al*.[135] This study was a retrospective analysis of medical claims with and without UC identified from a population of approximately 500,000 employees, retirees, and dependent in the US. This does not specifically relate to a moderate to severe UC population and the use of a 1-year study to estimate long-term risk is concerning, particularly given the availability of other longer studies undertaken in more relevant UC populations (see Section 6.2.2.3).

6.2 *De novo* Assessment Group model

6.2.1 Introduction

In light of the limitations of the models submitted by the manufacturers (see Section 6.1), the Assessment Group developed a *de novo* health economic model to assess the cost-effectiveness of second-line infliximab, adalimumab and golimumab, conventional non-biologic therapies and immediate colectomy for the treatment of patients with moderate to severe UC.

6.2.2 Methods

6.2.2.1 Model scope

The scope of the economic analysis follows the NICE Reference Case (summarised in Box 4).

Box 4: Scope of the Assessment Group economic analysis

Population:

Patients with moderate-to-severe UC who have failed at least one prior therapy*

Interventions and comparators: adalimumab 160mg/80mg/40mg; infliximab 5mg/kg; golimumab 200mg/100mg/100mg(50mg); conventional non-biologic therapy (comprised of a mix of 5-ASAs, immunosuppressants and corticosteroids); elective surgery

Economic outcome: Incremental cost-per QALY gained

Perspective: NHS and PSS

Time horizon: Lifetime

Discount rate: 3.5%

** The base case analysis relates to an adult UC population; a secondary analysis is considered for the paediatric population.*

The analysis compares infliximab, adalimumab and golimumab against each other and against conventional non-biologic therapy (comprised of a mix of 5-ASAs, immunosuppressants and corticosteroids) and immediate colectomy. Infliximab is assumed to be given at a dose of 5mg/kg on three visits during induction and subsequently at a dose of 5mg/kg every 8 weeks for patients who go on to receive maintenance therapy. Adalimumab is assumed to be given at one dose of 160mg, one dose of 80mg and two doses of 40mg during the induction phase; a dose of 40mg EOW is assumed for patients who go on to receive maintenance therapy. A fixed proportion of adalimumab patients (27%) are assumed to escalate to a 40mg EW dosing regimen, based on data reported in the AbbVie submission.[63] Golimumab is assumed to be given as one dose of 200mg and one dose of 100mg during induction treatment, with subsequent maintenance therapy given at a dose of 100mg every 4 weeks for patients with body mass greater than or equal to 80kg or 50mg every 4 weeks for patients with body mass less than 80kg. Infliximab is assumed to be administered in a day case setting whilst the administration of golimumab and adalimumab is not assumed to require any additional NHS

resources (no costs are included for training patients to self-inject). Patients in the non-surgical treatment groups are assumed to receive conventional background therapies (5-ASAs, immunosuppressants and corticosteroids). Surgery is included in the economic analysis both as a comparator within the analysis and also as a downstream component of the pathway for patients in the biologic and non-biologic treatment groups.

The population within the economic analysis relates specifically to patients with moderate-to-severe UC who have failed at least one prior therapy, as reflected in the RCTs included in the systematic review of clinical effectiveness (see Chapter 5). Patient characteristics are based on the trials included in the systematic review.[45,48-50,125,126] Patients are assumed to enter the model aged 40 years with a mean body mass of 77kg. Thirty two percent of patients are assumed to have a body mass greater than 80kg. Whilst the main economic analysis relates to adult patients with UC, a scenario analysis is also presented which compares infliximab against conventional drug treatment and immediate colectomy in paediatric patients with UC (note that golimumab and adalimumab do not currently have marketing authorisations in paediatric patients[67,133]). This secondary analysis should be considered exploratory as the efficacy data are drawn from trials undertaken within an adult UC population. The economic evaluation takes the form of a cost-utility analysis; the primary health economic outcome is the incremental cost per QALY gained. All treatment options are evaluated within a fully incremental analysis within the base case. The perspective of the economic analysis relates to that of the NHS and PSS. All costs and outcomes discounted at 3.5%. Costs and health outcomes are evaluated over a lifetime horizon in the base case; shorter time horizons are considered as secondary scenario analyses.

6.2.2.2 Model structure

The Assessment Group model adopts a Markov structure with eight mutually exclusive health states (see Figure 88). The model health states are defined according to whether the patient is alive or dead, the non-surgical treatment the patient is currently receiving (biologic therapy or non-biologic therapy), their prior history of colectomy and their current level of disease control (remission, response and active UC). Remission and response to treatment are classified according to the Mayo score, as defined within the trials included in the systematic review (see Chapter 5). Remission is defined as a Mayo score ≤2 with no individual subscore >1. Response is defined as a decrease from baseline in the total Mayo score of at least 3 points and at least 30 percent, with an accompanying decrease in the subscore for rectal bleeding of at least 1 point or an absolute subscore for rectal bleeding of 0 or 1. As remission is a subset of the broader category of response, these are dealt with a mutually exclusive ordered categorical data (see Section 5.2.3.4). Patients without either response or remission are defined as having active (moderate-to-severe) UC. The model includes the following health states: (1) on biologic treatment – active UC; (2) on biologic treatment – response; (3) on biologic treatment – remission; (4) on conventional treatment – active UC; (5) on conventional

treatment – response; (6) on conventional treatment – remission; (7) post-surgery (with or without complications), and; (8) dead. Surgery is not included as a state but rather it is incorporated as an event; patients undergoing colectomy are assumed to transit to the post-surgery state if they survive their surgery and the dead state if they do not.

The model time horizon is divided into two main phases: (i) induction and (ii) maintenance. The model adopts an 8-week cycle length for the induction phase and a 26-week cycle length during the subsequent maintenance phase. During the induction phase, patients receiving biologic treatment who achieve response or remission are assumed to continue receiving the same biologic treatment as maintenance therapy. Patients who do not achieve response or remission during biologic induction therapy are assumed to discontinue that biologic and subsequently receive conventional non-biologic treatments. Patients in the conventional treatment group are assumed to continue receiving conventional therapy irrespective of their response to induction therapy. Patients in the immediate colectomy group are assumed to undergo surgery during the induction phase of the model and subsequently remain in the post-surgery state. All patients have a probability of dying from other causes during the induction cycle.

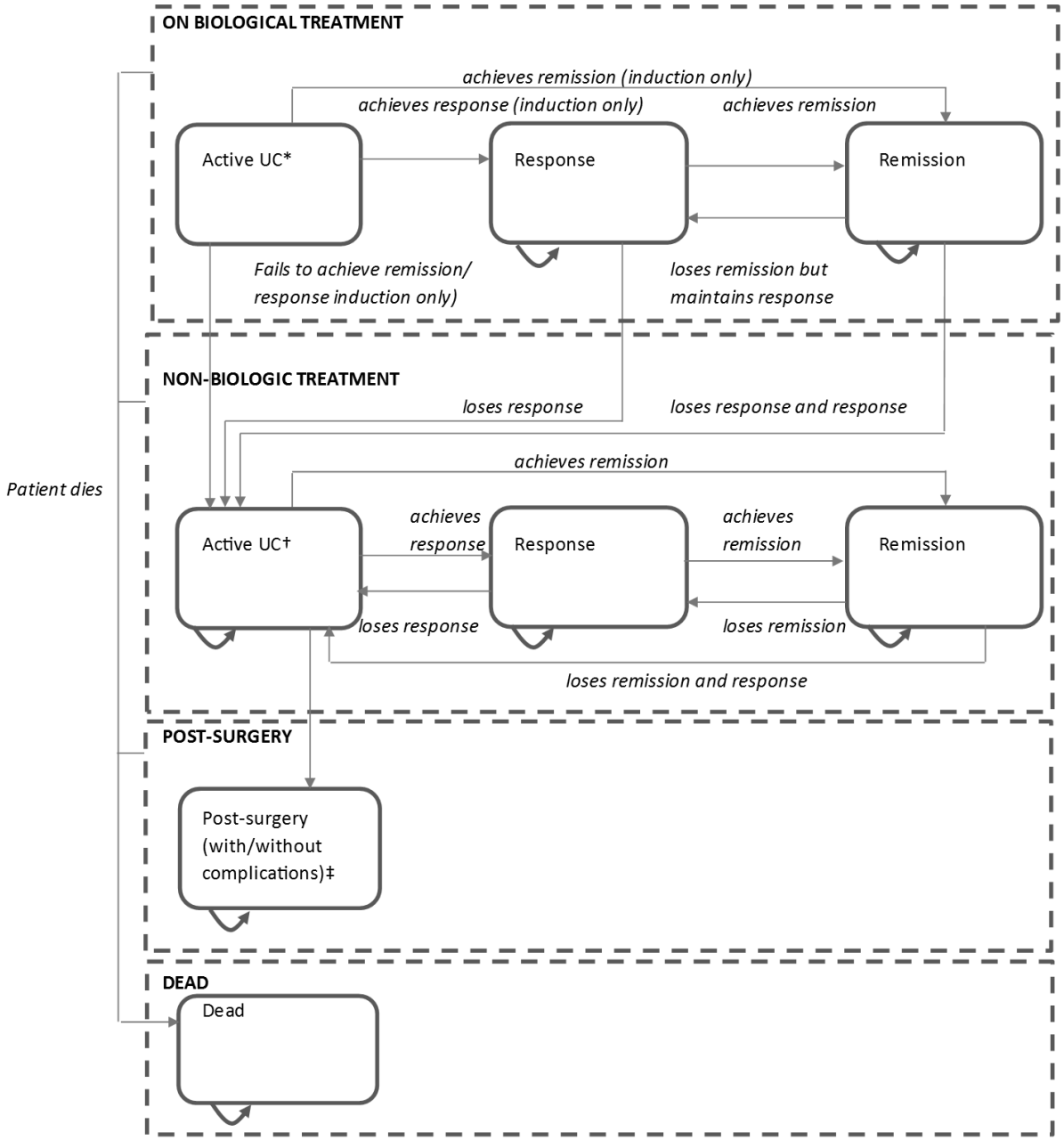
During the maintenance phase, patients receiving biologic therapy are assumed to continue receiving the same biologic treatment for as long as they continue to maintain response/remission. If patients receiving biologic therapy lose their response at any point they are assumed to transit to the active UC state and subsequently receive conventional therapy. Patients in the conventional treatment group, and those who have previously achieved but lost response to biologic therapy, are assumed to continue receiving conventional therapy irrespective of whether they achieve response or remission to that conventional therapy. A time-independent probability of undergoing surgery is applied to those patients receiving conventional treatment with active UC; the model assumes that this only possible within the active UC state. Patients in the immediate colectomy group, and those who have undergone surgery after receiving biologic/conventional treatment, remain in the post-surgery state for the remainder of the model time horizon. All patients have a probability of dying from other causes during each model cycle.

Differential levels of HRQoL are assigned to each model health state. Disutilities are assigned to those patients who develop chronic pouchitis – other complications of surgery are assumed to be transient and are assumed not to have a long-term impact upon patients' HRQoL. QALY gains in each arm of the model are driven by sojourn time in each of the model's health states and differential rates of surgery across the biologic groups, the conventional management group and the immediate colectomy group. Resource costs are assigned in terms of drug acquisition, drug administration

(infliximab only), surgery and related complications and UC health state costs (elective/emergency endoscopy, blood tests, consultant visits and hospitalisations).

The probability of residing in each health state during a given model cycle is estimated using simple matrix multiplication. Transitions between states are handled within a three-stage competing risks framework whereby (i) patients undergo transitions between each of the pre-surgical UC treatment states based on individual transition probabilities estimated using the NMAs (see Section 5.2.3) and the estimated colectomy rate, (ii) the populations of the post-surgery and dead states are adjusted to reflect surgical mortality rates, and (iii) the remaining surviving population is adjusted to account for other-cause mortality conditional on the patient cohort's current age. Given the different durations of the induction and maintenance phases, a half-cycle correction is not applied within the model.

Figure 88: Assessment Group model structure (induction and maintenance phases)



**Patients in the infliximab, adalimumab and golimumab groups begin in this portion of the model*

† Patients in the conventional non-biologic management group begin in this portion of the model

‡ Patients in the immediate colectomy group begin in this portion of the model

Key model assumptions

The Assessment Group model makes the following key assumptions:

- At the beginning of the maintenance phase, the decision to continue therapy with infliximab, adalimumab and golimumab is determined by the achievement of response or remission at the end of induction

- During each maintenance cycle, the decision to continue therapy with infliximab, adalimumab and golimumab is determined by the achievement/maintenance of response or remission at the end of the previous maintenance cycle
- Patients who discontinue biologic therapy are assumed to receive conventional treatment
- Patients with active UC receiving conventional treatment may undergo colectomy during any cycle; patients receiving biologic therapy will receive at least one cycle of conventional treatment before transiting to surgery
- Patients' HRQoL is assumed to be determined by their level of disease control, whether they have previously undergone colectomy and the incidence of post-surgical complications
- With the exception of chronic pouchitis, all surgery-related complications are assumed to occur during the first cycle following surgery
- With the exception of chronic pouchitis, surgical complications are assumed to be transient and can be resolved either through further surgery or through medical management
- The medical management of surgery-related complications is assumed to require a 7-day admission on a gastroenterology ward
- The incidence of chronic pouchitis is assumed to be associated with ongoing additional treatment costs and a decrement in patients' level of HRQoL.

6.2.2.3 Evidence used to inform the model's parameters

Table 65 summarises the source used to inform the groups of parameters within the model. These are described in further detail in the following sections.

Table 65: Summary of evidence sources used to inform the model's parameter values

Parameter group	Source(s) used to inform parameter values
Patient characteristics (starting age, mean body mass, proportion of patients with body mass > 80kg, proportion of patients who are female).	Patient age, mean body mass and the probability that a patient is female were derived from the RCTs included in the systematic review of clinical effectiveness (see Chapter 5). The proportion of patients with body mass>80kg was taken from the MSD golimumab model.[67]
Pre-surgical health state transition rates induction phase	*De novo* network meta-analysis of induction trials
Pre-surgical health state transition rates maintenance phase	*De novo* network meta-analysis of maintenance trials
Surgery rate during each maintenance cycle	Solberg *et al*[11]
Probability of peri-operative mortality	UK IBD Audit 2012[145]
Probability of other-cause mortality conditional on age and sex	ONS life tables for England and Wales 2009-2011[146]
Health state utilities for all pre-surgical and post-surgical states	Woehl *et al*[109]
Disutility associated with chronic pouchitis	Arseneau *et al*[110]
Biologic drug regimen schedules	Based on the SmPCs and trials for infliximab, adalimumab, and golimumab[133,147,148]
Biologic drug regimen usage, duration and dosing	Expert opinion (personal communication: Professor Alan Lobo, Consultant Gastroenterologist, Sheffield Teaching Hospitals).
Probability of surgery-related complications and proportion of cases requiring surgery/medical treatment	Arai *et al*[108]
Use of other related resources for the management of UC	Tsai *et al*[82]
Relative risk of hospitalisation for biologics versus conventional treatment	MSD submissions[65,67]
Unit costs	BNF[38] and NHS Reference Costs 2013[131]

Patient characteristics

Patient characteristics were based on data reported within the trials included in the systematic review[45,48-50,125,126] (see Chapter 5). Patients are assumed to enter the model aged 40 years. Forty three of patients are assumed to be female and patients are assumed to have a mean body mass of 77kg. Thirty two percent of patients are assumed to have a body mass greater than 80kg; this estimate was drawn from the MSD golimumab model.[67]

Transition probabilities for biologic and non-biologic therapies

The methods for the NMA models are described in Section 5.2.3.4. Table 66 presents the means and 95% credible intervals for transitions within the model (note all patients in the colectomy group who survive their surgery are assumed to transit immediately to the post-surgery group).

Table 66: Transition probabilities applied in the Assessment Group model

	Conventional non-biological treatment	Infliximab 5mg/kg	Adalimumab 160/80/40mg	Golimumab 200/100/50mg	Golimumab 200/100/100mg
Induction phase					
TP no response to no response	0.64 [0.57, 0.71]	0.29 [0.17, 0.44]	0.49 [0.33, 0.64]	0.45 [0.26, 0.64]	
TP no response to response	0.26 [0.21, 0.31]	0.35 [0.28, 0.41]	0.32 [0.25, 0.39]	0.33 [0.24, 0.39]	
TP no response to remission	0.10 [0.06, 0.15]	0.36 [0.21, 0.52]	0.19 [0.09, 0.32]	0.22 [0.09, 0.39]	
Maintenance phase 1					
TP no response to no response	0.85 [0.75, 0.92]	1.00*	1.00*	1.00*	1.00*
TP no response to response	0.10 [0.04, 0.17]	0.00	0.00	0.00	0.00
TP no response to remission	0.06 [0.02, 0.11]	0.00	0.00	0.00	0.00
TP response to no response	0.52 [0.43, 0.62]	0.43 [0.22, 0.66]	0.51 [0.23, 0.78]	0.40 [0.17, 0.66]	0.37 [0.15, 0.62]
TP response to response	0.27 [0.20, 0.34]	0.28 [0.19, 0.37]	0.26 [0.14, 0.35]	0.28 [0.18, 0.37]	0.29 [0.18, 0.38]
TP response to remission	0.21 [0.12, 0.31]	0.29 [0.11, 0.52]	0.23 [0.05, 0.49]	0.31 [0.11, 0.59]	0.35 [0.13, 0.62]
TP remission to no response	0.35 [0.17, 0.57]	0.32 [0.08, 0.65]	0.43 [0.10, 0.80]	0.18 [0.03, 0.46]	0.18 [0.03, 0.47]
TP remission to response	0.18 [0.07, 0.32]	0.17 [0.06, 0.30]	0.17 [0.05, 0.30]	0.14 [0.03, 0.28]	0.14 [0.03, 0.28]
TP remission to remission	0.47 [0.23, 0.71]	0.51 [0.18, 0.83]	0.41 [0.08, 0.80]	0.69 [0.32, 0.93]	0.68 [0.32, 0.93]
Maintenance phase 2					
TP no response to no response	0.97 [0.93, 1.00]	1.00*	1.00*	1.00*	1.00*
TP no response to response	0.02 [0.00, 0.05]	0.00	0.00	0.00	0.00
TP no response to remission	0.01 [0.00, 0.04]	0.00	0.00	0.00	0.00
TP response to no response	0.34 [0.07, 0.71]	0.25 [0.01, 0.72]	0.45 [0.06, 0.89]	0.30 [0.02, 0.75]	0.41 [0.05, 0.85]
TP response to response	0.37 [0.12, 0.60]	0.34 [0.06, 0.62]	0.33 [0.07, 0.56]	0.35 [0.08, 0.62]	0.34 [0.08, 0.58]
TP response to remission	0.29 [0.03, 0.72]	0.41 [0.03, 0.89]	0.22 [0.01, 0.72]	0.35 [0.02, 0.84]	0.25 [0.01, 0.74]
TP remission to no response	0.30 [0.17, 0.45]	0.25 [0.03, 0.61]	0.08 [0.01, 0.29]	0.33 [0.08, 0.66]	0.27 [0.05, 0.60]
TP remission to response	0.16 [0.03, 0.45]	0.14 [0.02, 0.41]	0.08 [0.00, 0.34]	0.16 [0.02, 0.41]	0.15 [0.02, 0.42]
TP remission to remission	0.54 [0.24, 0.73]	0.61 [0.17, 0.93]	0.83 [0.45, 0.99]	0.52 [0.14, 0.85]	0.59 [0.17, 0.89]

** Patients on biologic treatment in active UC (no response) are assumed to discontinue and subsequently receive conventional non-biologic treatments*

It should be noted that beyond 1-year, the model repeatedly uses the transition probabilities derived within the maintenance phase 2 NMA.

Surgery rate

The rate at which patients with moderate-to-severe UC progress to colectomy was based on estimates from the literature. A focussed Medline search was undertaken to identify studies reporting long-term rates of colectomy in patients with moderate-to-severe UC. Medline was searched from inception to April 2014 using a simple search comprised of two search terms: "ulcerative colitis/ exp" and "colectomy rate.tw." Studies were considered for inclusion in the economic model if they reported on long-term colectomy rates and if they either related to the moderate-to-severe population as a collective group of patients, or if they reported on colectomy rates in moderate and severe UC populations separately.

The Medline search identified 70 citations. Of these, only 6 studies were identified which reported on long-term colectomy rates for patients in a selected moderate to severe UC population (see Table 67).

Table 67: Summary of studies reporting on long-term colectomy rates in UC population

Study	Population	Follow-up duration	Reported colectomy rate
Actis 2007[149]	Patients admitted consecutively to study unit with an attack of UC and treated with ciclosporin between January 1991 and December 1999 (responders available for analysis, n=34)	7-years	24/34 (65%)
Gower-Rousseau *et al*[150]	All patients from the EPIMAD registry diagnosed with UC between January 1988 and December 2002 and who were less than 17 years old at the time of diagnosis (n=113)	Median 6.42 years (range – 3.83 years to 10.42 years).	Approximately 25% (Kaplan-Meier estimate)
Molnar *et al*[151]	UC patients admitted between 1998 and 2005 to tertiary clinic because of severe exacerbation of UC requiring parenteral corticosteroid therapy (n=183).	Average 4.4 years (range 1.1 years to 10 years)	16/110 (14.5%) steroid-responders 29/73 (39.7%) steroid- refractory Overall = 24.6%
Mocciaro *et al*[152]	Two historical cohorts of UC patients with severe relapse refractory to iv steroid treatment administered according to the "Oxford regimen" (n=65).	Mean 6.23 years ±5.07 years	Infliximab group = 60% Ciclosporin group = 30%
Gustavsson *et al*[14]	158 patients with UC treated in 1975–1982 with iv corticosteroid treatment.	Median 14.42 years (range 0.33 to 22.58 years)	All UC (n=147): colectomy rate= approximately 50% Mild UC (n=20): colectomy rate= approximately 40% Moderate UC (n=45): colectomy rate= approximately 50% Severe UC (n=61): colectomy rate= approximately 62%
Solberg *et al*[11]	Population-based cohort of 843 patients with inflammatory bowel disease was enrolled in South-Eastern Norway	Cohort followed-up at 1, 5 and 10 years.	Cumulative colectomy rate after 10 years = 9.8% (95% CI: 7.4-12.4%)

Several studies report estimates for patients who have been hospitalised for UC flare; these are likely to over-estimate the true colectomy rate in the moderate to severe population. On consideration of the remaining studies, the study reported by Solberg *et al*[11] was selected for inclusion in the model as this study was large (423 patients completed 10-year follow-up) and did not specifically relate to patients who had experienced UC flare. A constant 6-month colectomy rate of 0.0051 was applied within the model. Uncertainty surrounding this probability was modelled using a beta distribution.

Mortality

The model includes two types of mortality: peri-operative mortality associated with colectomy and other-cause mortality. Additional risks of death, e.g. due to the increased risk of colorectal cancer, are excluded from the model as this risk is likely to be small. Peri-operative mortality rates were taken from the third round of the UK IBD Audit[145] Within the 2012 publication of the UK IBD audit, there were 28 deaths reported amongst 807 elective and emergency surgical episodes in adult patients with UC; a probability of death of 0.03 is assumed within the cycle in which the patient undergoes surgery. Other-cause mortality was modelled according to age- and sex-specific life tables from the Office for National Statistics (ONS).[146] The annual probability of death during each model cycle was adjusted to reflect the duration of induction and maintenance cycles (8 weeks and 26 weeks respectively) using standard methods.[153] Uncertainty surrounding the peri-operative mortality rate was modelled using a beta distribution. No uncertainty was modelled for other-cause mortality.

Probability of experiencing surgery-related complications

The trials used to inform the efficacy parameters do not include details of surgery-related complications. Instead, the model uses data reported in Arai *et al*[108] to inform parameters relating to the probability of experiencing transient and chronic surgery-related complications and the probabilities that these complications are treated using medical or surgical approaches. Given the types of complications reported in Arai *et al*[108] (see Table 68), the model assumes that all are transient with the exception only of pouchitis. Therefore, the model assumes that 47.3% (140/296) patients will develop transient complications, with a further 5% of patients developing chronic pouchitis. Based on the reported timing of complications within the Arai *et al* study,[108] the model assumes that all transient complications will arise and will be resolved during the first cycle following surgery (in those patients who survive their surgery). Chronic pouchitis is assumed to continue for the remainder of the patient's lifetime. The model assumes that 19% of complications require further surgery, whilst the remaining 81% require medical treatment only.

Table 68: Surgery-related complication frequency and treatment approach[108]

Complication type	Complication frequency		Treatment approach	
	N early	N late	Medical	Surgical
Anastomotic stricture	7	56	63	0
Staple line ulcer	9	31	38	2
Pouchitis	0	16	16	0
Bowel obstruction	6	15	16	5
Proctitis	1	17	18	0
Pelvic sepsis	12	2	1	13
Peritoneal abscess	3	0	0	3
Anal fistula	0	12	2	8
Incisional hernia	1	11	0	4
Total	39	160	154	35

HRQoL

Within the model HRQoL is assigned according to the level of disease control achieved with drug therapy (active UC, response, remission), whether the patient has previously undergone colectomy and whether the patient is experiencing post-surgical complications. The same utility values are used for all biologic and non-biologic drug treatments. The Assessment Group undertook a systematic review of studies reporting valuations of states relating to different levels of UC control and post-surgery.

Searches were undertaken to identify utilities literature relating to UC, specifically using the EQ-5D instrument. The following electronic databases were searched from inception for utilities literature:

- MEDLINE, MEDLINE in-Process and Other Non-Indexed Citations: Ovid. 1946-present
- EMBASE: Ovid. 1974-2013 present
- Cochrane Library: Wiley Interscience
    - Cochrane Database of Systematic Reviews (CDSR). 1996-present
    - Database of Abstracts of Reviews of Effects (DARE). 1995-present
    - Cochrane Central Register of Controlled Trials (CCRT). 1995-present
    - Cochrane Methodology Register. 1904-present
    - Health Technology Assessment Database (HTA). 1995-present
    - NHS Economic Evaluation Database (NHS EED). 1995-present
- CINAHL: EBSCO. 1982-present
- Web of Science Citation Index: Web of Knowledge. 1900-present
- Conference Proceedings Citation Index: Web of Knowledge. 1990-present
- BIOSIS Previews: Web of Knowledge. 1969-present
- EconLit: Ovid. 1886- present

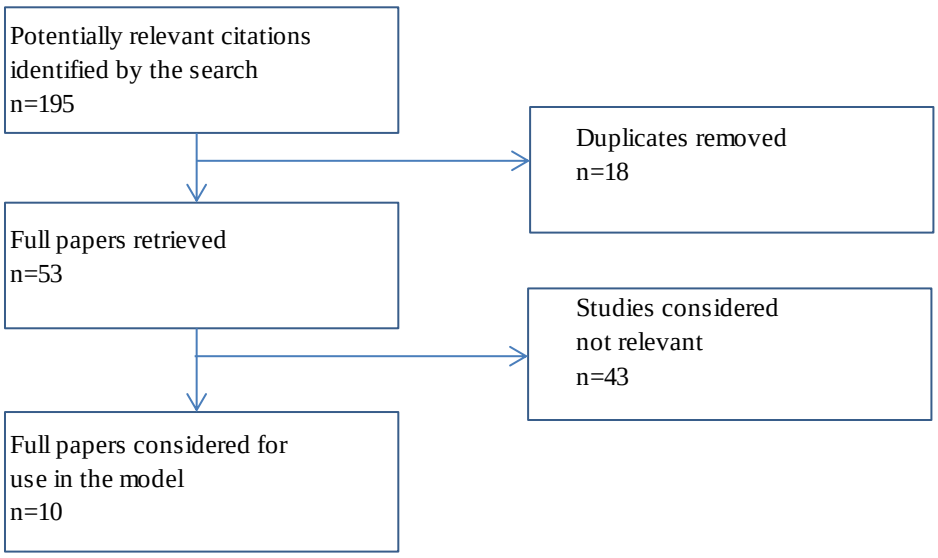
The MEDLINE search strategies are presented in Appendix 11. The search strategy combined freetext and MeSH (Medical subject headings) or thesaurus terms relating to UC combined with terms for specific utility measures or more general utility terms. The search strategy was translated across all databases. No date or language restrictions were applied. Literature searches were conducted during January and February 2014. References were collected in a bibliographic management database, and duplicates were removed. The results of the search are summarised in Table 69.

Table 69: EQ-5D utilities search results

Database	Date range	Date searched	Number of results
Medline (Ovid)	1946-present	29/01/14	52
Embase (Ovid)	1974-present	29/01/14	113
CINAHL (EBSCO)	1982-present	04/02/14	0
SCI & SSCI (WOK)	1900-present	04/02/14	5
BIOSIS (WOK)	1969-present	04/02/14	4
CDSR (Wiley)	1991-present	29/01/14	0
CENTRAL (Wiley)	1991-present	29/01/14	2
Cochrane HTA (Wiley)	1991-present	29/01/14	0
Cochrane DARE (Wiley)	1991-present	29/01/14	0
Cochrane EED (Wiley)	1991-present	29/01/14	0
EconLit (Ovid)	1886-present	29/01/14	1

Studies were considered potentially includable if they reported EQ-5D utility estimates for multiple UC health states or if they reported valuations of post-surgery states. The study selection process is shown in Figure 89.

Figure 89: Study selection results



Of the 177 de-duplicated, potentially relevant studies, the full papers of 53 citations were retrieved for more detailed examination by the Assessment Group based on their titles and abstracts. Of these 53

citations, 10 studies reported EQ-5D estimates for one or more health states relevant to the model.[48,50,109,140,154-159] Seven of these studies reported estimates for multiple pre-surgery UC health states; the remaining three studies reported estimates only for post-surgery state only (see Table 70). Of the 10 potentially relevant EQ-5D studies, those reported by Woehl *et al*[109] and Swinburn *et al*[140] appear to be the most useful as they are UK-based, included a fairly large number of patients (n=180 and n=230 respectively) and have the greatest coverage of the health states in the model (see Table 70).

The study reported by Swinburn *et al*[140] examines the impact of colectomy on the HRQoL of patients with UC. Three hundred and thirty participants were recruited into the study, comprising 230 UC patients (30 of whom had previously undergone surgery) together with 100 age- and gender-matched controls. EQ-5D utilities were collected via online survey. For both post-surgery patients versus non-surgery patients and post-surgery patients versus controls, EQ-5D utility scores were compared across IBDQ disease severity. Seventy eight patients had remission, 47 patients had mild disease, 31 patients had moderate disease and 44 patients had severe disease. The utility for patients post-surgery was reported to be 0.59 (95% c.i. 0.55-0.63). For patients who had not undergone surgery, the scores for each disease severity are: remission utility=0.91 (95% c.i. 0.87-0.95), mild disease utility=0.80 (95% c.i. 0.70-0.85), moderate disease utility=0.68 (95% c.i. 0.58-0.78) and severe disease utility=0.45 (95% c.i. 0.35-0.55). Across the total UC pre-surgery population, the mean EQ-5D utility was reported to be 0.75 (95% c.i. 0.71-0.79). Similarly, for the matched controls, the mean EQ-5D utility was estimated to be 0.79 (CI 0.75-0.83). Swinburn *et al*[140] report that on average, post-surgery patients reported lower HRQoL scores than non-surgery patients (p=0.016) and matched controls (p=0.03).

Woehl *et al*[109] collected EQ-5D utility scores from 180 patients with active UC. Within this study population, the mean age was 55.0 years (s.d.=14.2) and the mean age at diagnosis was 34.1 years (s.d.=14.6). UC disease severity groups were categorised by SCAI-2 and were compared against patients with IPAA and ileostomy. The mean EQ-5D score was 0.73 (s.d.=0.29). Mean EQ-5D utilities were reported to be 0.87 (s.d.=0.15) for remission, 0.76 (s.d.=0.18) for mild disease, and 0.41 (s.d.=0.34) for moderate/severe disease. Patients who had undergone IPAA reported an EQ-5D utility of 0.71 (s.d.=0.29) whilst patients with an ileostomy reported an EQ-5D score of 0.72 (s.d.=0.35). Therefore, the health utility scores for these surgery states were slightly below a mild disease severity. The difference between these five groups was statistically significant (p=0.001).

In the base case analysis of the Assessment Group model, the Woehl *et al* study was selected for use as the valuation for the surgery state (0.71 to 0.72) is more consistent with the other post-surgery valuations identified[157-159] as compared against the Swinburn *et al*[140] study.

Table 70: Studies included in the systematic review of utility values

	1	2	3	4	5	6	7	8	9	10
State / Study	ACT1/2	PURSUIT	Swinburn†	Woehl	Casellas†	Leidl†	Vaizey	Van der Valk	Richards	Kuruvilla
Study characteristics										
Sample size	486*	464	230	180	528	232	173	982	56	59
Country	Various	Various	UK	UK	Spain	Germany (UK tariff)	UK	Netherlands	UK	US
Health state valuations										
Remission	0.84-0.88	0.86-0.89	0.91	0.87	1.00	0.91	0.86	NR	NR	NR
Response	0.79-0.82	0.80	0.80	0.76	0.70	0.74	0.77	NR	NR	NR
Active UC	NR	NR	0.55	0.41	0.55	0.63	0.66	NR	NR	NR
Post-surgery	NR	NR	0.59	0.71-0.72	NR	NR	NR	0.85‡	0.85	0.90‡
Post-surgery complications	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

NR=not reported
**Licensed arms only*
† Approximate estimates based on graph
‡ Same value reported for pouch and for ileostomy

In order to maintain the ordinal ranking of remission, response and active UC states, remission was modelled as a baseline utility score parameter, with disutilities used to value the reductions in health associated with the loss of remission and the loss of response relative to a baseline of remission. The utility parameter for response is therefore modelled as *Utility(remission) minus disutility(loss of remission)* whilst the utility parameter for active UC is modelled as *Utility(remission) minus disutility(loss of remission) minus disutility(loss of response).* The utility score for post-surgery was based on the mean value reported by Woehl *et al*[109] (this parameter was not characterised as a health decrement). Uncertainty surrounding the parameters describing remission utility and post-surgery utility was modelled using beta distributions, assuming that an equal number of patients were in each UC state. The disutility parameters were based on the mean and variance of the differences between the health states; this method ensures that the notionally better health state always has a monotonically better valuation than that for the notionally worse health state.

As the studies identified for inclusion in the review did not identify any studies which employed the EQ-5D to value the health loss associated with surgery-related complications, other sources were required. The Assessment Group model adopts a similar approach to the AbbVie model to value this health decrement based on the difference between the surgery and chronic pouchitis valuations reported by Arseneau *et al*[110] (TTO valuation – 0.57-0.40=0.17). It should be noted that this study used scenario-based TTO elicitation methods and therefore deviates from the NICE Reference Case.[78] Health losses associated with transient complications of surgery are assumed not to last long enough to decrease HRQoL. The utility values in the model were not adjusted by age. The final utility vector for each treatment option is shown in Table 71.

Table 71: Utility vectors for all states and treatment options

	Receiving biologic treatment			Receiving standard care			
Treatment	No response	Response	Remission	No response	Response	Remission	Post-surgery*
Conventional management	0.41	0.76	0.87	0.41	0.76	0.87	0.70
Infliximab	0.41	0.76	0.87	0.41	0.76	0.87	0.70
Adalimumab	0.41	0.76	0.87	0.41	0.76	0.87	0.70
Golimumab	0.41	0.76	0.87	0.41	0.76	0.87	0.70
Colectomy	0.41	0.76	0.87	0.41	0.76	0.87	0.70

**Including a proportion of patients with chronic pouchitis*

Resource costs

The model includes resource costs related to drug acquisition, drug administration (infliximab only), consultant visits, emergency and elective endoscopy, hospitalisations, blood tests, surgery and the management of surgery-related complications.

Biologic drug resource use, acquisition and cost

Table 72 shows the dosing regimens and associated costs for each of the biologic options within the model. With the exception of golimumab induction therapy, the biologic regimen assumed reflects the wording of the SmPC for that product.[133,147,148] It should be noted that the SmPC for golimumab recommends that continued therapy should be reconsidered in patients who show no evidence of therapeutic benefit within 12-14 weeks (after 4 doses). However, the PURSUIT-SC trial[48] evaluated clinical benefits at 6 weeks (after 2 doses). The costs and benefits of golimumab induction are modelled in line with the design of the PURSUIT-SC trial and therefore the costs within the induction phase relate only to the first two doses of golimumab. The dose adjustments for adalimumab were based on the estimate reported within the AbbVie submission.[63]

Table 72: Dosing regimens and costs for biologic options

Treatment group	Mean doses and frequency within cycle	Cycle cost
Induction phase (8 week duration)		
Infliximab 5mg/kg	12 x 100mg infliximab (3 outpatient appointments)	£5,035.44 (acquisition) + £893.18 (administration = £5,928.44
Adalimumab 160mg/80mg	4 x 40mg adalimumab + 2 x 40mg adalimumab (self-administered)	£2,817.12
Golimumab 200mg/100mg	4 x 50mg golimumab + 2 x 50mg golimumab (self-administered)	£4,577.82
Maintenance phase (26 week duration)		
Infliximab 5mg/kg	13.04 x 100mg infliximab (3.26 outpatient appointments)	£5,473.79 (acquisition) + £970.94 (administration) = £6,444.73
Adalimumab 40mg EOW dosing (72.6% patients)	13.04 x 40mg adalimumab	£4,593.54
Adalimumab 40mg EW dosing (27.4% patients)	26.08 x 40mg adalimumab	£9,187.08
Golimumab 100mg (body mass>80kg, 31.60% patients)	13.04 x 50mg golimumab	£9,952.67
Golimumab 50mg (body mass<80kg, 68.40% patients)	6.52 x 50mg golimumab	£4,976.34

Non-biologic drug resource use

The costs of conventional therapies in each UC state are shown in Table 73. These resource costs were assumed to be the same for all biologic options and for the conventional management group. The treatments, dosing and their frequencies were based on expert opinion (personal communication: Professor Alan Lobo, Consultant Gastroenterologist, Sheffield Teaching Hospitals) and BNF dosing recommendations. Where multiple products are available, the least expensive was included in the analysis. The model assumes that all patients would receive 5-ASAs using Ipocol® at a dose of 2.4mg/day during induction and 1.2mg/day during maintenance. Ninety percent of patients are also assumed to receive topical 5-ASAs (80% enemas, 10% suppositories) during induction; these are assumed to be given for a maximum of 28-days per cycle. Following loss of response, the same therapies may also be used to re-induce response/remission; these same assumptions are applied during each maintenance cycle to the conventional management active UC state only. Eighty percent of patients are assumed to receive 2.5mg/kg azathioprine daily, with the remaining 20% receiving 6-MP at a dose of 1.5mg/kg daily (note it is likely that a lower proportion of patients will actually fulfil the criteria for treatment hence this may be an overestimate). All patients are also assumed to require one course of oral prednisolone during induction with the dose starting at 40mg and tapered by 5mg each week until the dose is zero (again, the same assumption is made with respect to re-induction of response/remission in patients in the conventional management active UC state during each maintenance cycle). The model does not include estimates of uncertainty around drug usage.

Table 73: Conventional drug regimen costs per induction/maintenance cycle

Drug - brand	Regimen assumed (% use)	Unit cost	Cost per cycle (for those receiving treatment)
Induction phase (8 weeks)			
5-ASAs (oral) – Ipocol®	2.4mg/day for 56 days (100% patients)	400mg (120 tabs) = £17.68	£49.50
5-ASAs (enema) – Asacol®	1 metered application/day for 28 days (80% patients)	1mg (14 application canister) = £26.72	£53.44
5-ASAs (suppository) – Asacol®	1.5g/day for 28 days (10% patients)	250mg (20 suppository pack) = £4.82	£20.24
Azathioprine – non-proprietary	2.5mg/kg daily for 56 days (80% patients)	50mg (56 tabs) = £3.85	£14.82
6-mercaptopurine – Puri-Nethol®	1.5mg/kg daily for 56 days (20% patients)	50mg (25 tabs) = £50.47	£261.15
Prednisolone (oral) – non-proprietary	40mg tapered by 5mg/week until dose is zero after 56 days (100% patients)	5mg (28 tabs) = £1.03	£9.27
Maintenance phase (26 weeks)			
5-ASAs (oral) – Ipocol®	1.2mg/day for 182.63 days (100% patients)	400mg (120 tabs) = £17.68	£80.72
5-ASAs (enema) – Asacol®*	1 metered application/day for 28 days (80% patients)	1mg (14 application canister) = £26.72	£53.44
5-ASAs (suppository) – Asacol®*	1.5g/day for 28 days (10% patients)	250mg (20 suppository pack) = £4.82	£20.24
Azathioprine – non-proprietary	2.5mg/kg daily for 182.63 days (80% patients)	50mg (56 tabs) = £3.85	£48.34
6-mercaptopurine – Puri-Nethol®	1.5mg/kg daily for 182.63 days (20% patients)	50mg (25 tabs) = £50.47	£851.66
Prednisolone (oral) – non-proprietary*	40mg tapered by 5mg/week until dose is zero after 56 days (100% patients)	5mg (28 tabs) = £1.03	£9.27

** Costs included for patients in conventional management active UC state only to re-induce response/remission*

Other UC health state resource use

Health state costs relating to the use of elective and emergency endoscopy, hospitalisations, consultant visits and blood tests were taken from the previous economic analysis reported by Tsai *et al*[82] (see Table 74). As Tsai *et al*[82] did not report any uncertainty around these resource use estimates, the standard error was arbitrarily assumed to be equal to 10% of the mean.

Table 74: UC resource use per year[82]

Resource item	Remission	Response	No response	Post-surgery remission	Post-surgery complications
Consultant visit	2.00	4.50	6.50	1.50	1.75
Hospitalisation episode	0.30	0.30	0.30	0.00	3.25
Blood tests	3.25	3.90	6.50	1.50	3.25
Elective endoscopy	0.20	0.50	2.00	1.25	0.65
Emergency endoscopy	0.00	0.25	0.75	0.50	0.13

The MSD submissions to NICE included a meta-analysis in which relative risks were derived for hospitalisations for adalimumab 160mg/80mg/40mg and infliximab 5mg/kg versus conventional treatment.[65,67] The MSD NMA did not include estimates of the relative risk of hospitalisation for golimumab versus conventional treatment as this was not measured in the PURSUIT-Maintenance trial. The relative risk for golimumab was assumed to be the same as that for adalimumab (the least favourable option); this assumption favours golimumab compared against the other options included in the economic analysis. The relative risks used in the model are shown in Table 75.

Table 75: Relative risks of hospitalisation for infliximab, adalimumab and golimumab[65,67]

Drug	Relative risk of hospitalisation	Estimated standard error	Comment
Infliximab	0.64	0.13	Taken from MSD submission NMA[65,67]
Adalimumab	0.70	0.12	
Golimumab	0.70	0.12	Assumed to be the same as relative risk for adalimumab

Unit costs for UC health state resource use, surgery and complications

The unit costs for UC health state resource use, surgery and complications were taken from NHS Reference Costs and are shown in Table 76.

Table 76: Unit costs for UC health state resource use, surgery and complications

Item	Unit cost	Standard error	Source
Consultant visit	£123.43	£3.30	NHS Reference Costs 2013,[131] WF01A, Gastroenterology, Non-Admitted Face to Face Attendance, Follow-up
Hospitalisation	£2,722.96	£101.66	NHS Reference Costs 2013,[131] FZ37N, Gastroenterology, Inflammatory Bowel Disease, with Single Intervention, with CC Score 0-3
Elective endoscopy	£462.36	£14.96	NHS Reference Costs 2013,[131] FZ51Z, Gastroenterology, Diagnostic Colonoscopy, 19 years and over
Emergency endoscopy	£512.62	£26.20	NHS Reference Costs 2013,[131] FZ51Z, Gastroenterology, Diagnostic Colonoscopy, 19 years and over
Blood tests	£1.94	£0.26	NHS Reference Costs 2013,[131] DAPS03, Integrated Blood Services
Surgery	£8,792.85	£473.03	NHS Reference Costs 2013,[131] FZ73F, Colorectal Surgery, Very Complex Large Intestine Procedures with CC Score 0-2
Medical management of complications*	£4,170.95	£464.59	NHS Reference Costs 2013,[131] WA12D, Colorectal Surgery, Complications of Procedures, with CC Score 0

** Assumes a length of stay of 7 days*

6.2.2.4 Methods for model evaluation

The model is fully probabilistic. The base case analysis relates to use of infliximab, adalimumab and golimumab within an adult population, based on the expectation of the mean. The cost-effectiveness of competing options are evaluated within a fully incremental analysis according to standard cost-effectiveness decision rules.[77] Results of the probabilistic analyses are presented separately for patients for whom colectomy is a potential option and those for whom it is not. Decision uncertainty is represented using cost-effectiveness planes and CEACs.

A secondary analysis is presented for the paediatric population; this analysis compares infliximab versus standard non-biological treatments versus colectomy. Given the absence of head-to-head trials of infliximab versus any other therapy, this analysis is exactly the same as the base case analysis except that the patients' starting age is set to 15 years (the median age within Hyams *et al*).

In addition to the main analyses, a number of secondary scenario analyses and sensitivity analyses are presented (see Box 5). It should be noted that whilst the base case economic analysis utilises the results of the NMA models, sensitivity analyses #4 present pairwise estimates of cost-effectiveness using direct head-to-head transition probabilities sourced from

the individual clinical trials of infliximab, adalimumab and golimumab. The pairwise analysis of infliximab uses simple pooling of the ACT1/2 trial data. The pairwise analysis of adalimumab is based on the anti-TNF-naïve subgroup from ULTRA2. The golimumab analysis is based on the data from the PURSUIT-SC trial and the PURSUIT-Maintenance trial.

Unless otherwise stated, all results are discounted at a rate of 3.5%.

Box 5: Sensitivity/scenario analyses undertaken using the Assessment Group model

- *Base case analysis*: NMA using anti-TNF-α naïve subgroup from ULTRA2[46,126] and excluding Suzuki *et al*[47] (probabilistic)
- *Sensitivity analysis 1:* NMA using ITT population from ULTRA2[46] and excluding Suzuki *et al*[47] (probabilistic)
- *Sensitivity analysis 2:* NMA using anti-TNF-α naïve subgroup from ULTRA2[46,126] and including Suzuki *et al*[47] (probabilistic)
- *Sensitivity analysis 3:* NMA using ITT population from ULTRA2[46,126] and including Suzuki *et al*[47] (probabilistic)
- *Sensitivity analysis 4:* Pairwise head-to-head comparisons of infliximab, golimumab and adalimumab versus conventional management using direct trial evidence from ACT1/2, ULTRA2 and PURSUIT[46,48,50,126] (deterministic)
- *Sensitivity analysis 5*: Base case using point estimates of parameters (deterministic)
- *Sensitivity analysis 6:* Time horizon=20 years (deterministic)
- *Sensitivity analysis 7:* Time horizon=10 years (deterministic)
- *Sensitivity analysis 8:* Time horizon=5 years (deterministic)
- *Sensitivity analysis 9:*All utilities except post-surgical complications drawn from Swinburn *et al*[140] (deterministic)
- *Sensitivity analysis 10:* Utilities of response/remission drawn from ACT1 trial[50] (deterministic)
- *Sensitivity analysis 11:* Utilities of response/remission drawn from PURSUIT-Maintenance trial[48] (deterministic)
- *Sensitivity analysis 12:* Relative risk of hospitalisation for golimumab vs conventional treatment assumed to be 1.0 (deterministic)
- *Sensitivity analysis 13:* UC health state costs doubled (deterministic)
- *Sensitivity analysis 14:* UC health state costs halved (deterministic)
- *Sensitivity analysis 15:* Probability of chronic pouchitis doubled (deterministic)
- *Sensitivity analysis 16:* Probability of chronic pouchitis halved (deterministic)
- *Sensitivity analysis 17:* Cost of surgery doubled (deterministic)
- *Sensitivity analysis 18:* Cost of surgery halved (deterministic)
- *Sensitivity analysis 19:* Probability of undergoing surgery in drug groups halved (deterministic)

6.2.2.5 Model verification and validation

The Assessment Group undertook a number of measures to ensure the credibility of the model (author/advisor initials are shown in brackets).

- Peer review of economic analysis by two internal clinical advisors (SH, AL), one external clinical expert (MH) and one external methodological expert (SD)
- Verification of executable model by a second modeller not involved in its implementation (HB)
- Double-programming of separate Markov models for all five treatment options by the lead modeller (PT)
- Scrutiny of implemented model coding and formulae by lead modeller (PT)
- Double-checking of accuracy of all model inputs against sources
- Comparison of model results using point estimates of parameters and expectation of the mean
- Comparison of mean of all parameter samples against point estimates of parameters
- Examination of all identified sources of discrepancy
- Model testing using sensitivity analysis and use of extreme parameter values
- Comparison of model results against manufacturers' models (see Section 6.1)

6.2.3 Assessment Group model results

6.2.3.1 Central estimates of cost-effectiveness (base case analysis - adults)

Table 77 presents the base case results generated using the probabilistic version of the model within an adult population in whom colectomy is an acceptable option. The base case analysis of the model suggests that colectomy is expected to produce 14.72 QALYs at a cost of approximately £41,900 over the patient's remaining lifetime. All medical options are expected to produce substantially fewer QALYs at a greater cost than colectomy, hence colectomy is expected to dominate infliximab, adalimumab, golimumab and conventional non-biologic treatments.

Table 77: Probabilistic cost-effectiveness results, base case analysis, adult patients in whom colectomy is an option (medical and surgical treatments)

Option	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Colectomy	14.72	£41,920.71	-	-	dominating
Adalimumab	10.83	£89,288.61	-	-	dominated
Infliximab	10.82	£94,664.81	-	-	dominated
Golimumab	10.65	£88,107.84	-	-	dominated
Conventional treatment	10.48	£71,592.46	-	-	dominated

Figure 90 presents cost-effectiveness acceptability curves (CEACs) for infliximab, adalimumab, golimumab, conventional treatment and colectomy for the adult population. Assuming a willingness to pay (WTP) threshold of £20,000 per QALY gained, the probability that colectomy produces the greatest amount of expected net benefit is approximately 0.97. The probability that any of the biologic treatments produce the greatest amount of expected net benefit at this threshold is approximately zero. Assuming a WTP threshold of £30,000 per QALY gained, the probability that colectomy produces the greatest amount of expected net benefit is approximately 0.96. The probability that any of the biologic treatments produce the greatest amount of expected net benefit at this threshold is approximately zero.

Figure 90: CEACs, base case analysis, adult patients in whom colectomy is an option (medical and surgical treatments)

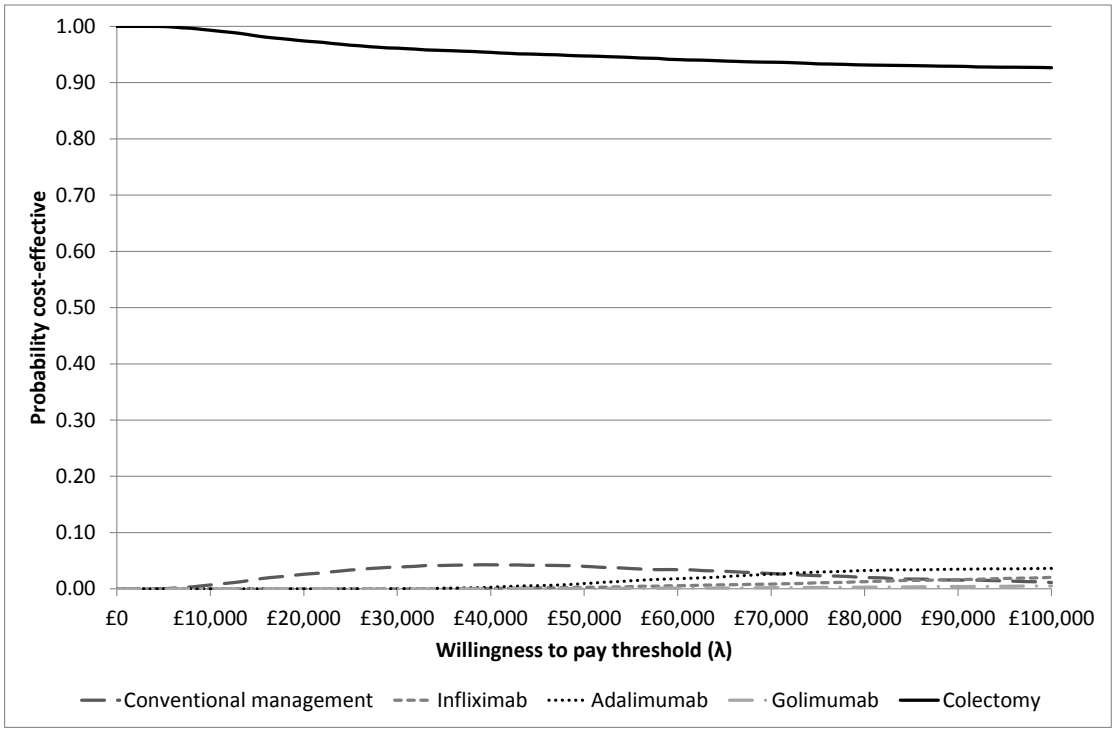


Table 78 presents the probabilistic base case model results within an adult population in whom colectomy is not an acceptable option, thus relevant treatment options are restricted to medical treatments only (infliximab, adalimumab, golimumab and conventional non-biologic treatments). The model results suggest that within this population infliximab is expected to be dominated by adalimumab (note - the difference in QALYs is very small), whilst golimumab is expected to be ruled out due to extended dominance. The incremental cost-effectiveness of adalimumab versus conventional treatment is expected to be £50,624 per QALY gained.

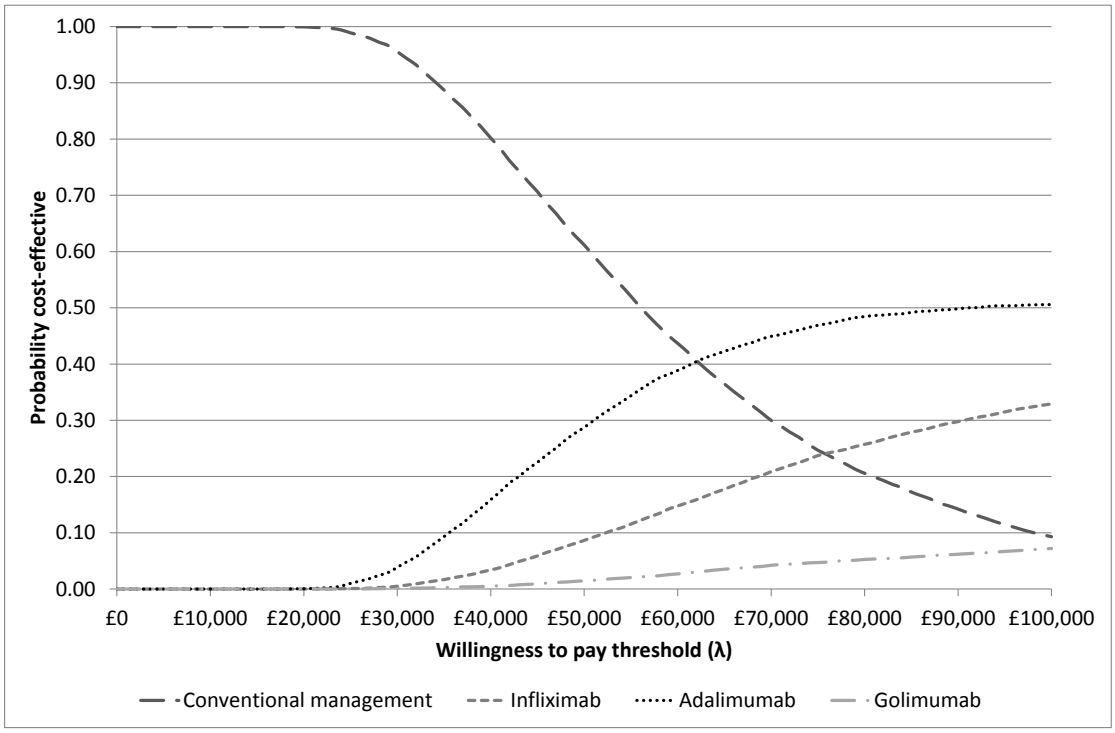
Table 78: Probabilistic cost-effectiveness results, base case analysis, adult patients in whom colectomy is not an option (medical treatments only)

Option	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Adalimumab	10.83	£89,288.61	0.35	£17,696.15	£50,623.66
Infliximab	10.82	£94,664.81	-	-	dominated
Golimumab	10.65	£88,107.84	-	-	ext dom
Conventional treatment	10.48	£71,592.46	-	-	-

Ext dom – extendedly dominated

Figure 91 presents CEACs for infliximab, adalimumab, golimumab and conventional treatment within a population in whom colectomy is not an option. Assuming a WTP threshold of £20,000 per QALY gained, the probability that conventional non-biologic treatment produces the greatest expected net benefit is approximately 1.0. Assuming a WTP threshold of £30,000 per QALY gained, the probability that conventional management produces the greatest expected net benefit is approximately 0.96.

Figure 91: CEACs, base case analysis, adult patients in whom colectomy is not an option (medical treatments only)



6.2.3.2 Central estimates of cost-effectiveness (base case analysis – paediatric population)

Table 79 presents the base case results generated using the probabilistic version of the model within a paediatric population in whom colectomy is an acceptable option. The base case analysis of the model suggests that colectomy is expected to produce 17.55 QALYs at a cost

of approximately £47,900 over the patient's remaining lifetime. Infliximab and conventional management are expected to produce substantially fewer QALYs at a greater cost than colectomy, hence colectomy is expected to dominate these medical options.

Table 79: Probabilistic cost-effectiveness results, base case analysis, paediatric patients in whom colectomy is an option (medical and surgical treatments)

Option	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Colectomy	17.55	£47,871.23	-	-	dominating
Infliximab	13.01	£106,759.45	-	-	dominated
Conventional treatment	12.67	£83,491.53	-	-	dominated

Figure 92 presents CEACs for infliximab, conventional treatment and colectomy for the paediatric population. Assuming a WTP threshold of £20,000 per QALY gained, the probability that colectomy produces the greatest amount of expected net benefit is approximately 0.98. Assuming a WTP threshold of £30,000 per QALY gained, the probability that colectomy produces the greatest amount of expected net benefit is approximately 0.96. The probability that infliximab produces the greatest amount of expected net benefit at these thresholds is approximately zero.

Figure 92: CEACs, base case analysis, paediatric patients in whom colectomy is an option (medical and surgical treatments)

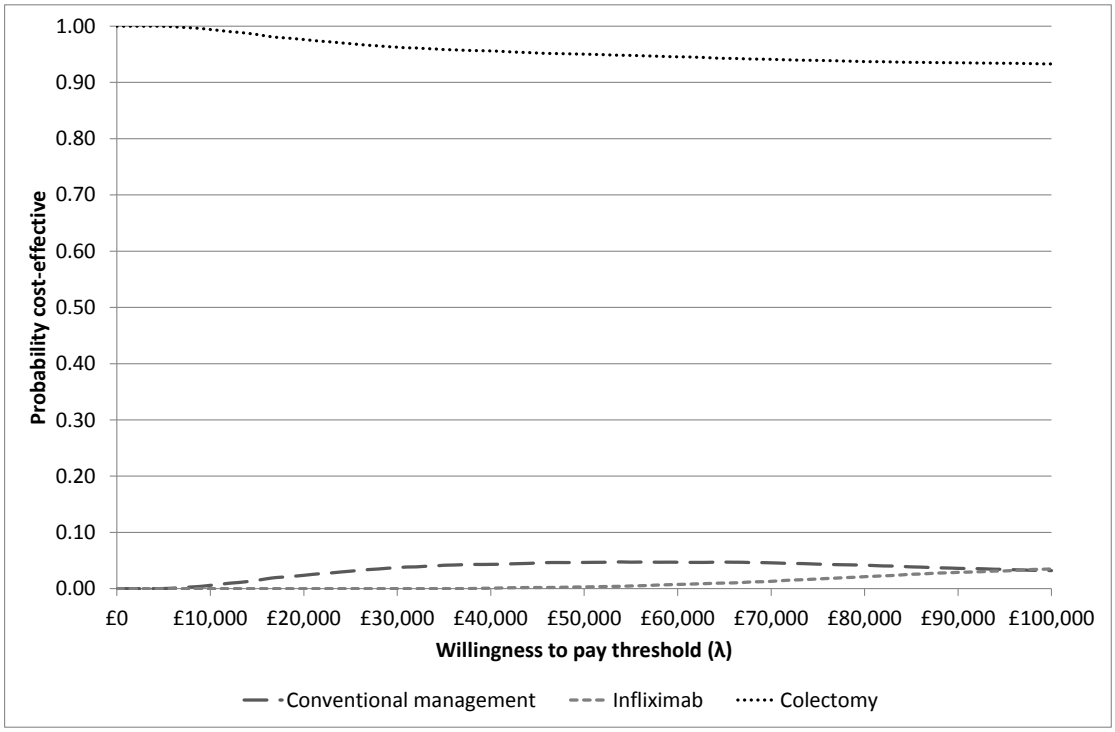


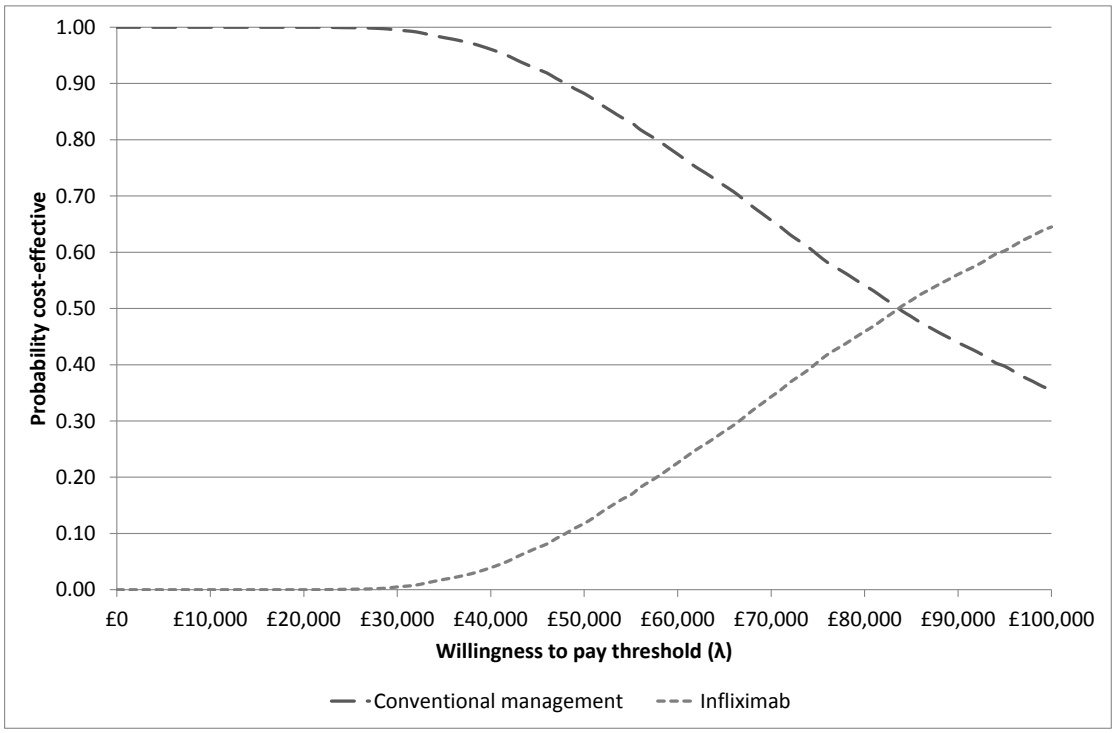
Table 80 presents the probabilistic base case model results within a paediatric population in whom colectomy is not an acceptable option, thus relevant treatment options are restricted to infliximab and conventional non-biologic treatments. The model indicates that within this population, infliximab is expected to produce an additional 0.34 QALYs at an additional cost of £23,268 over the patient's remaining lifetime; the ICER for infliximab versus conventional management is expected to be £68,364 per QALY gained.

Table 80: Probabilistic cost-effectiveness results, base case analysis, paediatric patients in whom colectomy is not an option (medical treatments only)

Option	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Infliximab	13.01	£106,759.45	0.34	£23,267.92	£68,364.03
Conventional treatment	12.67	£83,491.53	-	-	-

Figure 93 presents CEACs for infliximab and conventional treatment for the paediatric population. Assuming WTP thresholds of £20,000 and £30,000 per QALY gained, the probability that conventional management produces the greatest amount of expected net benefit is approximately 1.0. The probability that infliximab produces the greatest amount of expected net benefit at these thresholds is approximately zero.

Figure 93: CEACs, base case analysis, paediatric patients in whom colectomy is not an option (medical and surgical treatments)



6.2.3.3 NMA sensitivity analyses (sensitivity analyses 1-3, probabilistic)

Table 81 summarises the results of the economic analysis for the adult population based on the three alternative sensitivity analyses.

Table 81: Results of probabilistic NMA sensitivity analyses

	Colectomy	Infliximab	Adalimumab	Golimumab	Conventional management
Adult population in whom colectomy is an option					
NMA sensitivity analysis #1 – ULTRA2 ITT population, excluding Suzuki et al					
ICER	dominating	dominated	dominated	dominated	dominated
P(optimal £20,000/QALY)	0.97	0.00	0.00	0.00	0.03
P(optimal £30,000/QALY)	0.96	0.00	0.00	0.00	0.04
NMA sensitivity analysis #2 - ULTRA2 anti-TNF-α naïve subgroup, including Suzuki et al					
ICER	dominating	dominated	dominated	dominated	dominated
P(optimal £20,000/QALY)	0.96	0.00	0.00	0.00	0.04
P(optimal £30,000/QALY)	0.94	0.00	0.00	0.00	0.06
NMA sensitivity analysis #3 – ULTRA2 ITT population, including Suzuki et al					
ICER	dominating	dominated	dominated	dominated	dominated
P(optimal £20,000/QALY)	0.95	0.00	0.00	0.00	0.05
P(optimal £30,000/QALY)	0.93	0.00	0.00	0.00	0.93
Adult population in whom colectomy is an not option					
NMA sensitivity analysis #1 – ULTRA2 ITT population, excluding Suzuki et al					
ICER	n/a	£251,121 [1]	£54,309 [2]	ext dom [3]	- [4]
P(optimal £20,000/QALY)	n/a	0.00	0.00	0.00	1.00
P(optimal £30,000/QALY)	n/a	0.00	0.02	0.00	0.98
NMA sensitivity analysis #2 - ULTRA2 anti-TNF-α naïve subgroup, including Suzuki et al					
ICER	n/a	£525,806 [1]	£56,656 [2]	ext dom [3]	- [4]
P(optimal £20,000/QALY)	n/a	0.00	0.00	0.00	1.00
P(optimal £30,000/QALY)	n/a	0.00	0.01	0.00	0.98
NMA sensitivity analysis #3 – ULTRA2 ITT population, including Suzuki et al					
ICER	n/a	dominated [2]	£56,013.52 [1]	ext dom [3]	- [4]
P(optimal £20,000/QALY)	n/a	0.00	0.00	0.00	1.00
P(optimal £30,000/QALY)	n/a	0.00	0.02	0.00	0.98

Ext dom – extendedly dominated; n/a – not applicable
Where different to the base case analysis, the QALY rank is shown in parentheses [ ]

In the circumstances whereby colectomy is an option, the three NMA sensitivity analyses produce very similar results to the base case analysis. In all three analyses, colectomy is consistently expected to dominate all medical treatment options. Assuming a WTP threshold of £20,000 per QALY gained, the probability that colectomy produces the greatest amount of net benefit is expected to be between 0.95 and 0.97. Assuming a WTP threshold of £30,000 per QALY gained, the probability that colectomy produces the greatest amount of net benefit is expected to be between 0.93 and 0.96. Where colectomy is not an acceptable option, the results are influenced by which studies are included in the NMA, as the difference in

effectiveness between adalimumab and infliximab is very small. Golimumab is consistently expected to be ruled out of the analysis due to extended dominance. In sensitivity analyses 1 and 2, infliximab is expected to produce slightly more QALYs than adalimumab, however the ICER for infliximab versus adalimumab is expected to be greater than £251,000 per QALY gained. In these sensitivity analyses, the ICER for adalimumab versus conventional non-biologic treatment is expected to be greater than £54,000 per QALY gained. In sensitivity analysis 3, infliximab is expected to be ruled out due to simple dominance; the ICER for adalimumab versus conventional non-biologic treatments is expected to be approximately £56,000 per QALY gained.

6.2.3.4 Head-to-head analyses (sensitivity analysis 4)

Table 82 presents the results of the economic analysis using the direct head-to-head trial data.

Table 82: Head-to-head analysis – adult population, infliximab versus conventional management

Option	QALYs	Costs	Incremental QALYs	Incremental cost	ICER
Infliximab versus conventional management and colectomy					
Colectomy	14.69	£41,962.22	-	-	dominating
Infliximab	11.62	£85,116.81	-	-	dominated
Conventional treatment	11.44	£68,041.14	-	-	dominated
Adalimumab versus conventional management and colectomy					
Colectomy	14.69	£41,962.22	-	-	dominating
Adalimumab	10.23	£88,393.06	-	-	dominated
Conventional treatment	10.02	£73,206.96	-	-	dominated
Golimumab versus conventional management and colectomy					
Colectomy	14.69	£41,962.22	-	-	dominating
Golimumab	10.16	£89,228.93	-	-	dominated
Conventional treatment	9.99	£73,321.49	-	-	dominated

The head-to-head analyses indicate that colectomy is expected to dominate all medical options. Within a population in whom colectomy is not an acceptable option, the incremental cost-effectiveness of infliximab versus conventional non-biologic treatments is estimated to be £96,682 per QALY gained, the ICER for adalimumab versus conventional non-biologic treatments is estimated to be £70,075 per QALY gained and the ICER for golimumab versus conventional non-biologic treatments is estimated to be £90,720 per QALY gained.

Sensitivity analyses 5-19 - Other deterministic sensitivity analyses (medical and surgical options)

Table 83 presents the results of the remaining deterministic sensitivity analysis (analyses 5-19).

Table 83: Other deterministic sensitivity analyses

Sensitivity analysis	Incremental cost per QALY gained				
	Infliximab	Adalimumab	Golimumab	Conventional management	Colectomy
SA5: Base case using point estimates of parameters	dominated	dominated	dominated	dominated	dominating
SA6: Time horizon=20 years	dominated	dominated	dominated	dominated	dominating
SA7: Time horizon=10 years	dominated	dominated	dominated	dominated	dominating
SA8: Time horizon=5 years	dominated	dominated	dominated	-	£1,554
SA9: All utilities except post-surgical complications drawn from Swinburn *et al*[140†]	£179,374 [1]	£80,315 [2]	dominated [3]	ext dom [4]	-
SA10: Utilities of response/remission drawn from ACT1 trial[50] (0.88, 0.82 for remission and response respectively)	dominated	dominated	dominated	dominated	dominating
SA11: Utilities of response/remission drawn from PURSUIT-Maintenance trial[48] (0.89, 0.80 for remission and response respectively)	dominated	dominated	dominated	dominated	dominating
SA12: Relative risk of hospitalisation for golimumab vs conventional treatment assumed to be 1.0	dominated	dominated	dominated	dominated	dominating
SA13: UC health state resource use doubled	dominated	dominated	dominated	dominated	dominating
SA14: UC health state resource use halved	dominated	dominated	dominated	dominated	dominating
SA15: Probability of chronic pouchitis doubled	dominated	dominated	dominated	dominated	dominating
SA16: Probability of chronic pouchitis halved	dominated	dominated	dominated	dominated	dominating
SA17: Cost of surgery doubled	dominated	dominated	dominated	dominated	dominating
SA18: Cost of surgery halved	dominated	dominated	dominated	dominated	dominating
SA19: Probability of undergoing surgery in drug groups halved	dominated	dominated	dominated	dominated	dominating

Ext dom – extendedly dominated
** Excluded as adalimumab and golimumab have marketing authorisation in adult populations only*
† QALY rank shown in parentheses [ ]

The analyses indicate that the model results remain largely unaffected by changes in the model time horizon, assumed patient age, utilities for remission and response, assumptions regarding UC resource use, and the colectomy rate. The model is however very sensitive to assumptions regarding the relative utilities of remission, response, active UC and post-surgery. Within the sensitivity analysis in which utility values are drawn from Swinburn *et al*[140] (analysis number 11), the rank ordering of QALY gains is altered such that colectomy moves from being the most effective option to the least effective option. In this scenario, golimumab and conventional non-biologic treatments are expected to be ruled out of the analysis, the ICER for adalimumab versus colectomy is estimated to be approximately £80,315 per QALY gained whilst the ICER for infliximab versus adalimumab is estimated to be approximately £179,374 per QALY gained.

6.3 Budget impact analysis

This section presents an analysis of the expected net budget impact of introducing infliximab, adalimumab and golimumab for the treatment of moderate to severe UC in England and Wales. Budget impact estimates are presented annually for a 5-year period. The analysis makes the following assumptions:

- The prevalence of UC is 240 per 100,000 population
- The incidence of UC is 10 per 100,000 population
- The population of England and Wales is approximately 56million
- 14.5% of all UC patients would be eligible to receive biologic treatments[80]
- All patients who are eligible for treatment with biologics will receive these therapies
- Discounting is not applied

These assumptions suggest an eligible prevalent UC cohort of approximately 19,488 patients and an eligible incident cohort of approximately 812 patients per year. Based on the cost distribution over time estimated for each treatment using the Assessment Group model, combined with the estimated eligible prevalent and incident cohorts in each year, this gives rise to the budget impact estimates presented in Table 84. Assuming full uptake of these drugs, the estimated net budget impact of infliximab, adalimumab and golimumab for the treatment of moderate to severe UC is estimated to be between £269million and £434million.

Table 84: Estimated absolute and net budgetary impact of introducing biologics for the treatment of moderate to severe UC in England and Wales

Year	Infliximab	Adalimumab	Golimumab	Conventional non-biologic treatments
Absolute budget impact for each treatment				
Year 1	£320,110,846	£212,330,484	£274,167,240	£74,464,026
Year 2	£144,962,276	£113,527,742	£129,696,877	£68,924,692
Year 3	£123,381,870	£107,709,058	£108,800,972	£73,417,250
Year 4	£112,139,838	£104,975,394	£100,118,987	£76,740,469
Year 5	£106,761,267	£103,549,916	£97,224,320	£79,522,798
Net budget impact for costs of biologic less costs of conventional treatments				
Year 1	£245,646,820	£137,866,457	£199,703,214	-
Year 2	£76,037,584	£44,603,050	£60,772,185	-
Year 3	£49,964,620	£34,291,808	£35,383,722	-
Year 4	£35,399,369	£28,234,925	£23,378,518	-
Year 5	£27,238,469	£24,027,118	£17,701,522	-
Total cost over 5 years	£434,286,862	£269,023,359	£336,939,161	

6.4 Discussion

The manufacturers of adalimumab, infliximab and golimumab submitted economic models to assess the cost-effectiveness of biologic therapies versus conventional treatment. The MSD infliximab submission model indicates that the estimated ICER for infliximab versus standard non-biologic treatment (colectomy) is £37,682 per QALY gained.[65] The MSD golimumab submission reports an estimated ICER of £27,322 per QALY gained.[67] The AbbVie submission reports a base case ICER of £34,590 per QALY gained.[63] The Assessment Group scrutinised these models and critiqued the evidence and assumptions which underpin the cost-effectiveness estimates reported by the manufacturers. A number of problems with these models were identified by the Assessment Group, particularly with respect to the exclusion of relevant treatment options specified in the final NICE scope[41] and the use of a short time horizon. In addition, the MSD model does not include a fully incremental analysis, confuses evidence from populations with varying degrees of UC severity and inadequately reflects likely UK treatment pathways. As a consequence of these problems, the Assessment Group do not believe that the cost-effectiveness evidence produced by either manufacturer adequately addresses the specified decision problem.

In light of the problems with the manufacturers' submitted economic analyses, the Assessment Group developed a *de novo* cost-effectiveness model to assess infliximab, adalimumab, golimumab, conventional non-biologic treatments and elective surgery within the moderate to severe UC population. The Assessment Group model differs from both the manufacturers's models in that all relevant medical and surgical options are evaluated over a lifetime horizon, as specified in the final NICE scope. Underpinning the Assessment Group

model is a series of complex network meta-analyses which synthesise all relevant evidence relating to infliximab, adalimumab, golimumab and conventional non-biologic therapies. A summary of the key differences between the Assessment Group model and the manufacturers' models is presented in Table 85.

Table 85: Summary of key differences between the Assessment Group model and the manufacturers' models

Element of evaluation	Assessment Group model (base case)	MSD models[65,67]	AbbVie model[63]
Options evaluated	(i) infliximab (ii) adalimumab (iii) golimumab (iv) conventional non-biologic treatments (v) colectomy	(i) infliximab (ii) adalimumab (iii) golimumab (iv) conventional non-biologic treatments	(i) adalimumab (ii) conventional non-biologic treatments
Time horizon	Lifetime	10 years	10 years
Source of efficacy evidence	NMA using unpublished ordered categorical data	NMA using published binomial data (includes manipulation of PURSUIT-Maintenance trial data)	Unpublished data from ULTRA2 and ULTRA1/2 extension study supplemented using estimates from Kane *et al*[134] and Odes *et al*[120]
Treatment options following failure of biologic	Conventional non-biologic treatments and possible colectomy	Relapse management and imminent colectomy	Conventional non-biologic treatments and possible colectomy
Possible transitions between active UC states (remission, response, no response)	All transitions in matrix allowed	Patients losing remission transit to response, patients achieving response cannot achieve remission	All transitions in matrix allowed
Source of health utilities	Woehl *et al*[160] (chronic pouchitis valued using Arseneau *et al*[110])	ACT1/PURSUIT-Maintenance, Woehl *et al*,[160] complications valued using Arseneau *et al*[110,161]	Swinburn *et al*,[63] Tsai *et al*,[82] complications valued using weighted average of Arseneau *et al*,[110] Hu *et al*[141] and Smith and Roberts[142]

The base case analysis of the Assessment Group model suggests that within an adult UC population, colectomy is expected to produce 14.72 discounted QALYs at a discounted cost of approximately £41,900 over the patient's remaining lifetime. All medical options are expected to produce substantially fewer QALYs at a greater cost than colectomy, hence colectomy is expected to dominate infliximab, adalimumab, golimumab and conventional non-biologic treatments. Importantly however, elective colectomy may not be considered an

acceptable or preferable option for some proportion of patients. In circumstances whereby only drug options are considered acceptable, the base case version of the Assessment Group model suggests that infliximab and golimumab are expected to be ruled out due to dominance, whilst the incremental cost-effectiveness of adalimumab versus conventional non-biologic treatment is expected to be approximately £50,600 per QALY gained.

The Assessment Group also undertook a separate probabilistic economic analysis of infliximab, conventional non-biologic treatments and colectomy within a paediatric population (mean age=15 years). Where colectomy is an acceptable treatment option, the economic analysis suggests that this option dominates infliximab and conventional non-biologic treatments. Where colectomy is not an acceptable option, the economic analysis suggests that the incremental cost-effectiveness of infliximab versus conventional treatments is approximately £68,400 per QALY gained. This analysis is however based on adult efficacy evidence hence it should be interpreted with some degree of caution.

Three separate probabilistic sensitivity analyses were undertaken using data from the Japanese trial reported by Suzuki *et al*[47] and using the ITT data rather than anti-TNF-α naïve patients from ULTRA2.[46,126] Across these three scenarios, the general conclusions of the economic analysis remain unchanged. The Assessment Group also undertook separate comparisons of (i) infliximab versus colectomy and conventional treatments, (ii) adalimumab colectomy and versus conventional treatments and (iii) golimumab versus colectomy and conventional treatments using the head-to-head trials rather than the NMA models. These analyses indicate that where colectomy is an acceptable option, it is expected to dominate the drug options. Where colectomy is not an acceptable option, the ICERs produced from these analyses are all in excess of £70,000 per QALY gained.

A number of simple sensitivity analyses were also undertaken using the point estimates of model parameters. Across these scenarios, the model results appear to be insensitive to changes in these assumptions, with the exception of the HRQoL values assumed. Within the scenario whereby utilities from Swinburn *et al*[63] are used in the model (rather than those reported by Woehl *et al*[162] as per the base case analysis), colectomy produces the <u>lowest</u> QALY gain and conventional management and golimumab are ruled out due to extended dominance. Within this scenario, the incremental cost-effectiveness of adalimumab versus elective colectomy is estimated to be £80,315 per QALY gained, whilst the incremental cost-effectiveness of infliximab versus adalimumab is estimated to be £179,374 per QALY gained. Whilst these results are very different to the Assessment Group's base case analysis, the economic conclusions that should be drawn from this sensitivity analysis are not.

7. ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES

7.1 Surgery and patient choice

Surgery may be considered as an option for patients with UC for a number of indications including due to complications of disease, perceived risk of or identified dysplasia/neoplasia or due to lack or loss of efficacy of medical treatments. For a proportion of patients without emergency symptoms, surgery may not represent an acceptable treatment option.

7.2 Administration route – i.v. infusions versus s.c. injection

The method of drug administration differs between the interventions included in this appraisal. Infliximab is given via i.v. infusion whilst adalimumab and golimumab are administered via subcutaneous injection. Infliximab therefore requires outpatient attendances, additional nursing care and monitoring. These resources are not required for the administration of adalimumab or golimumab. Pre-infusion prophylaxis may be required to minimise the risk of infusion reactions associated with infliximab.

7.3 Training for subcutaneous injections

Where considered appropriate by the physician, patients, family members and/or carers require training for the administration of subcutaneous injections. This training is associated with additional resource use and costs.

7.4 Screening for TB and other infectious diseases (e.g. hepatitis B)

All of the biologic therapies considered in this assessment report may be associated with the reactivation of TB. Patients should be screened for TB (and other infectious disease e.g. hepatitis B) before initiation of treatment.

7.5 Off-license use in children for golimumab and adalimumab

Currently, infliximab is the only biologic treatment option that is licensed for the treatment of moderate to severe UC in children in the UK.

8. DISCUSSION

8.1 Statement of principal findings

8.1.1 Principal findings – clinical effectiveness

A total of ten RCTs were identified in the clinical effectiveness systematic review, of which nine related to adults and one was conducted in a paediatric population. All of the adult RCTs were performed against placebo (with the exception of UC-SUCCESS) and were a maximum of one year in study duration. No head-to-head RCTs were identified in which interventions of interest were assessed against each other.

The risks of bias associated with included RCTs were assessed using the Cochrane risk of bias instrument. Only three RCTs could be considered as being at overall low risk of bias (as allocation concealment, blinded outcome assessment and completeness of outcome data were all judged as low risk). It should be noted that one of the maintenance trials (PURSUIT-Maintenance) re-randomised patients who had previously responded to golimumab induction therapy in two previous trials; the extent of this potential bias on patient outcomes is unclear.

The outcome measures pre-specified in the final NICE scope were all addressed by the included trial evidence, with the exception of rates of relapse. Clinical response and remission data based on complete Mayo scores were well reported across trials. Evidence was identified to demonstrate that patients receiving infliximab, adalimumab or golimumab were more likely than patients receiving placebo to achieve clinical response and remission at induction and maintenance time points. Patients in the UC-SUCCESS trial who received combination treatment with infliximab and azathioprine experienced the most favourable rates of steroid-free remission compared with infliximab and azathioprine treatment groups. Seven RCTs performed in adult populations contributed data on clinical response and remission at induction or maintenance time points to network meta-analyses.

Based on the NMA, in the induction phase, all treatments were associated with statistically significant beneficial effects relative to placebo with the greatest effect being associated with infliximab.Infliximab was also associated with the highest probability of moving from no response to response and from no response to remission. The effects of adalimumab and golimumab on these two probabilities were broadly comparable.

For patients classified as responders at the end of the induction phase, treatment effects were not statistically significant, although the greatest effect was associated with golimumab 100mg at 8-32 weeks. Golimumab 100mg was associated with the highest probability of

moving from response to remission and staying in response, and the smallest probability of moving from response to no response. However, at 32-52 weeks, only infliximab and golimumab 50mg were associated with beneficial effects on clinical response, although the effects were not statistically significant. Infliximab was associated with the highest probability of moving from response to remission and staying in response, and the smallest probability of moving from response to no response at 32-52 weeks. The probabilities of staying in response were comparable among treatments at both 8-32 weeks and 32-52 weeks.

For patients classified as being in remission at the end of the induction phase, all treatments except for adalimumab were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 50mg and 100mg, although the effects were not statistically significant at 8-32 weeks. Golimumab 50mg and 100mg were associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response and from remission to no response. At 32-52 weeks, all treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo, with the greatest effect being associated with adalimumab (the only treatment with statistically significant effect). Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response and from remission to no response.

Available data on hospitalisation outcomes were very limited, but suggested that outcomes may be more favourable for adalimumab-treated and infliximab-treated patients compared with placebo (with no data available from golimumab trials). Data on surgical intervention were also very sparse, with a potential inconclusive benefit for intervention-treated patients compared with placebo. No trials reported whether surgical outcomes were elective or emergency in nature. However, more data are required to demonstrate the impact of interventions on hospitalisation and surgical intervention more conclusively. Data were available from a single trial to support the use of infliximab in induction and maintenance treatment in a paediatric population.

The main safety issues highlighted in the RCT evidence appeared to be generally consistent with those previously discussed in the respective Summary of Product Characteristics (including serious infections, malignancies and administration site reactions). Deaths occurring during and after the study period were described in some trials evaluating golimumab (PURSUIT-Maintenance) and infliximab (ACT trials), of which infection or malignancy commonly appeared to be implicated. This underlines the importance of

monitoring and treating serious infections and malignancies in patients receiving immunosuppressive treatment.

The trials included in the clinical effectiveness systematic review typically excluded patients with ulcerative proctitis, patients with fulminant/acute severe disease, those with a history of or at imminent risk of bowel surgery, pregnant or lactating women, and patients with diseases of the central nervous system (e.g. demyelinating disease). Furthermore, patients with history of serious infection and/or immunodeficiency were also typically excluded, as were individuals with a history of malignancy or signs of dysplasia. Therefore, the effects of adalimumab, golimumab or infliximab in these UC populations are unknown.

Two biosimilars (Remsima and Inflectra) to Remicade were considered as part of the evidence base for infliximab as part of this assessment. The sponsor submission received from the manufacturers of Remsima (Celltrion) and the EPAR reports for Remsima and Inflectra indicated that both biosimilars were approved by the EMA on the basis of reported similar pharmacokinetic and efficacy (demonstrated in AS and RA patients) profiles to Remicade. No further trials of Remsima or Inflectra were identified in the course of this assessment.

8.1.2 Principal findings – cost-effectiveness

The manufacturers of adalimumab, infliximab and golimumab submitted economic models to assess the cost-effectiveness of biologic therapies versus conventional treatment. The MSD infliximab submission model indicates that the estimated ICER for infliximab versus standard non-biologic treatment (colectomy) is £37,682 per QALY gained.[65] The MSD golimumab submission reports an estimated ICER of £27,322 per QALY gained.[67] The AbbVie submission reports a base case ICER of £34,590 per QALY gained.[63]

The Assessment Group identified several issues with the manufacturers' submitted models, in particular, the exclusion of relevant treatment options specified in the final NICE scope[41] and the use of a short time horizon. Given the missing comparators within each of the manufacturers' submitted economic analyses, it is unclear how these models should be used to inform NICE decision-making.

The Assessment Group developed a *de novo* cost-effectiveness model to assess infliximab, adalimumab, golimumab, conventional non-biologic treatments and elective surgery within the moderate to severe UC population. The base case analysis of the Assessment Group model suggests that within an adult UC population, colectomy is expected to produce 14.72

discounted QALYs at a discounted cost of approximately £41,900 over the patient's remaining lifetime. All medical options are expected to produce substantially fewer QALYs at a greater cost than colectomy, hence colectomy is expected to dominate infliximab, adalimumab, golimumab and conventional non-biologic treatments. Where colectomy is not be considered an acceptable option, the base case analysis of the Assessment Group model suggests that infliximab and golimumab are expected to be ruled out due to dominance, whilst the incremental cost-effectiveness of adalimumab versus conventional non-biologic treatment is expected to be approximately £50,600 per QALY gained.

The Assessment Group also undertook a separate probabilistic economic analysis of infliximab, conventional non-biologic treatments and colectomy within a paediatric population (mean age=15 years). Where colectomy is an acceptable treatment option, the economic analysis suggests that this option dominates infliximab and conventional non-biologic treatments. Where colectomy is not an acceptable option, the economic analysis suggests that the incremental cost-effectiveness of infliximab versus conventional treatments is approximately £68,400 per QALY gained.

The results of the Assessment Group model were largely insensitive to changes in model parameter values with the exception of the HRQoL values for each state. The use of utility estimates from Swinburn *et al*[63] results in a situation whereby colectomy produces the lowest QALY gain and conventional management and golimumab are ruled out due to extended dominance. Within this scenario, the incremental cost-effectiveness of adalimumab versus elective colectomy is estimated to be £80,315 per QALY gained, whilst the incremental cost-effectiveness of infliximab versus adalimumab is estimated to be £179,374 per QALY gained.

8.2 Strengths and limitations of the assessment

The systematic review was based on rigorous methods, with comprehensive searches for evidence, a good level of consistency between reviewers in study selection and double checking of data extraction. Clinical response and remission data were widely reported across included trials and study authors were consistent in their use of the complete Mayo score, which aided the comparison of trials. Whilst no head-to-head RCT evidence was available, network meta-analyses were performed to permit a comparison of the efficacy of interventions in terms of clinical response and remission.

The economic analysis presented by the Assessment Group has several strengths:

- The treatment pathway represented within the model was based on considerable expert opinion of several leading UC experts

- The Assessment Group model is underpinned by a complex network meta-analysis across all drug options thereby synthesising relevant efficacy outcomes data within a single network of evidence
- The model generally adheres to NICE's Reference Case and fully addresses the decision problem set out in the final NICE scope
- Where appropriate and possible, systematic search methods have been used to identify, select and use evidence to inform the model's parameters (efficacy, HRQoL and colectomy rates)
- The Assessment Group have undertaken extensive sensitivity analyses to examine the impact of alternative assumptions and sources of evidence on the robustness of the results of the model.

The Assessment Group model is also subject to a number of limitations:

- There is considerable uncertainty associated with Assessment Group's extrapolation of short-term trial data (maximum 54 weeks) to a lifetime horizon
- The model does not consider an explicit sequential pathway of non-biologic treatments; rather during any cycle, a proportion of patients are assumed to receive 5-ASAs, immunomodulators and prednisolone.
- Evidence relating to complications of colectomy was identified through consideration of approaches used within previous models rather than through a full systematic review; these assumptions were however tested within the sensitivity analyses

8.3 Uncertainties

Key uncertainties in this assessment include:

- the optimal duration of intervention treatment in responding patients
- the maintenance of efficacy outcomes and safety of interventions beyond the limited study lengths available
- the maintenance of outcomes in responding patients following cessation of anti-TNF-α treatment

9. CONCLUSIONS

Based on the NMA for clinical response and remission, in the induction phase, all treatments were associated with statistically significant beneficial effects relative to placebo, with the greatest effect being associated with infliximab. For patients classified as responders at the end of the induction phase, treatment effects were not statistically significant, although the greatest effect at 8-32 weeks was associated with golimumab 100mg. At 32-52 weeks, only infliximab and golimumab 50mg were associated with beneficial effects on clinical response. For patients classified as being in remission at the end of the induction phase, all treatments except for adalimumab were associated with beneficial treatment effects relative to placebo with the greatest effect being associated with golimumab 50mg and 100mg, although the effects were not statistically significant at 8-32 weeks. At 32-52 weeks, all treatments except golimumab 50mg were associated with beneficial treatment effects relative to placebo, with the greatest effect being associated with adalimumab (the only treatment with statistically significant effect). Adalimumab was associated with the highest probability of staying in remission, and the smallest probability of moving from remission to response and from remission to no response.

Available data on hospitalisation outcomes were very limited, but suggested that outcomes may be more adalimumab-treated and infliximab-treated patients compared with placebo (with no data available from golimumab trials). Data on surgical intervention were also very sparse, with a potential inconclusive benefit for intervention-treated patients compared with placebo. No trials reported whether surgical outcomes were elective or emergency in nature. However, more data are required to demonstrate the impact of interventions on hospitalisation and surgical intervention more conclusively. Data were available from a single trial to support the use of infliximab in induction and maintenance treatment in a paediatric population.

The main safety issues highlighted in the RCT evidence appeared to be generally consistent with those previously discussed in the respective Summary of Product Characteristics (including serious infections, malignancies and administration site reactions). Deaths occurring during and after the study period were described in some trials evaluating golimumab (PURSUIT-Maintenance) and infliximab (ACT trials), of which infection or malignancy commonly appeared to be implicated. This underlines the importance of monitoring and treating serious infections and malignancies in patients receiving immunosuppressive treatment.

The base case analysis of the Assessment Group model suggests that within an adult UC population, colectomy is expected to dominate infliximab, adalimumab, golimumab and

conventional non-biologic treatments. Where elective colectomy is not an acceptable option, the Assessment Group model suggests that infliximab and golimumab are expected to be ruled out due to dominance, whilst the incremental cost-effectiveness of adalimumab versus conventional non-biologic treatment is expected to be approximately £50,600 per QALY gained. The base case analysis of the Assessment Group model suggests that within a paediatric UC population, colectomy is expected to dominate infliximab and conventional non-biologic treatments. Where colectomy is not an acceptable option, the incremental cost-effectiveness of infliximab versus conventional treatments is approximately £68,400 per QALY gained.

9.1 Implications for service provision

The Assessment Group is unaware of any further implications for service provision beyond those addressed in Chapter 7 of this report.

9.2 Suggested research priorities

- Surgical intervention and hospitalisation to be incorporated as outcomes in future RCTs and associated extension studies of interventions in the treatment of moderately to severely active UC after the failure of conventional therapy
- Head-to-head RCTs of interventions under assessment against each other in the treatment of moderate to severe UC after the failure of conventional therapy
- RCTs of longer duration of follow-up to assess maintenance of outcomes over the longer term
- Assessment of maintenance of outcomes in responding patients following cessation of anti-TNF-α treatment
- Assessment of efficacy, safety and immunogenicity following reintroduction of interventions after interruption in treatment
- Assessment of efficacy of interventions under assessment in specific subgroups (e.g. according to disease duration, as specified in the appraisal scope)
- RCTs evaluating use of interventions under assessment in biologic switching (i.e. after failure of first anti-TNF-α agent)
- Consideration of unified universally agreed primary end points in future UC RCTs
- Further exploration of comparative clinical and economic outcomes associated with medical versus surgical treatments for patients with moderate to severe UC
- Definition of factors that predict an improved patient response to anti-TNF-α treatment

10. APPENDICES

Appendix 1: Final Protocol

Technology Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence

Final protocol 22nd November 2013

1. Title of the project:

Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262)

2. Name of TAR team and 'lead'

TAR team
School of Health and Related Research Technology Assessment Group,
University of Sheffield
Project lead
Rachel Archer
Research Fellow
School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA
Tel: (+44) (0)114 222 0793
Fax: (+44) (0)114 272 4095
Email: r.archer@sheffield.ac.uk

3. Plain English Summary

Ulcerative colitis is recognised as the most common form of inflammatory bowel disease in the UK, having an incidence of approximately 10 per 100,000 per year and a prevalence of approximately 240 per 100,000.[1] Peak incidence is between 15 and 25 years of age, with a potential second peak between 55 and 65 years.[1] The majority (approximately 80%) of incident cases are reported to be of mild or moderate severity. An estimated 132,600 people in England and Wales have been diagnosed with ulcerative colitis. It is a chronic disease of unknown cause with symptoms including the development of bloody diarrhoea, abdominal

pain, weight loss, fatigue, anaemia and an urgent need to defecate. Symptoms may vary according to the degree and severity of bowel inflammation. The condition has no current cure and the disease course is relapsing-remitting in pattern. A range of factors have been suggested as potentially influencing the risk of relapse.[2] There is evidence to indicate that severity of disease may be associated with younger age at diagnosis.[3,4] Complications of ulcerative colitis include primary sclerosing cholangitis (inflamed and damaged bile ducts), bowel cancer, osteoporosis and toxic megacolon (swelling of colon due to trapped gases). The aim of clinical management is to induce and maintain disease remission and to avoid potential complications and surgical intervention.[5]

4. Decision problem

4.1 Purpose of the decision to be made

This assessment will address the question "what is the clinical effectiveness and cost-effectiveness of infliximab, adalimumab and golimumab in the treatment of moderately to severely active ulcerative colitis after the failure of conventional therapy as compared against each other and standard clinical management?"

4.2 Clear definition of interventions

Three interventions will be considered within this assessment. Infliximab, adalimumab and golimumab are monoclonal antibodies which inhibit the activity of TNF-α.

<u>(1) Infliximab (Remicade, Merck Sharp and Dohme)</u>

Infliximab has a UK marketing authorisation for the treatment of moderately to severely active ulcerative colitis in adults, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[6]

Infliximab also has a UK marketing authorisation for the treatment of severely active ulcerative colitis in children and adolescents aged 6 to 17 years, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[6]

Infliximab for the treatment of ulcerative colitis is administered by intravenous infusion at a dosage of 5 mg/kg followed by additional 5 mg/kg infusion doses at 2 and 6 weeks after the initial infusion, then every 8 weeks thereafter.[6]

Biosimilar versions of infliximab (Remsima, Celltrion Healthcare; Inflectra, Hospira) are also licensed for the same indications. These will also be included as part of the evidence base for infliximab in this assessment.

(2) Adalimumab (Humira, AbbVie)

Adalimumab has a UK marketing authorisation for the treatment of moderately to severely active ulcerative colitis in adults who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[7]

Adalimumab for the treatment of ulcerative colitis is administered subcutaneously according to an induction dose regimen of 160 mg at Week 0 and 80 mg at Week 2 followed by a recommended maintenance dosage of 40 mg every other week (increased to 40 mg every week if clinical response is insufficient).[7]

(3) Golimumab (Simponi, Merck Sharp and Dohme)

Golimumab has a UK marketing authorisation for the treatment of moderately to severely active ulcerative colitis in adults who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.[8]

Golimumab for the treatment of ulcerative colitis is administered subcutaneously according to body weight. Patients with body weight less than 80 kg receive an initial dose of 200 mg, followed by 100 mg at week 2, then 50 mg every 4 weeks, thereafter. Patients with body weight greater than or equal to 80 kg receive an initial dose of 200 mg, followed by 100 mg at week 2, then 100 mg every 4 weeks, thereafter.[8]

4.3 Place of the intervention in the treatment pathway(s)

As outlined in the final scope and NICE clinical guideline 166 ('Ulcerative colitis: Management in adults, children and young people'),[1] conventional treatment options for moderately to severely active (non-systemic) ulcerative colitis include the use of oral or topical aminosalicylates, corticosteroids and/or immunosuppressants (NB: Some conventional treatment options did not have marketing authorisation at the time of clinical guideline publication [June 2013]). Recommended conventional treatment options may vary according

to the extent and location of colitis.[1] Colectomy may be considered in the event of inadequate control of symptoms and/or poor quality of patient life on conventional treatment.

Infliximab, adalimumab and golimumab will be assessed in this current technology assessment in line with licensed indications as treatment options for moderately to severely active ulcerative colitis after the failure of conventional therapy.
Infliximab was not previously recommended by NICE for the treatment of "subacute" manifestations of moderately to severely active ulcerative colitis (NICE technology appraisal guidance 140).[9] NICE technology appraisal 262 (adalimumab for the treatment of moderately to severely active ulcerative colitis) was terminated as no evidence submission was provided by the manufacturer.[10]

4.4 Relevant comparators

Interventions may be compared against each other. Other relevant comparators include standard clinical management options, which, as described in the final scope, may include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors or elective surgical intervention.

Emergency surgical intervention will not be considered as a comparator in this assessment (as acute severe ulcerative colitis is stated in the final scope as being outside the remit of this assessment).

4.5 Population and relevant sub-groups

The assessment will consider the following two populations:

(1) Adults aged 18 years and over with moderately to severely active ulcerative colitis who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.

It is anticipated that severity of disease in adults will be defined according to the modified Truelove and Witts' severity index (1955) (as referred to in the final NICE scope and as categorised and tabulated in NICE clinical guideline 166).[1]

The following interventions are indicated for use in adults:

- Adalimumab
- Infliximab
- Golimumab

(2) Children and adolescents aged 6 to 17 years (inclusive) with severely active ulcerative colitis, who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies.

It is anticipated that the severity of ulcerative colitis in children and adolescents will be made using the Paediatric Ulcerative Colitis Activity Index (PUCAI) (as categorised and tabulated in NICE clinical guideline 166).[1]

The following intervention is indicated for use in children and adolescents:

- Infliximab

Specific subgroups and treatment effect modifiers of interest include duration of disease, as specified in the final scope.

4.6 Key factors to be addressed

The objectives of the assessment are to:

- evaluate the clinical effectiveness of each intervention
- evaluate the adverse effect profile of each intervention
- evaluate the incremental cost-effectiveness of each intervention compared (i) against each other and (ii) against all comparators
- estimate the overall NHS budget impact in England and Wales

4.7 Factors that are outside the scope of the appraisal

The evaluation of interventions in the following groups are outside of the appraisal scope and will not be considered in this assessment:

- Children with mildly or moderately active ulcerative colitis (as defined by the PUCAI measure)
- Adults with mildly active ulcerative colitis (as defined by the modified Truelove and Witts' [1955] criteria)
- Adults and children with acute severe (systemic) ulcerative colitis

5. Methods for the synthesis of evidence of clinical effectiveness

A systematic review of the evidence for clinical effectiveness will be undertaken following the general principles outlined in 'Systematic Reviews: CRD's guidance for undertaking reviews in health care'[11] and the principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (http://www.prisma-statement.org/).[12]

5.1. Search strategy

A comprehensive search will be undertaken to systematically identify clinical effectiveness literature relating to infliximab, adalimumab and golimumab within their licensed indications for the treatment of moderately to severely active ulcerative colitis after the failure of conventional therapy.

The search strategy will comprise the following main elements:

- Searching of electronic databases
- Contact with experts in the field
- Scrutiny of bibliographies of retrieved papers

Search strategies will be used to identify relevant trials (as specified under the inclusion criteria below) and systematic reviews/meta-analyses (for the identification of additional trials). The following databases will be searched:

- MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)
- Embase (Ovid)
- Cochrane Database of Systematic Reviews (Wiley Interscience)
- Cochrane Central Register of Controlled Trials (Wiley Interscience)
- Cumulative Index to Nursing and Allied Health Literature (EBSCO)
- Science Citation Index (ISI Web of Knowledge)
- Social Sciences Citation Index (ISI Web of Knowledge)
- BIOSIS (Web of Knowledge)
- Centre for Reviews and Dissemination Database of Abstracts of Reviews of Effectiveness and Health Technology Assessment (CRD DARE and HTA)

Current research registers (e.g. UK Clinical Research Network Portfolio Database, ClinicalTrials.gov) will also be searched for ongoing and recently completed research projects. Citation searches of key included studies will also be undertaken using the Web of Science Citation Index Expanded and Conference Proceedings Citation Index - Science.

Searches will not be restricted by language or date or publication type. The MEDLINE search strategy is presented in Appendix 1. High precision search filters designed to retrieve clinical trials and systematic reviews will be used on MEDLINE and other databases, where appropriate. The search will be adapted for other databases. Industry submissions and relevant systematic reviews will also be handsearched in order to identify any further relevant clinical trials. A comprehensive database of relevant published and unpublished articles will be constructed using Reference Manager bibliographic software, (version 12.0; Thomson Reuters, Philadelphia, PA).

5.2 Inclusion and exclusion criteria

5.2.1 Inclusion criteria

Inclusion criteria have been defined in line with the final scope provided by NICE and are outlined below.

5.2.1.1 Populations

(1) Adults aged 18 years and over with moderately to severely active (non-systemic) ulcerative colitis (defined as patients with moderately active disease according to the modified Truelove and Witts' criteria [1955] only) whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant of or have medical contraindications to such therapies.

ii) Children aged 6 to 17 years with severely active (non-systemic) ulcerative colitis (as classified by the PUCAI measure) whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant of or have medical contraindications to such therapies.

5.2.1.2 Interventions

For adults (defined by the Assessment Group as aged 18 years and over):

- Adalimumab
- Infliximab
- Golimumab

For children and adolescents aged 6 to 17 years (inclusive):

- Infliximab

Interventions will be assessed in line with licensed indications.

5.2.1.3 Comparators

Interventions may be compared with each other. Interventions will be compared with standard clinical management, which may include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclomethasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors or elective surgical intervention.

5.2.1.4 Outcomes

The outcome measures to be considered include:

- Mortality
- Measures of disease activity
- Rates of and duration of response, relapse and remission
- Rates of hospitalisation
- Rates of surgical intervention (both elective and emergency)
- Time to surgical intervention (both elective and emergency)
- Adverse events of treatment (including leakage and infections following surgery)
- Health-related quality of life

Mucosal healing will not be included as an outcome in this assessment.

5.2.1.5 Study design

Randomised controlled trials (RCTs) will be included in the clinical effectiveness systematic review. If no RCTs are identified for an intervention, non-randomised studies may be considered for inclusion. Non-randomised studies may also be included, where necessary, as a source of additional evidence (e.g. relating to adverse events, long-term effectiveness etc) associated with the interventions.

Studies published as abstracts or conference presentations will only be included if sufficient details are presented to allow an assessment of the methodology and results to be undertaken.

5.2.2 Exclusion criteria

The following types of studies will be excluded:

- Studies which include adults with mildly active ulcerative colitis (as defined by the modified Truelove and Witts' [1955] criteria)
- Studies which include children with mildly or moderately active ulcerative colitis (as defined by the PUCAI measure)

- Studies which include adults with severely active ulcerative colitis as defined by the modified Truelove and Witts' [1955] criteria (representing patients who are systemically ill and are excluded as being outside the remit of this appraisal)
- Studies which include adults, adolescents or children with acute severe ulcerative colitis, whose disease is systemic (as shown by tachycardia, fever, anaemia or a raised erythrocyte sedimentation rate) (representing patients who are excluded as being outside the remit of this appraisal)
- Studies which include patients with inflammatory bowel disease other than ulcerative colitis (e.g. Crohn's disease)
- Studies where interventions are administered not in accordance with licensed indications
- Systematic reviews and clinical guidelines (these may be used as sources of references)
- Studies which are considered methodologically unsound in terms of study design or the method used to assess outcomes
- Studies which are only published in languages other than English
- Studies based on animal models
- Preclinical and biological studies
- Narrative reviews, editorials, opinions
- Reports published as abstracts or conference presentations only, where insufficient details are reported to allow an assessment of study quality or results.

Trials retrieved for full paper screening which are subsequently excluded will be listed in an appendix to the report with reasons justifying their exclusion.

5.2.3 Study selection

Retrieved studies will be selected for inclusion according to the inclusion and exclusion criteria specified in Sections 5.2.1 and 5.2.2. Studies will be assessed for relevance first by title/abstract, and then finally by full text, excluding at each step studies which do not satisfy the inclusion criteria. One reviewer will examine titles and abstracts for inclusion, and a second reviewer will check at least 10% of citations. Full manuscripts of selected citations will be retrieved and assessed by one reviewer against the inclusion and exclusion criteria. Discrepancies will be resolved by discussion, with involvement of a third team member when necessary.

5.3 Data extraction strategy

Data will be extracted by one reviewer using a standardised data extraction form. A draft data extraction form is presented in Appendix 2. Data will be extracted with no blinding to authors or journal. Where multiple publications of the same study are identified, data will be extracted and reported as a single study. A second reviewer will check at least 10% of data extraction forms. Discrepancies will be resolved by discussion. The Assessment Group's approach to handling data obtained from the manufacturers' submissions is detailed in Section 7.

5.4 Quality assessment strategy

The methodological quality of each included RCT will be assessed using the Cochrane Risk of Bias tool[13] or (adapted) criteria based on those proposed by the NHS Centre for Reviews and Dissemination for RCTs.[11] The purpose of such quality assessment is to provide a narrative account of trial quality for the reader and, where meta-analysis is appropriate, to inform potential exclusions from any sensitivity analysis. Each included study will be quality assessed by one reviewer and a second reviewer will check at least 10% of quality assessment forms.

5.5. Methods of analysis/synthesis

Pre-specified outcomes will be tabulated and discussed in a narrative synthesis.

If considered appropriate, meta-analysis may be carried out using fixed and/or random effects models using the Cochrane Collaboration Review Manager© software (version 5.1). Heterogeneity may be explored through consideration of the study populations, methods, and interventions and, in statistical terms, by the χ^2 test for homogeneity and the I^2 statistic. If appropriate, a simultaneous comparison of all interventions will be performed. This will be done using a random effects network meta-analysis assuming that the trials form a connected network of evidence. Network meta-analyses will be implemented using the freely available software WinBUGS 1.4.3.

5.6 Methods for estimating quality of life

Health-related quality of life (HRQoL) data available from studies included in the clinical effectiveness systematic review will be extracted. In the absence of such evidence, the mathematical model may use evidence on HRQoL drawn from alternative sources.

6. Methods for synthesising evidence of cost-effectiveness

6.1 Identifying and systematically reviewing published cost-effectiveness studies

A comprehensive search will be undertaken to systematically identify cost-effectiveness literature relating to infliximab, adalimumab and golimumab within their licensed indications for the treatment of moderately to severely active ulcerative colitis after the failure of conventional therapy.

The search strategy will comprise the following main elements:

- Searching of electronic databases
- Contact with experts in the field
- Scrutiny of bibliographies of retrieved papers

Search strategies will be used to identify relevant economic papers.

The following databases will be searched:

- MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)
- Embase (Ovid)
- Cumulative Index to Nursing and Allied Health Literature (EBSCO)
- Science Citation Index (ISI Web of Knowledge)
- Social Sciences Citation Index (ISI Web of Knowledge)
- Centre for Reviews and Dissemination Database of Abstracts of Reviews of Effectiveness, Health Technology Assessment and NHS Economic Evaluations Database (CRD DARE, HTA and EED)
- EconLit (Ovid)
- BIOSIS (Web of Knowledge)

Citation searches of key included studies will also be undertaken using the Web of Science Citation Index Expanded and Conference Proceedings Citation Index - Science.

Searches will not be restricted by language or date or publication type. The MEDLINE search strategy is presented in Appendix 1. High precision search filters designed to identify existing economic evaluations of interventions for the treatment of moderately to severely active ulcerative colitis will be used on MEDLINE and other databases, where appropriate. The search will be adapted for other databases. A comprehensive database of relevant published

and unpublished articles will be constructed using Reference Manager bibliographic software, (version 12.0; Thomson Reuters, Philadelphia, PA).

Additional searches, for example to inform the decision-analytic model, where required in the course of the project, will be undertaken through consultation between the team.

Any existing health economic analyses identified by the searches will be critically appraised using published checklists.[14,11] In addition, any economic analyses presented in the sponsor submissions to NICE will also be critically appraised using these checklists. Existing cost-effectiveness analyses may also be used to identify sources of evidence to inform structural assumptions and parameter values for the Assessment Group model.

6.2 Development of a *de novo* economic model

A *de novo* economic evaluation will be undertaken from the perspective of the UK NHS and Personal Social Services (PSS). The model will draw together evidence concerning treatment efficacy, withdrawal, treatment-related adverse events, relevant imaging/diagnostic interventions, chronic care costs, and HRQoL. Costs on drug acquisition, administration, hospitalisation, adverse events and primary care will be identified through literature searches and national formularies. In line with current recommendations, costs and health outcomes will be discounted at 3.5%. The primary health economic outcome of the model will be expressed in terms of the incremental cost per quality-adjusted life year (QALY) gained. The cost-effectiveness of all interventions and comparators will be compared incrementally against each other.

Sensitivity analysis will be undertaken to examine the key determinants of cost-effectiveness. Probabilistic sensitivity analysis (PSA) will be undertaken to generate information on the likelihood that each treatment produces the greatest amount of net benefit. The results of this PSA will be presented as cost-effectiveness acceptability curves (CEACs).

7. Handling the company submission(s)

Data submitted by the manufacturers/sponsors will be considered if received by the TAR team no later than 14th March 2014. Data arriving after this date will not be considered. If the data meet the inclusion criteria for the review, they will be extracted and quality assessed in accordance with the procedures outlined in this protocol. Any economic evaluations included in the company submission, provided it complies with NICE's advice on economic model submission, will be assessed for clinical validity, reasonableness of assumptions, and appropriateness of the data used in the economic model. If the TAR team judge that the

existing economic evidence is not robust, then further work will be undertaken, either by adapting what already exists or by developing a *de novo* model.

Any 'commercial in confidence' data taken from a company submission will be underlined and highlighted in turquoise in the assessment report (followed by an indication of the relevant company name, e.g. in brackets). Any academic in confidence data will be underlined and highlighted in yellow.

8. Competing interests of authors

None

9. Appendices

Appendix 1: Search strategy

Database: Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R) <1946 to Present>

1. Colitis, Ulcerative/
2. ulcerative colitis.tw.
3. colitis ulcerosa.tw.
4. uc.tw.
5. colitis ulcerative.tw.
6. Colitis/
7. colitis.tw.
8. colitides.tw.
9. Inflammatory Bowel Diseases/
10. inflammatory bowel disease$.tw.
11. ibd.tw.
12. or/1-11
13. adalimumab.af.
14. humira.af.
15. d 2e7.af.
16. d2e7.af.
17. 331731-18-1.rn.
18. infliximab.af.
19. remicade.af.

20. 170277-31-3.rn.
21. ta650.af.
22. ta 650.af.
23. inx.af.
24. remsima.af.
25. inflectra.af.
26. ct p13.af.
27. ctp13.af.
28. golimumab.af.
29. simponi.af.
30. cnto148.af.
31. cnto 148.af.
32. 476181-74-5.rn.
33. or/13-32
34. 12 and 33

Search strings 1-11 are terms for the condition, ulcerative colitis, with string 12 combining these terms with OR.
Search strings 13-32 are terms for the interventions, adalimumab, infliximab and golimumab, with string 33 combining these terms with OR.
Search string 34 combines the condition and intervention terms together to retrieve studies about the condition and intervention.

The filters provided below will each be combined with the search above to retrieve trials, systematic reviews and economic literature on the condition and intervention.

<u>RCT search filter for Ovid MEDLINE(R)</u>

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. groups.ab.
9. or/1-8

10. exp animals/ not humans.sh.
11. 9 not 10

Systematic Reviews search filter for Ovid MEDLINE(R)

1. Meta-Analysis/
2. meta analy$.tw.
3. metaanaly$.tw.
4. meta analysis.pt.
5. (systematic adj (review$1 or overview$1)).tw.
6. exp Review Literature/
7. or/1-6
8. cochrane.ab.
9. embase.ab.
10. (psychlit or psyclit).ab.
11. (psychinfo or psycinfo).ab.
12. (cinahl or cinhal).ab.
13. science citation index.ab.
14. bids.ab.
15. cancerlit.ab.
16. or/8-15
17. reference list$.ab.
18. bibliograph$.ab.
19. hand-search$.ab.
20. relevant journals.ab.
21. manual search$.ab.
22. or/17-21
23. selection criteria.ab.
24. data extraction.ab.
25. 23 or 24
26. review.pt.
27. 25 and 26
28. comment.pt.
29. letter.pt.
30. editorial.pt.
31. animal/
32. human/
33. 31 not (31 and 32)

34. or/28-30,33
35. 7 or 16 or 22 or 27
36. 35 not 34

Economic search filter for Ovid MEDLINE(R)

1. exp "costs and cost analysis"/
2. economics/
3. exp economics, hospital/
4. exp economics, medical/
5. economics, nursing/
6. exp models, economic/
7. economics, pharmaceutical/
8. exp "fees and charges"/
9. exp budgets/
10. budget$.tw
11. ec.fs
12. cost$.ti
13. (cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab
14. (economic$ or pharmacoeconomic$ or pharmaco-economic$).ti
15. (price$ or pricing$).tw
16. (financial or finance or finances or financed).tw
17. (fee or fees).tw
18. (value adj2 (money or monetary)).tw
19. quality-adjusted life years/
20. (qaly or qalys).af.
21. (quality adjusted life year or quality adjusted life years).af.
22. or/1-21

Appendix 9.2. Draft data extraction form

DRAFT DATA EXTRACTION FORM (VERSION 1.1)	
TRIAL DETAILS	
Author, year	
Objective	
Study design (e.g. RCT)	
Publication type (i.e. full report or abstract)	
Country of corresponding author	
Sources of funding	
INTERVENTIONS	
Focus of interventions (comparisons)	
Description	
Intervention group	
Intervention name	
Intervention dosing regimen and route of administration	
Comparator group	
Comparator name	
Comparator dosing regimen and route of administration	
Geographical Setting (number of study sites, geographical location details)	
Length of study and latest time point available with data	
Duration of treatment	
Length of follow-up (if different)	
STUDY CHARACTERISTICS	
Method of randomisation	
Description	
Generation of allocation sequences	
Allocation concealment	
Blinding level	
Numbers included in the study	
Numbers randomised	
POPULATION CHARACTERISTICS	
Target population (describe)	
Inclusion / exclusion criteria (n)	
Diagnosis method applied	
Recruitment procedures used (participation rates if available)	
Characteristics of participants at baseline	
Age	
Gender	
Ethnicity	
Extent of disease severity at baseline	
Duration of disease	
Comorbidities at baseline	
Details of any previous colorectal surgical intervention for ulcerative colitis	
Any details of previous conventional treatments (including type, dose and duration)	
Proportion receiving steroids at baseline	
Details of any other medication at baseline and whether discontinued	
Concomitant medications during study	
Any other relevant information	
Were intervention and control groups comparable?	
OUTCOMES	
Measures of disease activity	
Mortality	
Rates of and duration of response, relapse and remission	
Rates of hospitalisation	
Rates of surgical intervention	
Time to surgical intervention	
Adverse events of treatment (including leakage and infections following surgery)	
Health-related quality of life	

Any evidence of selective reporting of outcomes?	
ANALYSIS	
Statistical techniques used	
Intention to treat analysis?	
Power calculation?	
Any rescue therapy / early escape option?	
Attrition rates	
Was attrition adequately dealt with?	
Number (%) followed-up	
RESULTS	
Measures of disease activity	
Mortality	
Rates of and duration of response, relapse and remission	
Rates of hospitalisation	
Rates of surgical intervention	
Time to surgical intervention	
Adverse events of treatment (including leakage and infections following surgery)	
Health-related quality of life	
Other information	
SUMMARY	
Authors' overall conclusions	
Reviewers' comments	

Appendix 9.3. Timetable/milestones

Milestone	Date
Draft protocol	1st November 2013
Final protocol	22nd November 2013
Progress report	21st March 2014
Draft assessment report	27th May 2014
Final Assessment report	24th June 2014

10. References

(1) NICE. Ulcerative colitis: Management in adults, children and young people. CG166. 2013.

(2) Hoie O, Wolters F, Riis L, Aamodt G, Solberg C, Bernklev T et al. Ulcerative colitis: patient characteristics may predict 10-yr disease recurrence in a European-wide population-based cohort. *Am J Gastroenterol* 2007; 102(8):1692-1701.

(3) Roth LS, Chande N, Ponich T, Roth ML, Gregor J. Predictors of disease severity in ulcerative colitis patients from Southwestern Ontario. *World J Gastroenterol* 2010; 16(2):232-236.

(4) Lee JH, Cheon JH, Moon CM, Park JJ, Hong SP, Kim TI et al. Do patients with ulcerative colitis diagnosed at a young age have more severe disease activity than patients diagnosed when older? *Digestion* 2010; 81(4):237-243.

(5) Ng SC, Kamm MA. Therapeutic strategies for the management of ulcerative colitis. *Inflamm Bowel Dis* 2009; 15(6):935-950.

(6) EMC. Remicade Summary of Product Characteristics. 2013. 30-10-2013.

(7) EMC. Humira Summary of Product Characteristics. 2013. 30-10-2013.

(8) EMC. Simponi Summary of Product Characteristics. 2013. 30-10-2013.

(9) NICE. Infliximab for subacute manifestations of ulcerative colitis. TA140. 2011.

(10) NICE. Adalimumab for the treatment of moderate to severe ulcerative colitis (terminated appraisal). TA262. 2013.

(11) Centre for Reviews and Dissenination. Systematic reviews: CRD's guidance for undertaking reviews in health care. 2009.

(12) Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *Annals of Internal Medicine* 2009; 151(4):1.

(13) The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions. England: John Wiley & Sons Ltd; 2008.

(14) Eddy DM. Technology assessment: the role of mathematical modeling. Assessing medical technology. 1985. Washington, DC, National Academy Press.

(15) Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ 313[7052], 275-283. 1996.

Appendix 2: MEDLINE search for clinical effectiveness evidence

Database: Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R) <1946 to Present>
Search Strategy:
---

1. Colitis, Ulcerative/
2. ulcerative colitis.tw.
3. colitis ulcerosa.tw.
4. uc.tw.
5. colitis ulcerative.tw.
6. Colitis/
7. colitis.tw.
8. colitides.tw.
9. Inflammatory Bowel Diseases/
10. inflammatory bowel disease$.tw.
11. ibd.tw.
12. (col* and ulcer*).tw.
13. colitis gravis.tw.
14. proctocolitis.tw.
15. or/1-14
16. adalimumab.af.
17. humira.af.
18. d 2e7.af.
19. d2e7.af.
20. 331731-18-1.rn.
21. infliximab.af.
22. remicade.af.
23. 170277-31-3.rn.
24. ta650.af.
25. ta 650.af.
26. inx.af.
27. remsima.af.
28. inflectra.af.
29. ct p13.af.
30. ctp13.af.
31. golimumab.af.
32. simponi.af.
33. cnto148.af.
34. cnto 148.af.
35. 476181-74-5.rn.
36. tnf inhibitor$.tw.
37. anti tnf.tw.
38. antitnf.tw.
39. tnf antagonist$.tw.
40. tnf-alpha blocker$.tw.
41. antitumo?r necrosis factor.tw.

42. Biosimilar Pharmaceuticals/
43. (biosimilar$ or biologic$).tw.
44. or/16-43
45. 15 and 44

Terms 1-14 are terms for the condition (ulcerative colitis) which are then combined using OR in term 15. Terms 16-43 are terms for the interventions (infliximab, adalimumab and golimumab) which are then combined using OR in term 44. Terms 15 and 44 are then combined using AND to find studies on the condition and interventions in term 45.
To retrieve RCTs and systematic reviews specially designed highly sensitive search filter were combined with term 45. RCT filter and systematic review filter below.
RCT filter

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. 49 placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. groups.ab.
9. or/1-8
10. exp animals/ not humans.sh.
11. 9 not 10

Systematic Review Filter

1. Meta-Analysis/
2. meta analy$.tw.
3. metaanaly$.tw.
4. meta analysis.pt.
5. (systematic adj (review$1 or overview$1)).tw.
6. exp Review Literature/
7. or/1-6
8. cochrane.ab.
9. embase.ab.
10. (psychlit or psyclit).ab.
11. (psychinfo or psycinfo).ab.
12. (cinahl or cinhal).ab.
13. science citation index.ab.
14. bids.ab.
15. cancerlit.ab.
16. or/8-15
17. reference list$.ab.
18. bibliograph$.ab.
19. hand-search$.ab.
20. relevant journals.ab.
21. manual search$.ab.
22. or/17-78
23. selection criteria.ab.

24. data extraction.ab.
25. 23 or 24
26. review.pt.
27. 25 and 26
28. comment.pt.
29. letter.pt.
30. editorial.pt.
31. animal/
32. human/
33. 31 not (31 and 32)
34. or/28-30,33
35. 7 or 16 or 22 or 27
36. 35 not 34

Appendix 3: Table of excluded studies

Author and Year (NCT number)	Reason for exclusion
Actis *et al.*, 2002 [163]	Not randomised controlled trial
Actis, 2003 [164]	Not randomised controlled trial
Afif *et al.*, 2009 [165]	Not randomised controlled trial
Allez *et al.*, 2010 [166]	Not randomised controlled trial
Anon, 2007 [167]	Not randomised controlled trial
Armuzzi *et al.*, 2004 [168]	Not protocol-eligible population. No prior immunosuppressant use reported.
Baert *et al.*, 2007 [169]	Not randomised controlled trial
Barbato *et al.*, 2006 [170]	Not randomised controlled trial
Barreiro-de *et al.*, 2009 [171]	Not randomised controlled trial
Baumgart, 2010 [172]	Not randomised controlled trial
Ben-Horin,2012 [173]	Not treatment of interest (rituximab)
Bengi & Akpinar, 2012 [174]	Not randomised controlled trial
Biancone *et al.*, 2009 [175]	Not randomised controlled trial
BMJ, 2012 [176]	Not randomised controlled trial
Bordeianou, 2009 [177]	Not randomised controlled trial
Borruel *et al.*, 2013 [178]	Not randomised controlled trial
Brooklyn *et al.*, 2006 [179]	Not ulcerative colitis trial population
Bujanover & Weiss, 2008 [180]	Not randomised controlled trial
Busquets & Aldeguer, 2012 [181]	Not randomised controlled trial
Carbone *et al.*, 2009 [182]	Not randomised controlled trial
Cariñanos *et al.*, 2011 [183]	Not randomised controlled trial
Casteele *et al.*, 2012 [184], [185]	Evaluation of two IFX dosing strategies
Charles *et al.*, 2010 [186]	Not randomised controlled trial
Chen *et al.*, 2013 [187]	Not randomised controlled trial
Chey & Shah, 2005 [188]	Not randomised controlled trial
Chowers *et al.*, 2010 [189]	Not randomised controlled trial
Chuang *et al.*, 2010 [190]	Not randomised controlled trial
Cohen, 2003 [191]	Not randomised controlled trial
Colombel *et al.*, 2011 [192]	No protocol-eligible outcome data
Colombel *et al.*, 2011 [192]	Not randomised controlled trial
Colombel *et al.*, 2012[193]	Parallel publication, duplicate outcome data
Cottone, Orlando & Mocciaro, 2009	Not randomised controlled trial

Author and Year (NCT number)	Reason for exclusion
[194]	
Croft *et al.*, 2013 [195]	Population outside scope of appraisal (use of biologic in acute severe ulcerative colitis following failure of i.v. steroids)
Cross, Lapshin & Finkelstein, 2008 [196]	Not randomised controlled trial
Danese, 2013[197]	Unable to obtain
De Vos *et al.*, 2012 [198]	Not randomised controlled trial
de Vries, 2012 [199]	Not randomised controlled trial
D'Haens, 2005 [200]	Not randomised controlled trial
Dean *et al.*, 2012 [201]	Not randomised controlled trial
Dignass *et al.*, 2012 [202]	Not randomised controlled trial
Domènech *et al.*, 2010 [203]	Not randomised controlled trial
Dranitsaris *et al.*, 2012 [204]	Not randomised controlled trial
Eidelwein *et al.*, 2005 [205]	Not randomised controlled trial
Erikkson *et al.*, 2012 [206]	Not randomised controlled trial
Esteve *et al.*, 2011 [207]	Not randomised controlled trial
EUCTR2007-006692-37-GB [208]	Not ulcerative colitis trial population
EUCTR2007-007702-30-IT [208]	Not randomised controlled trial
EUCTR2007-000842-11-AT [208]	Not randomised controlled trial
EUCTR2008-007519-34-SE [208]	Not randomised controlled trial
EUCTR2011-002411-29-SE [208]	Not randomised controlled trial
EUCTR2011-006084-22-GB [208]	Not randomised controlled trial
Fanjiang *et al.*, 2007 [209]	Not randomised controlled trial
Fasanmade *et al.*, 2009 [210]	No protocol-eligible outcome data
Fasanmade *et al.*, 2010 [211]	No protocol-eligible outcome data
Feagan *et al.*,2005 [212]	Not randomised controlled trial
Feagan, 2006 [213]	Not randomised controlled trial
Florholmen *et al.*, 2011 [214]	Population outside scope of appraisal (use of biologic in acute severe ulcerative colitis following failure of i.v. steroids)
Ford *et al.*, 2013 [4]	Not randomised controlled trial
Gao & Jiang, 2013 [215]	Not available in English language
Gavalas *et al.*, 2007 [216]	Population outside scope of appraisal (use of biologic in

Author and Year (NCT number)	Reason for exclusion
	acute severe ulcerative colitis)
Gearry & Falvey, 2012 [217]	Not randomised controlled trial
Gies *et al.*, 2010 [218]	Not randomised controlled trial
Ginard *et al.*, 2008 [219]	Not randomised controlled trial
Grosen & Julsgaard, 2013 [220]	Not randomised controlled trial
Gustavsson *et al.*, 2010 [221]	Follow-on study to excluded Järnerot *et al.*, 2005 [111]
Ha *et al.*, 2009 [222]	Not randomised controlled trial
Halpin *et al.*, 2010 [223]	Not randomised controlled trial
Halpin & Hamlin, 2012 [224]	Not randomised controlled trial
Hämäläinen *et al.*, 2011 [225]	Not randomised controlled trial
Hanauer, 2005 [226]	Not randomised controlled trial
Hanauer *et al.*, 2008 [227]	Not randomised controlled trial
Hanauer, Rubin & Sandborn, 2008 [228]	Not randomised controlled trial
Heraganahally *et al.*, 2009 [229]	Not randomised controlled trial
Herrlinger *et al.*, 2010 [230]	Not randomised controlled trial
Honeywell, Touchstone & Caspi, 2007 [231]	Unable to obtain
Hyams *et al.*, 2010 [232]	Not randomised controlled trial
Hyams *et al.*, 2011 [233]	Unable to obtain
Assasi, N (INAHTA), 2011 [234]	Not randomised controlled trial
Jackson, 2007 [235]	Not randomised controlled trial
Järnerot *et al.*, 2005 [111]	Population outside scope of appraisal (use of biologic in acute ulcerative colitis following failure of i.v. steroids)
Järnerot, 2006 [236]	Not randomised controlled trial
Jiménez 2004 [237]	Not randomised controlled trial
Joob and Wiwanikit, 2013 [238]	Not randomised controlled trial
JPRN-UMIN000006169 [208]	Not randomised controlled trial
JPRN-UMIN000007256 [208]	Not randomised controlled trial
JPRN-UMIN000007806 [208]	Not randomised controlled trial
JPRN-UMIN000010205 [208]	Not randomised controlled trial
JPRN-UMIN000013033 [208]	Not randomised controlled trial
Kaser & Tilg, 2008 [239]	Not randomised controlled trial
Kaur & Targan, 2013 [240]	Not randomised controlled trial

Author and Year (NCT number)	Reason for exclusion
Kerbleski & Gottlieb, 2009 [241]	Not randomised controlled trial
Klotz, Teml & Schwab M, 2007 [242]	Not randomised controlled trial
Kohn *et al.*, 2004 [243]	Not randomised controlled trial
Kohn *et al.*, 2007 [244]	Not randomised controlled trial
Kohn, 2008 [245]	Not randomised controlled trial
Laharie *et al.*, 2012 [246,247]	Population outside scope of appraisal (use of biologic in acute severe ulcerative colitis following failure of i.v. steroids)
Leal *et al.*, 2012 [248]	Not randomised controlled trial
LeBlanc *et al.*, 2013 [249]	Not randomised controlled trial
Leblanc *et al.*, 2011 [250]	Not randomised controlled trial
Levesque & Sandborn, 2012 [251]	Not randomised controlled trial
Levy, 2009 [252]	Not randomised controlled trial
Li *et al.*, 2013 [253]	No protocol-eligible outcome data
Lichtenstein, 2001 [254]	Not randomised controlled trial
Lichtenstein, 2009 [255]	Not randomised controlled trial
Liu *et al.*, 2013 [256]	Not randomised controlled trial
Löfberg *et al.*, 2012 [257]	Not ulcerative colitis trial population
Lorenzo-Zúñiga *et al.*, 2013 [258]	Not randomised controlled trial
Mallow *et al.*, 2013 [259]	Not randomised controlled trial
Mallow *et al.*, 2013 [260]	Not randomised controlled trial
Maser *et al.*, 2008 [261]	Not randomised controlled trial
Matsumoto, 2007 [262]	Not randomised controlled trial
Mazumdar & Greenwald, 2009 [263]	Not randomised controlled trial
McCann & Smith, 2012 [264]	Not randomised controlled trial
Molnár *et al.*, 2010 [265]	Not randomised controlled trial
Molnár *et al.*, 2011 [266]	Not randomised controlled trial
Molnár *et al.*, 2011 [267]	Not randomised controlled trial
Moss & Farrell, 2006 [268]	Not randomised controlled trial
Nakase *et al.*, 2010 [269]	Not randomised controlled trial
National Institute for Health Research, 2011 [270]	Not randomised controlled trial
National Institute for Health Research, 2011 [271]	Not randomised controlled trial
National Institute for Health	Not randomised controlled trial

Author and Year (NCT number)	Reason for exclusion
Research, 2013 [272]	
NCT00207688 [208]	Not randomised controlled trial
NCT00421642 [208]	Not randomised controlled trial
NCT00488774 [273]	Unlicensed route of administration for intervention
NCT00573794 [273]	Not randomised controlled trial
NCT00586599 [273]	Not randomised controlled trial
NCT00586807 [273]	Not randomised controlled trial
NCT00606346 [273]	Not randomised controlled trial
NCT00705484 [273]	Not randomised controlled trial
NCT00745329 [273]	Not randomised controlled trial
NCT00791557 [273]	Not randomised controlled trial
NCT00955123 [273]	Not randomised controlled trial
NCT00984568 [273]	Evaluation of different IFX treatment strategies
NCT01346826 [273]	Evaluation of accelerated IFX infusions
NCT01408810 [273]	Not randomised controlled trial
NCT01417728 [273]	Not randomised controlled trial
NCT01494857 [273]	Not randomised controlled trial
NCT01550965 [273]	Not randomised controlled trial
NCT01585155 [273]	Not randomised controlled trial
NCT01670240 [273]	Evaluation of biologic in treatment of chronic pouchitis (trial currently recruiting)
NCT01716039 [273]	Study evaluating ADA-MTX interaction
NCT01787786 [273]	Not randomised controlled trial
NCT01846026 [273]	Not protocol-eligible intervention
NCT01848561 [273]	Not randomised controlled trial
NCT01851343 [208]	Not randomised controlled trial
NCT01900574 [273]	Not randomised controlled trial
NCT01947816 [273]	Not randomised controlled trial
NCT01960426 [273]	Evaluation of two dosing methods
NCT01971814 [273]	Not randomised controlled trial
NCT01988961 [273]	Not randomised controlled trial
NCT02057016 [208]	Not randomised controlled trial
NCT02073526 [208]	Not randomised controlled trial
Nguyen & Prather, 2009 [274]	Not randomised controlled trial
Nielsen & Jess, 2013 [275]	Unable to obtain
Ochsenkühn *et al.*, 2004 [276]	Population outside scope of appraisal (i) use of biologic in acute severe ulcerative colitis, ii) no patients were receiving immunosuppressants/immunomodulators/more than 10mg/day prednisolone at baseline and therefore not inadequate responders/stated intolerant to conventional therapy options)
Orlando *et al.*, 2012 [277]	Not ulcerative colitis trial population
Oussalah *et al.*, 2008 [278]	Not randomised controlled trial
Oussalah *et al.*, 2010 [279]	Not randomised controlled trial
Panncione *et al.*, 2011 [280]	Parallel publication, duplicate outcome data

Author and Year (NCT number)	Reason for exclusion
Panncione *et al.*, 2013 [281]	Parallel publication, duplicate outcome data
Pardi & Sandborn, 2008 [282]	Not randomised controlled trial
Pastore *et al.*, 2010 [283]	Not randomised controlled trial
Pastorelli *et al.*, 2009 [284]	Not randomised controlled trial
Pearce & Lawrance, 2007 [285]	Not randomised controlled trial
Peyrin-Biroulet *et al.*, 2007 [286]	Not randomised controlled trial
Pola *et al.*, 2013 [287]	Not randomised controlled trial
Reinisch *et al.*, 2012 [288]	No protocol-eligible outcome data
Rizzello *et al.*, 2013 [289]	Not randomised controlled trial
Rostholder *et al.*, 2012 [290]	Not randomised controlled trial
Rubin *et al.*, 2012 [291]	Not ulcerative colitis trial population
Russell & Katz, 2004 [292]	Not randomised controlled trial
Rutgeerts, 2002 [293]	Not randomised controlled trial
Rutgeerts *et al.*, 2010 [294]	Not randomised controlled trial
Rutgeerts *et al.*, 2013a [295]	Parallel publication, duplicate outcome data
Rutgeerts *et al.*, 2013b [296]	Parallel publication, duplicate outcome data
Salvana & Salata, 2009 [297]	Not randomised controlled trial
Sandborn *et al.*, 2007 [298]	Not randomised controlled trial
Sandborn *et al.*, 2009 [68]	Not randomised controlled trial
Sandborn, 2012 [299]	Not randomised controlled trial
Sandborn *et al.*, 2011[300]	Parallel publication, duplicate outcome data
Sandborn *et al.*, 2011[301]	Parallel publication, duplicate outcome data
Sandborn *et al.*, 2012 [302]	Parallel publication, duplicate outcome data
Sandborn *et al.*, 2012[303]	Parallel publication, duplicate outcome data
Sandborn *et al.*, 2012 [304]	Parallel publication, duplicate outcome data
Sandborn *et al.*, 2012[60]	Parallel publication, duplicate outcome data
Sandborn *et al.*, 2012 [305]	Unable to obtain
Sandborn *et al.*, 2012 [306]	Unable to obtain
Sandborn & Loftus, 2004 [307]	Not randomised controlled trial
Sands *et al.*, 2001 [308]	Population outside scope of appraisal (use of biologic in acute severe ulcerative colitis following failure of i.v. steroids)
Scholmerich, 2009 [309]	Not randomised controlled trial
Sciaudone *et al.*, 2010 [310]	Not randomised controlled trial

Author and Year (NCT number)	Reason for exclusion
Sciaudone *et al.*, 2011 [311]	Not randomised controlled trial
Seirafi *et al.*, 2011 [312]	Not randomised controlled trial
Siemanowski & Regueiro, 2007 [313]	Not randomised controlled trial
Simmons & Jewell, 2002 [314]	Not randomised controlled trial
Singh & Loftus, 2013 [315]	Not randomised controlled trial
Sjöberg *et al.*, 2012 [316]	Not randomised controlled trial
Smith, 2013 [317]	Unable to obtain
Sokol *et al.*, 2010 [318]	Not randomised controlled trial
Stein *et al.*, 2013 [319]	Unable to obtain
Su *et al.*, 2002 [320]	Not randomised controlled trial
Taxonera *et al.*, 2011 [321]	Not randomised controlled trial
Thorlund *et al.*, 2013 [322]	Not randomised controlled trial
Toedter *et al.*, 2010 [323]	No protocol-eligible outcome data
Toedter *et al.*, 2011 [324]	No protocol-eligible outcome data
Travis, 2011 [325]	Not randomised controlled trial
Tursi *et al.*, 2010 [326]	Not randomised controlled trial
Van Assche, 2008 [327]	Not randomised controlled trial
Van Assche, Vermeire & Rutgeerts, 2008 [328]	Not randomised controlled trial
van Casteren-Messidoro & Zelinkova 2012 [329]	Not randomised controlled trial
Velayos & Mahadevan, 2007 [330]	Not randomised controlled trial
Vermeire *et al.*, 2011 [331]	Not treatment of interest (PF-00547,659)
Warner & Harris, 2012 [332]	Not randomised controlled trial
Waters *et al.*, 2008 [333]	Unable to obtain
Waters *et al.*, 2009. [334]	Not randomised controlled trial
Willert & Lawrance, 2008 [335]	Not randomised controlled trial
Wolf, 2007 [336]	Not randomised controlled trial
Wolf *et al.*, 2012[337]	Parallel publication, duplicate outcome data
Yamamoto & Shiraki, 2013 [338]	Not randomised controlled trial
Yamamoto-Furusho & Uzcanga, 2008 [339]	Not randomised controlled trial
Yapali & Hamzaoglu, 2007 [340]	Not randomised controlled trial

Appendix 4: Table of numbers withdrawing and reasons for withdrawal

Table 84. Participants withdrawing from treatment arms, reasons for withdrawal, and risk of attrition bias assessment judgement

Study and RM No.	Treatment arm	No. completing - n/N (%)	Reasons for withdrawal	Attrition bias judgement
ULTRA1[45]	PBO	ITT-A3 (amendment): 121/130 (93%)	ITT-A3 (amendment): adverse event, 5/130 (4%); lack of efficacy, 4/130 (3%) - overall, 9/130 (7%)	Low risk - <10% attrition in each group, numbers balanced across groups, ITT analysis presented
ULTRA1 [45]	ADA 160/80mg	ITT-A3 (amendment): 118/130 (91%)	ITT-A3 (amendment): adverse event, 6/130 (5%); withdrew consent, 2/130 (1.5%); lost to follow-up, 2/130 (1.5%); lack of efficacy, 2/13 (1.5%) - overall, 12/130 (9%)	
ULTRA2 [46]	PBO	135/260 (52%) switched to OL of which n=84 dose escalated to 40g wk Overall, 131/260 (50.4%); completed on 16/80/40 dosing, 56/260 (21.5%); switched to open label 40mg every other week at week 12 (n=135), 30/260 (11.5%); dose escalated to 40mg weekly (n=68), 45/260 (17.3%)	Site non-compliance, 10/258 (4%); lack of efficacy, 63/258 (24%); adverse event, 12/258 (5%); withdrew consent, 8/258 (3%); lost to follow-up, 1/258 (<1%); protocol violation, 1/258 (<1%); other, 9/258 (3%)	High risk - although ITT analysis was undertaken, there was a high level of attrition and an imbalance between treatment groups (PBO, 50%; ADA, 59%)
ULTRA2 [46]	ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW beginning at week 4	116/258 (45%) switched to OL of which n=68 dose escalated to 40g wk Overall, 154/258 (59.7%); completed on 16/80/40 dosing, 82/258 (31.7%); switched to open label 40mg every other week at week 12 (n=116), 32/258 (12.4%); dose escalated to 40mg weekly (n=68), 40/258 (15.5%)	Site non-compliance, 14/260 (5%); lack of efficacy, 70/260 (27%); adverse event, 25/260 (10%); withdrew consent, 4/260 (1.5%); protocol violation, 5/260 (2%); other, 11/260 (4%)	

Study and RM No.	Treatment arm	No. completing - n/N (%)	Reasons for withdrawal	Attrition bias judgement
Suzuki [47]	PBO	Week 8: 92/96 (96%) Week 52: 73/96 (77%)	Week 8: total discontinued, 4/96 (4%) - lack of efficacy n=2, adverse event n=2 Week 52: total discontinued, 23/96 (23%) - withdrew consent n=2, lack of efficacy n=14, adverse event n=7, moved to rescue therapy n=63	Induction: Low risk - <10% attrition in each group and numbers reasonably balanced across groups. All patients accounted for in the primary outcome analysis Maintenance: High risk - PBO, 23%; ADA, 33%
Suzuki [47]	ADA80/ 40mg	Week 8: 85/87 (98%) (UNLICENCED) Week 52: 58/87 (67%)	Week 8: total discontinued, 2/87 (2%) - withdrew consent n=1, lack of efficacy n=1 Week 52: total discontinued, 29/87 (33%) - withdrew consent n=3, lack of efficacy n=17, adverse event n=9, moved to rescue therapy n=50	
Suzuki [47]	ADA160/ 80mg	Week 8: 86/90 (96%) Week 52: 60/90 (67%)	Week 8: total discontinued, 4/90 (4%) - lack of efficacy n=1, adverse event n=3 Week 52: total discontinued, 30/90 (33%) - lack of efficacy n=16, adverse event n=13, other n=1, moved to rescue therapy n=46	
ULTRA3 [61]	PBO	91/121 (75%)	Lack of efficacy, 21/121 (17%); adverse event, 16/121 (13%); withdrew consent, 3/121 (2%); protocol violation, 1/121 (1%)	Extension study not included in RoB assessment
ULTRA3 [61]	ADA 80/40mg	86/118 (73%)	Lack of efficacy, 17/118 (14%); adverse event, 12/118 (10%); withdrew consent, 5/118 (4%); lost to follow-up, 1/118 (1%); protocol violation, 1/118 (1%); other, 4/118 (3%)	
ULTRA3 [61]	ADA 160/80mg	95/121 (79%)	Lack of efficacy, 15/121 (%); adverse event, 10/121 (%); withdrew consent, 4/121 (%); lost to follow-up, 1/121 (%); protocol violation, 1/121 (%)	
PURSUIT-SC [48]	Phase II PBO	41/42 plus 26/31 enrolled whilst Phase II data being analysed	2/42 'other' reasons	Low risk - ITT reported and withdrawal <10% across all groups and n balanced

Study and RM No.	Treatment arm	No. completing - n/N (%)	Reasons for withdrawal	Attrition bias judgement
PURSUIT-SC [48]	Phase II GOL 200/100 mg all randomised	41/42 plus 31/31 enrolled whilst Phase II data being analysed	1/42 withdrew consent	
PURSUIT-M [49]	PBO	PBO 115/156 (73%) randomised completed through week 54	Discontinued treatment prior to week 52 (n=43): 17 adverse event, 19 unsatisfactory therapeutic effect, 1 lost to follow-up, 6 other Terminated study before week 54 (n=18): 5 withdrew consent, 3 lost to follow-up, 10 other	High risk – although ITT reported, withdrawal >10% across all groups
PURSUIT-M [49]	GOL 50 mg	GOL 50 mg. 120/154 (78%) randomised completed through week 54.	Discontinued treatment prior to week 52 (n=43): 12 adverse event, 17 unsatisfactory therapeutic effect, 2 lost to follow-up, 12 other Terminated study before week 54 (n=18): 10 withdrew consent, 2 lost to follow-up, 6 other	
PURSUIT-M [49]	GOL 100 mg	GOL 100 mg. 116/154 (75%) randomised completed through week 54.	Discontinued treatment prior to week 52 (n=45): 12 adverse event, 22 unsatisfactory therapeutic effect, 1 lost to follow-up, 10 other Terminated study before week 54 (n=21): 11 withdrew consent, 2 lost to follow-up, 8 other	
UC-SUCCESS [52]	AZA	53/79 (66%)	Adverse event, 11/80 (14%); withdrew consent, 8/80 (10%); non-compliance with protocol, 5/80 (6%); protocol ineligible, 3/80 (4%);	High risk - although ITT analysis was undertaken, there was a high level of attrition and an imbalance between treatment groups (AZA, 34%; IFX, 18%; IFX/AZA, 21%)

Study and RM No.	Treatment arm	No. completing - n/N (%)	Reasons for withdrawal	Attrition bias judgement
UC-SUCCESS [52]	IFX	65/78 (82%)	Adverse event, 7/79 (9%); clinical event, 1/79 (1%); lost to follow-up, 1/79 (1%); withdrew consent, 3/79 (4%); non-compliance with protocol, 1/79 (1%); protocol ineligible, 1/79 (1%);	
UC-SUCCESS [52]	IFX/AZA	63/80 (79%)	Adverse event, 8/80 (10%); withdrew consent, 4/80 (5%); non-compliance with protocol, 1/80 (1%); protocol ineligible, 2/80 (3%); administrative reasons, 2/80 (3%)	
Probert [51]	PBO	20/20 (100%)	No withdrawals reported	Low risk - all patients accounted for in the primary outcome analysis
Probert [51]	IFX	23/23 (100%)		
Hyams [53]	IFX/5mg/ q8w	18/22 (82%) completed infusions and follow up	Adverse event, 3/22 (14%); lack of efficacy, 1/22 (5%)	High risk - numbers withdrawing >10% and unbalanced across groups (8qw, 21%; q12w, 51%)
Hyams [53]	IFX/5mg/ q12w	12/23 (52%) complete infusions; 11/23 (49%) completed follow-up	Adverse event, 6/23 (26%); lack of efficacy, 4/23 (17%); other, 1/23 (4%)	
ACT1 [50]	PBO	47 completed study infusions, of whom 46/121 completed follow-up. 74 discontinued study infusions, of whom 18 completed follow-up.	In ACT 1, similar numbers of patients in each group discontinued treatment because of an adverse event	High risk - although ITT reported, >50% in PBO and >30% in IFX 5 and 10mg did not complete
ACT1 [50]	IFX 5mg/kg	76/121 completed study infusions, of whom 76 completed follow-up. 45 discontinued study infusions, of whom 6 completed follow-up		
ACT2 [50]	PBO	67 completed study infusions, of whom 64/123 completed follow-up. 56 discontinued study infusion, 9 completed follow-up.	In ACT 2, more patients in the placebo group than in the two infliximab groups discontinued treatment because of an adverse event	High risk - although ITT reported, >50% in PBO and >30% in IFX 5 and 10mg did not complete
ACT2 [50]	IFX 5mg/kg	97 completed study infusions, of whom 94/121 completed follow-up. 24 discontinued study infusions, of whom 3 completed follow-up.		

Appendix 5: Additional efficacy outcomes tables

Additional efficacy outcomes (adult population trials)

Study name	Treatment arm	Time point	Outcome measure
ULTRA1	PBO	Week 8	[45] Subgroup analysis results: Remission n/N (%): Mayo <10: 10/83 (12.0%) Mayo ≥10: 2/47 (4.3%) Extensive colitis : 11/73 (15.1%) No extensive colitis: 1/57 (1.8%) Corticosteroid (without IMM%): 6/55 (10.9%) IMM (Azathioprine and 6-mercaptopurine%) (without corticosteroid%): 0/18 (0%) 2/25 (8.0%) 6/28 (21.4%) IMM + corticosteroid: 2/34 (5.9%) No corticosteroid + no IMM: 4/23 Aminosalicylates: 11/98 (11.2%) No aminosalicylates: 1/32 (3.1%) CRP <10 mg/l: 7/95 (7.4%) CRP ≥10 mg/l: 4/32 (12.5%) Weight < 70.0 kg: 5/35 (14.3%) Weight ≥70.0 kg, <82.0 kg: 3/43 (7.0%) Weight ≥82.0 kg: 4/52 (7.7%) Change from baseline in CRP mg/L: median -0.09 (range -274.79 to 88.71) Rectal bleeding subscore ≤1, 86/130 (66.2%) PGA subscore ≤1, 61/130 (46.9%) Stool frequeny subscore, 49/130 (37.7%) [62] change from baseline: Haemoglobin g/L, 4.4; p-value vs. PBO, <0.001 Haematocrit fraction, 0.014; p-value vs. PBO, <0.001 Red blood cells x1012/L, 0.16; p-value vs. PBO, <0.01 Total protein g/L, 1.5; p-value vs. PBO, <0.05 Albumin g/L, 1.3

Study name	Treatment arm	Time point	Outcome measure
			CRP mg/L, -0.47
ULTRA1	ADA 160/80mg	Week 8	[45] Subgroup analysis results: Remission n/N (%): Mayo <10: 17/85 (20.0%), difference from placebo (95% CI) 1.5 (e8.7 to 11.8) 8.0 (e3.1 to 19.0) Mayo ≥10: 7/45 (15.6%), difference from placebo (95% CI) 11.3 (-0.8 to 23.4) Extensive colitis: 12/60 (20.0%), difference from placebo (95% CI) 4.9 (-8.1 to 18.0) No extensive colitis: 12/70 (17.1%), difference from placebo (95% CI) 15.4 (5.9 to 24.9) Corticosteroid (without IMM%): 10/48 (20.8%), difference from placebo (95% CI) 11.3 (-3.0 to 25.7) IMM (Azathioprine and 6-mercaptopurine%) (without corticosteroid%): 6/28 (21.4%), difference from placebo (95% CI) 20.7 (5.9 to 35.4) IMM + corticosteroid: 2/23 (8.7%), difference from placebo (95% CI) 4.1 (-11.2 to 19.5) No corticosteroid + no IMM: 6/31 (19.4%), difference from placebo (95% CI) -1.0 (-20.5 to 18.5) Aminosalicylates: 18/105 (17.1%), difference from placebo (95% CI) 5.9 (-3.6 to 15.5) No aminosalicylates: 6/25 (24.0%), difference from placebo (95% CI) 20.9 (3.1 to 38.7) CRP <10 mg/l: 21/101 (20.8%), difference from placebo (95% CI) 13.4 (3.9 to 22.9) CRP ≥10 mg/l: 2/25 (8.0%), difference from placebo (95% CI) -4.5 (-20.1 to 11.1) Weight < 70.0 kg: 5/35 11/45 (24.4%), difference from placebo (95% CI) 10.2 (-6.9 to 27.2) Weight ≥70.0 kg, <82.0 kg: 8/33 (24.2%), difference from placebo (95% CI) 17.3 (0.8 to 33.8) Weight ≥82.0 kg: 5/52 (9.6%), difference from placebo (95% CI) 1.9 (-8.9 to 12.7) Change from baseline in CRP mg/L: median -0.77 (range -95.09 to 130.41) Rectal bleeding subscore ≤1, 101/130 (77.7%) PGA subscore ≤1, 78/130 (60.0%) Stool frequeny subscore, 63/130 (48.5%) [62] change from baseline: Haemoglobin g/L, 4.9; p-value vs. PBO, <0.001 Haematocrit fraction, 0.014; p-value vs. PBO, <0.001 Red blood cells x1012/L, 0.19; p-value vs. PBO, <0.001 Total protein g/L, 1.7; p-value vs. PBO, <0.01 Albumin g/L, 1.7; p-value vs. PBO, <0.01 CRP mg/L, -0.87; p-value vs. PBO, <0.001

Study name	Treatment arm	Time point	Outcome measure
ULTRA2	PBO	Week 8	[46] No prior anti-TNF-α treatment: PGA ≤1, 63/145 (43.4%) SFS ≤1, 43/145 (29.7%) RBS ≤1, 86/145 (59.3%) Prior anti-TNF-α treatment: PGA ≤1, 29/101 (28.7%) SFS ≤1, 27/101 (26.7%) RBS ≤1, 57/101 (56.4%)
ULTRA2	PBO	Week 32	EPAR (Humira)[35] Number and percentage of subjects taking corticosteroids at baseline who discontinued corticosteroid use and achieved clinical remission per Mayo score at week 32 (ITT Analysis): Clinical remission at week 32 - discontinued CS at any time prior to week 32, 10/140 (7.1%) Clinical remission at week 32 - discontinued CS for ≥90 days prior to week 32, 9/140 (6.4%)
ULTRA2	PBO	Week 52	[46] Discontinued corticosteroid use before week 52 and achieved clinical remission at week 52 among patients with baseline corticosteroid use, 5/81 (6.2%) Discontinued corticosteroid use for ≥90 days before week 52 and achieved remission at week 52 among patients with baseline corticosteroid use, 5/81 (6.2%) Discontinued corticosteroid use and achieved sustained clinical remission at both weeks 32 and 52 among patients with baseline corticosteroid use, 1/81 (1.2%) IBDQ responders at week 52, 31/145 (21.4%) Prior anti-TNF-α treatment: Discontinued corticosteroid use before week 52 and achieved clinical remission at week 52 Among patients with baseline corticosteroid use, 3/59 (5.1%) Discontinued corticosteroid use for ≥90 days before week 52 and achieved remission at week 52 among patients with baseline corticosteroid use, 3/59 (5.1%) Discontinued corticosteroid use and achieved sustained clinical remission at both weeks 32 and 52 among patients with baseline corticosteroid use, 1/59 (1.7%) IBDQ responders at week 52, 9/101 (8.9%)

Study name	Treatment arm	Time point	Outcome measure
			EPAR (Humira)[35] Discontinued cosrticosteroid use for ≥ 90 days before week 52 and achieved remission at week 52, 8/246 (5.7%) Discontinued cosrticosteroid use and sustained remission at both week 32 and week 52, 2/246 (1.4%) IBDQ responders at week 52, 40/246 (16.3%) Number and percentage of subjects taking corticosteroids at baseline who discontinued corticosteroid use and achieved clinical remission per Mayo score at week 52 (ITT Analysis): Clinical remission at week 52 - discontinued CS at any time prior to week 52, 8/140 (5.7%) Clinical remission at week 52 - discontinued CS for ≥90 days prior to week 52, 8/140 (5.7%) 306 week 52 corticosteroid-free remission: All PBO, 8/140 (5.7%) Anti-TNF naïve PBO, 5/81 (6.2%) Anti-TNF exposed PBO, 3/59 (5.1%) week 52 corticosteroid-free: All PBO, 32/140 (22.9%) Anti-TNF naïve PBO, 20/81 (24.7%) Anti-TNF exposed PBO, 12/59 (20.3%) 60 Post hoc analysis, week 52 n=246: Mean days in IBDQ remission (IBDQ score ≥170), 79.00 Mean serious-adverse-event adjusted days in clinical remission, 48.23
ULTRA2	ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW	Week 8	46 Serum Trough Concentrations Over Time by Remission Status, mean (SD) [min max], N_{nmiss}: 40 mg EOW patients who were remitters (n=43): 11.4 (5.15) [range 0.000 to 22.8], 41 40 mg EOW patients who were non-remitters (n=153): 8.49 (4.35) [range 0.000 to 21.8], 110 No prior anti-TNF-α treatment: PGA ≤1, 88/150 (58.7%); p-value vs. PBO, 0.009 SFS ≤1, 69/150 (46.0%); p-value vs. PBO, 0.004

Study name	Treatment arm	Time point	Outcome measure
			RBS ≤1, 116/150 (77.3%); p-value vs. PBO, 0.001 IBDQ responders, 102/150 (68.0%); p-value vs. PBO, 0.004 Prior anti-TNF-α treatment: PGA ≤1, 26/98 (26.5%); p-value vs. PBO, 0.731 SFS ≤1, 25/98 (25.5%); p-value vs. PBO, 0.844 RBS ≤1, 58/98 (59.2%); p-value vs. PBO, 0.695 IBDQ responders, 42/98 (42.9%); p-value vs. PBO, 0.370
ULTRA2	ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW	Week 32	[46] Serum Trough Concentrations Over Time by Remission Status, mean (SD) [min max], N_{nmiss}: 40 mg EOW patients who were remitters (n=43), 10.6 (5.64) [range 0.000 26.9], 39 40 mg EOW patients who were non-remitters (n=153), 6.95 (3.98) [0.000 to 18.1], 70
ULTRA2	ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW	Week 52	[46] Serum Trough Concentrations Over Time by Remission Status, mean (SD) [min max], N_{nmiss}: 40 mg EOW patients who were remitters (n=43), 10.8 (7.45) [0.000 to 39.3], 39 40 mg EOW patients who were non-remitters (n=153), 6.18 (4.22) [0.000 16.1], 62 No prior anti-TNF-α treatment: Discontinued corticosteroid use before week 52 and achieved clinical remission at week 52 Among patients with baseline corticosteroid use, 15/110 (13.6%); p-value vs. PBO, 0.096 Discontinued corticosteroid use for ≥90 days before week 52 and achieved remission at week 52 among patients with baseline corticosteroid use, 15/110 (13.6%); p-value vs. PBO, 0.096 Discontinued corticosteroid use and achieved sustained clinical remission at both weeks 32 and 52 among patients with baseline corticosteroid use, 11/110 (10.0%); p-value vs. PBO, 0.014 IBDQ responders at week 52, 48/150 (32.0%); p-value vs. PBO, 0.039 Prior anti-TNF-α treatment: Discontinued corticosteroid use before week 52 and achieved clinical remission at week 52 among patients with baseline corticosteroid use, 5/40 (12.5%); p-value vs. PBO, 0.263 Discontinued corticosteroid use for ≥90 days before week 52 and achieved remission at week 52 among patients with baseline corticosteroid use, 5/40 (12.5%); p-value vs. PBO, 0.263 Discontinued corticosteroid use and achieved sustained clinical remission at both weeks 32 and 52 Among patients with baseline corticosteroid use, 4/40 (10.0%); p-value vs. PBO, 0.155

Study name	Treatment arm	Time point	Outcome measure
			IBDQ responders at week 52, 17/98 (17.3%); p-value vs. PBO, 0.078 EPAR (Humira)[35] Discontinued corticosteroid use before week 52 and achieved resmission at week 52, 20/248 (13.3%); p-value vs. PBO, 0.035 Discontinued cosrticosteroid use for ≥ 90 days before week 52 and achieved remission at week 52, 20/248 (13.3%); p-value vs. PBO, 0.035 Discontinued cosrticosteroid use and sustained remission at both week 32 and week 52, 15/248 (10.0%); p-value vs. PBO, 0.002 IBDQ responders at week 52, 65/248 (26.2%); p-value vs. PBO, 0.007 Number and percentage of subjects taking corticosteroids at baseline who discontinued corticosteroid use and achieved clinical remission per Mayo score (ITT Analysis): Clinical remission at week 52 - discontinued CS at any time prior to week 52, 20/150 (13.3%); p-value vs. PBO, 0.035 Clinical remission at week 52 - discontinued CS for ≥90 days prior to week 52, 20/150 (13.3%); p-value vs. PBO, 0.035 59 week 52 corticosteroid-free remission - full Mayo: All ADA, 18/90 (20.0%); p-value vs. PBO, <0.05 Anti-TNF naïve ADA, 14/69 (20.3%); p-value vs. PBO, <0.05 Anti-TNF exposed ADA, 4/21 (19.0%) week 52 corticosteroid-free remission - partial Mayo: All ADA, 19/90 (21.1%); p-value vs. PBO, <0.001 Anti-TNF naïve ADA, 14/68 (20.6%); p-value vs. PBO, <0.05 Anti-TNF exposed ADA, 5/22 (22.7%); p-value vs. PBO, <0.05 week 52 corticosteroid-free - full Mayo: All ADA, 4/90 (45.6%); p-value vs. PBO, <0.001 Anti-TNF naïve ADA, 31/69 (44.9%); p-value vs. PBO, <0.05 Anti-TNF exposed ADA, 10/21 (47.6%); p-value vs. PBO, <0.05 week 52 corticosteroid-free - partial Mayo: All ADA, 43/90 (47.8%); p-value vs. PBO, <0.001

Study name	Treatment arm	Time point	Outcome measure
			Anti-TNF naïve ADA, 31/68 (45.6%); p-value vs. PBO, <0.05 Anti-TNF exposed ADA, 12/22 (54.5%); p-value vs. PBO, <0.05 [60] Post hoc analysis, week 52 n=248: Mean days in IBDQ remission (IBDQ score ≥170), 103.93; p-value vs. PBO, 0.025 Mean serious-adverse-event adjusted days in clinical remission, 81.21; p-value vs. PBO, <0.001
ULTRA3	ADA 40 mg EOW or EW	Week 52	[61] Clinical Remission at week 52 in patients Who Responded per Partial Mayo Score at week 8 - ITT-A3 protocol: Non-responder imputation, 76/196 (38.8%); modified Non-responder imputation, 84/196 (42.9%); as observed 76/131 (58.0%) Clinical Response at week 52 in patients Who Responded per Partial Mayo Score at week 8 - ITT-A3 protocol: Non-responder imputation, 113/196 (57.7%); modified Non-responder imputation, 131/196 (66.8%); as observed 113/131 (86.3%) Proportion of Patients in the ITT-A3 Population with Mayo Subscores Indicative of Mild Disease or Remission at week 52: Rectal bleeding subscore ≤1: non-responder imputation, 185/390 (47.4%); modified non-responder imputation, 246/390 (63.1%); as observed, 185/279 (67.0%) Stool frequency subscore ≤1: non-responder imputation, 145/390 (37.2%); modified non-responder imputation, 175/390 (44.9%); as observed, 145/276 (52.5%) Physician's global assessment ≤1: non-responder imputation, 169/390 (43.3%); modified non-responder imputation, 215/390 (55.1%); as observed, 169/276 (61.2%) Proportion of Patients in the ITT-A3 and ITT-E Populations with Mayo Subscores Indicative of Mild Disease or Remission at week 52: Rectal bleeding subscore ≤1: non-responder imputation, 270/575 (47.0%); modified non-responder imputation, 348/575 (60.5%); as observed, 270/290 (93.1%) Stool frequency subscore ≤1: non-responder imputation, 210/575 (36.5%); modified non-responder imputation, 251/575 (43.7%); as observed, 210/290 (72.4%) Physician's global assessment ≤1: non-responder imputation, 240/575 (41.7%); modified non-responder imputation, 299/575 (52.0%) 240/290 (82.8%)

Study name	Treatment arm	Time point	Outcome measure
			Steroid-Free Remission at week 52; Patients Using Steroids at Baseline in the ITT-A3 Population (modified non-responder imputation): Steroid-free at week 52, 131/234 (56.0%) Remission at week 52, 66/234 (28.2%) Steroid-free for ≥90 d at week 52, 118/234 (50.4%) Remission at week 52, patients who were steroid-free for ≥90 d, 61/234 (26.1%)
ULTRA3	ADA 40 mg EOW or EW	Weeks 0 to 156	[61] Remission per partial Mayo score presented graphically for weeks 0, 8, 12, 24, 36, 48, 60, 72, 84, 96, 108, 120, 132, 144, 156. Samples below from: Week 0: All non-responder imputation, 307/588 (52.2%) Entered ULTRA 3 on eow modified non-responder imputation, 252/447 (56.4%) Week 36: All non-responder imputation, 334/588 (56.8%) Entered ULTRA 3 on eow modified non-responder imputation, 254/447 (56.8%) Week 60: All non-responder imputation, 325/588 (55.3%) Entered ULTRA 3 on eow modified non-responder imputation, 229/447 (51.2%) (53.0%) Week 156: All non-responder imputation, 273/588 (46.4%) Entered ULTRA 3 on eow modified non-responder imputation, 187/447 (41.8%)
Suzuki	PBO	Week 8	[47] Rectal bleeding sub-score<=1, 65/96 (67.7%) PGA sub-score <=1, 43/96 (44.8%) Stool frequency sub-score<=1, 31/96 (32.3%) IBDQ response (increase in IBDQ score of ≥16 points from baseline), 38/96 (39.6%) Remission by baseline corticosteroid use, 10/58 (17.2%); non-use, 1/38 (2.6%) Remission by baseline immunomodulator use, 1/52 (1.9%); non-use, 10/44 (22.7%)
Suzuki	PBO	Week 32	[47] Rectal bleeding sub-score<=1, 27/96 (28.1%) PGA sub-score <=1, 27/96 (28.1%) Stool frequency sub-score<=1, 20/96 (20.8%)

Study name	Treatment arm	Time point	Outcome measure
			IBDQ response (increase in IBDQ score of ≥16 points from baseline), 2196 (21.9%) Steroid-free, 12/58 (n at baseline) (20.7%) Steroid-free remission, 5/58 (n at baseline) (8.6%)
Suzuki	PBO	Week 52	[47] Rectal bleeding sub-score<=1, 22/96 (22.9%) PGA sub-score <=1, 19/96 (19.8%) Stool frequency sub-score<=1, 13/96 (13.5%) IBDQ response (increase in IBDQ score of ≥16 points from baseline), 12/96 (12.5%) Steroid-free, 12/58 (n at baseline) (20.7%) Steroid-free remission, 4/58 (n at baseline) (6.9%) Remission by baseline corticosteroid use, 4/58 (6.9%); non-use, 3/38 (7.9%) Remission by baseline immunomodulator use, 1/52 (1.9%); non-use, 6/44 (13.6%)
Suzuki	ADA160/80mg	Week 8	[47] Rectal bleeding sub-score<=1, 64/90 (71.1%) PGA sub-score <=1, 55/90 (61.1%); p-value vs. PBO ≤ 0.05 Stool frequency sub-score<=1, 36/90 (40.0%) IBDQ response (increase in IBDQ score of ≥16 points from baseline), 38/90 (42.2%) Remission by baseline corticosteroid use, 5/57 (8.8%); non-use, 4/33 (12.1%) Remission by baseline immunomodulator use, 6/41 (14.6%), p-value vs. PBO ≤0.05; non-use, 3/49 (6.1%)
Suzuki	ADA80/40mg or ADA160/80 to week 8 then ADA40 EOW	Week 32	[47] Rectal bleeding sub-score<=1, 74/177 (41.8%) PGA sub-score <=1, 66/177 (37.3%) Stool frequency sub-score<=1, 57/177 (32.2%); p-value vs. PBO, ≤ 0.05 IBDQ response (increase in IBDQ score of ≥16 points from baseline), 55/177 (31.1%) Steroid-free, 35/120 (n at baseline) (29.2%) Steroid-free remission, 12/120 (n at baseline) (10.0%) (17.3%)
Suzuki	ADA80/40mg or ADA160/80 to week 8 then ADA40 EOW	Week 52	[47] Rectal bleeding sub-score<=1, 59/177 (33.3%) PGA sub-score <=1, 57/177 (32.2%); p-value vs. PBO, ≤ 0.05 Stool frequency sub-score<=1, 51/177 (28.8%); p-value vs. PBO, ≤ 0.05 IBDQ response (increase in IBDQ score of ≥16 points from baseline), 45/177 (25.4%); p-value vs. PBO, ≤ 0.01

Study name	Treatment arm	Time point	Outcome measure
			Steroid-free; p-value vs. PBO, ≤ 0.05; p-value vs. PBO, ≤ 0.05, 39/120 (n at baseline) (32.5%) Steroid-free remission; p-value vs. PBO, ≤ 0.05; p-value vs. PBO, ≤ 0.05, 17/120 (n at baseline) (14.2%) Remission by baseline corticosteroid use, 24/120 (20.0%); non-use, 17/57 (29.8%); p-value vs. PBO use and non-use, ≤0.05 Remission by baseline immunomodulator use, 24/79 (30.4%), p-value vs. PBO ≤0.001; non-use, 17/98 (17.3%)
PURSUIT-SC	Phase II PBO	Week 6	[48] Mean change (SD), - 1.8 (2.96), Median change from baseline in Mayo score (IQR), - 1.0 (- 4.0, 1.0)
PURSUIT-SC	Phase II GOL 200/100 mg all randomised	Week 6	[48] Mean (SD), - 2.6 (2.73) Median change from baseline in Mayo score (IQR), - 2.0 (- 4.0, 0.0) P=0.219)
PURSUIT-SC	Phase III PBO	Week 6	[48] Phase III: Mean change in CRP concn at week 6 (mg/L) PBO = + 1.59 Phase III: Week 6 Mayo score change from baseline PBO, Mean (SD)= - 1.6 (2.53), Median (IQR)= - 1.0 (- 3.0, 0.0) Phase III: Normal or inactive mucosal disease (endoscopy score = 0) at week 6 PBO = 10/251 (4.0)
PURSUIT-SC	Phase III GOL 200/100 mg phase III	Week 6	[48] Phase III: Mean change in CRP concn at week 2 (mg/L)GOL 200/100 = - 6.57 (P<0.0001) Phase III: Mean change in CRP concn at week 6 (mg/L) GOL 200/100 = - 3.35 (P<0.0001) Phase III: Week 6 Mayo score change from baseline GOL 200/100 mg, Mean (SD)= - 3.1 (2.90), Median (IQR)= - 3.0 (- 6.0, 0.0) (P<0.0001) Phase III: Normal or inactive mucosal disease (endoscopy score = 0) at week 6 GOL 200/100 mg = 21/253 (8.3) (P=0.0437)
PURSUIT-SC	PBO	Weeks 0 to 6	[341] Stool frequency at week 0 (all mean (SD)) PBO= 2.3 (0.8) Stool frequency at week 2 PBO=2.1 (0.9) Stool frequency at week 4 PBO=1.9 (0.9) Stool frequency at week 6 PBO=2 (1) (as reported) Rectal bleeding score at week 0 (all mean (SD)) PBO=1.50(0.86) Rectal bleeding score at week 2 PBO=1.20 (0.91) Rectal bleeding score at week 4 PBO=1.04 (0.94) Rectal bleeding score at week 6 PBO=1.04 (0.94)
PURSUIT-M	PBO	Week 54	[49] PBO: 24.1% (N=54) maintained clinical remission among those who were in clinical remission at baseline. 54% total patients receiving corticosteroids at baseline. PBO: Of these, 18.4% (N=87) were in corticosteroid-free clinical remission at Week 54

Study name	Treatment arm	Time point	Outcome measure
			(ie. achieved corticosteroid-free clinical remission at Week 54 among those who were receiving corticosteroids at baseline). Maintained clinical response through Week 54 and corticosteroid-free at Week 54 among those receiving corticosteroids at PURSUIT-M baseline. PBO= 18/87 (20.7) 169 (37.1%) patients in primary analysis popn had dose adjustment. PBO = 75 (48.7%) Reduction in median partial Mayo scores observed at baseline of PURSUIT-M among GOL-induction responders (ie decrease of 4 points from induction baseline) maintained in 100 mg and 50 mg groups through weeks 52 and 48 respectively (but in PBO group increased after week 8 and increased to value approaching that an induction baseline at week 54. Proportion of patients with normal or inactive mucosal disease (ie endoscopy score = 0) at week 54 = 13.0%
PURSUIT-M	GOL 50 mg	Week 54	[49] 36.5% (N=52) of patients in clinical remission at baseline maintained clinical remission (P=0.365). 28.2% (n=78, P=0.279) were in corticosteroid-free clinical remission at Week 54 Maintained clinical response through Week 54 and corticosteroid-free at Week 54 among those receiving corticosteroids at PURSUIT-M baseline. 30/78 (38.5) (P=0.026) GOL 50 mg. 51 (33.8%) had dose adjustment Proportion of patients with normal or inactive mucosal disease (ie endoscopy score = 0) at week 54 = 25.8% (P=0.011)
PURSUIT-M	GOL 100 mg	Week 54	[49] 23.3% were in corticosteroid-free clinical remission at Week 54 (N=82, P=0.423) Maintained clinical response through Week and corticosteroid-free at Week 54 among those receiving corticosteroids at PURSUIT-M baseline. 25/82 (30.5) (P=0.138) GOL 100 mg. 43 (28.5%) had dose adjustment. Proportion of patients with normal or inactive mucosal disease (ie endoscopy score = 0) at week 54 = 21.9% (P=0.033)
UC-SUCCESS	AZA	Week 8	[52] Patients with partial Mayo score decrease of ≥1 : 50/76 (65.79%); p-value between IFX, 0.002; IFX/AZA, 0.003 Patients with partial Mayo score decrease of ≥2 : 28/76 (36.84%); Change (SD) in partial Mayo scores from baseline , -2.81 (2.46) Fecal calprotectin ≤50ug/g, 12/62 (19.4%); ≤250ug/g, 24/62 (38.7%) ≥251 µg/g not extracted
UC-SUCCESS	AZA	Week 16	[52] Patients with Mayo score response: 38/76 (50.00%); p-value between IFX, 0.018; IFX/AZA, 0.001 Total Mayo score change from baseline, mean; n=71: -3.00 (baseline 8.50); p-value between IFX, 0.013; IFX/AZA, 0.001

Study name	Treatment arm	Time point	Outcome measure
			Change (SD) in partial Mayo scores from baseline, -2.34 (2.70) A post hoc analysis was conducted to determine the proportion of patients who achieved a Mayo endoscopy subscore of 0 only at week 16. A greater proportion of patients treated with IFX/AZA combination therapy (29.5%) achieved a Mayo endoscopy subscore of 0 than patients given monotherapy with IFX (11.7%; p=0 .006) and AZA (13.2%; p=0.014). The difference between the IFX group and the AZA group was not statistically significant (p=0.783). Fecal calprotectin ≤50ug/g, 12/66 (18.2%); ≤250ug/g, 29/66 (43.9%) [342] Change from baseline (assume mean): Stool frequency, -0.97 Rectal bleeding, -0.77 Physician global assessment, -0.59 Total Mayo, -3.00
UC-SUCCESS	IFX	Week 8	[52] Patients with partial Mayo score decrease of ≥1: 68/77 (88.31%); p-value between IFX/AZA, 0.654 Patients with partial Mayo score decrease of ≥2: 38/77 (49.35%) Change (SD) in partial Mayo scores from baseline , -3.52 (2.25) Fecal calprotectin ≤50ug/g, 15/66 (22.7%); ≤250ug/g, 33/66 (50.7%) ≥251 µg/g not extracted
UC-SUCCESS	IFX	Week 16	[52] Patients with Mayo score response: 53/77 (68.83%); p-value between IFX/AZA, 0.514 Total Mayo score change from baseline, mean; n=70: -4.27 (baseline 8.08); p-value between IFX/AZA, 0.001 Change (SD) in partial Mayo scores from baseline, -3.43 (2.26) Fecal calprotectin ≤50ug/g, 11/62 (17.7%); ≤250ug/g, 19/62 (30.6%) [342] Change from baseline (assume mean): Stool frequency, -1.23 Rectal bleeding, -1.14; p-value vs. AZA, <0.05 Physician global assessment, -1.06; p-value vs. AZA, <0.05 Total Mayo, -4.27; p-value vs. AZA, <0.05
UC-SUCCESS	IFX/AZA	Week 8	[52] Patients with partial Mayo score decrease of ≥1: 67/78 (85.90%) Patients with partial Mayo score decrease of ≥2: 41/78 (52.56%) Change (SD) in partial Mayo scores from baseline, -4.01 (SD, 2.04); p-value vs. AZA 0.005

Study name	Treatment arm	Time point	Outcome measure
			Fecal calprotectin ≤50ug/g, 26/63 (41.3%); ≤250ug/g, 42/63 (66.7%) ≥251 µg/g not extracted
UC-SUCCESS	IFX/AZA	Week 16	[52] Patients with Mayo score response: 60/78 (76.92%) Total Mayo score change from baseline, mean (SD); n=76: -5.28 (baseline 8.54) Change (SD) in partial Mayo scores from baseline, -4.09 (SD, 2.18); p-value vs. IFX <0.001 Fecal calprotectin ≤50ug/g, 22/70 (31.4%); ≤250ug/g, 41/70 (58.6%) [342] Change from baseline (assume mean): Stool frequency, -1.54; p-value vs. AZA, <0.05 Rectal bleeding, -1.25; p-value vs. AZA, <0.05 Physician global assessment, -1.30; p-value vs. AZA and IFX, <0.05 Total Mayo, -5.28; p-value vs. AZA and IFX, <0.05
Probert	PBO	Week 6	[51] UCSS Mean (SD), 5 (3); improvement in UCSS, 4 (SD 3); median improvement 3. Median between-group difference p=0.82 Baron score Mean SD, 1 (1); proportion of patients with a Baron score of 0, 6/20 (30%). 95% CI for difference −30% to 23%); p=0.96. Mean (SD) improvement in Baron score, 1 (SD 1). Baron score improved by a decrease in score of at least 1 in 3/20 (13%); seven (37%) remained the same, and one underwent colectomy; p=0.67 When remission rates of patients with total disease in each of the two groups were compared, no significant difference was found (p=0.9) Remission rate in patients receiving azathioprine, 2/6 (33%). 95% CI for difference −79% to 45%; p=0.89 Mean reduction in daily dose of glucocorticoid was equivalent to 14 mg prednisolone (SD 12); p=0.037 compared to IFX Week 6: CRP median value did not change (data not reported); p-value 0.96 but unclear if change from baseline or between IFX
Probert	IFX	Week 6	[51] UCSS Mean (SD), 5 (3); improvement in UCSS (n=18), 4 (SD 3); median improvement 2.5 for the 18 assessable patients Baron score Mean SD, 1 (1); proportion of patients with a Baron score of 0, 6/23 (26%). Mean (SD) improvement in Baron score, 1 (SD 1). Baron score improved by a decrease in score of at least 1 in 13/23 (57%); seven (30%) remained the same, and three (13%) deteriorated 5/14 (36%) with total colitis went into remission, 3/5 (60%) with left sided colitis and 1/4 (25%) with distal colitis (p=0.5) Remission rate in patients receiving azathioprine, 4/6 (67%) Mean reduction in daily dose of glucocorticoid was equivalent to 19 mg prednisolone (SD 15)

Study name	Treatment arm	Time point	Outcome measure
			Week 6: CRP median levels rose from 6.5 to 10 mg/l
ACT1	PBO	Week 2	[50] Partial Mayo score Median (IQR) at baseline , 6.0 (5.0-7.0 Partial Mayo score Median (IQR) at week 2 , 5.0 (4.0-6.0)
ACT1	PBO	Week 6	[50] Partial Mayo score Median (IQR) at week 6 , 5.0 (3.0-6.0)
ACT1	PBO	Week 8	[50] Refractory to corticosteroid therapy, 35.3 (12/34) Not refractory to corticosteroid therapy, 37.9 (33/87) Partial Mayo score Median (IQR) at week 8 , 5.0 (3.0-6.0) Daily corticosteroid dose in mg (median, IQR) at baseline. 20.0 (10.0-30.0) Daily corticosteroid dose in mg (median, IQR) at week 8. 20.0 (10.0-30.0)
ACT1	PBO	Week 30	[50] Partial Mayo score Median (IQR) at week 30 , 5.0 (3.0-6.0) Daily corticosteroid dose in mg (median, IQR). 10.0 (0.8-30.0)
ACT1	PBO	Week 54	[50] Partial Mayo score Median (IQR) at week 54 , 5.0 (4.0-7.0) Clinical remission and discontinued use of corticosteroids at week 54, 7/79 (8.9) Daily corticosteroid dose in mg (median, IQR). 20.0 (0.0-30.0) [343] Clinical response: Baseline IMM use: 26% (14/53) No baseline IMM use: 15% (10/68) Clinical remission: Baseline IMM use: 21% (11/53) No baseline IMM use: 13% (9/68)

Study name	Treatment arm	Time point	Outcome measure
ACT1	IFX 5mg/kg	Week 2	[50] Partial Mayo score Median (IQR) at baseline 6.0 (5.0-7.0) Partial Mayo score Median (IQR) at week 2 3.0 (2.0-5.0)
ACT1	IFX 5mg/kg	Week 6	[50] Partial Mayo score Median (IQR) at week 6 3.0 (2.0-5.0)
ACT1	IFX 5mg/kg	Week 8	[50] Refractory to corticosteroid therapy, 77.4 (24/31) (P<0.001) Not refractory to corticosteroid therapy, 66.7 (60/90) (P<0.001) Partial Mayo score Median (IQR) at week 8 2.0 (1.0-4.0) Daily corticosteroid dose in mg (median, IQR) at baseline. 20.0 (10.0-25.0) Daily corticosteroid dose in mg (median, IQR) at week 8. 20.0 (10.0-25.0)
ACT1	IFX 5mg/kg	Week 30	[50] Partial Mayo score Median (IQR) at week 30 3.0 (1.0-6.0) Clinical remission and discontinued use of corticosteroids at week 30, 17/70 (24.3) (P=0.030) Daily corticosteroid dose in mg (median, IQR) at week 30. 5.6 (0.0-20.0)
ACT1	IFX 5mg/kg	Week 54	[50] Partial Mayo score Median (IQR) at week 54 3.0 (1.0-6.0) Clinical remission and discontinued use of corticosteroids at week 54, 18/70 (25.7) (P=0.006) Daily corticosteroid dose in mg (median, IQR) at week 54. 5.0 (0.0-20.0) [50] Clinical response: Baseline IMM use: 48% (32/66) OR 2.62 (95% CI 1.20 to 5.71) No baseline IMM use: 42% (23/55) OR 4.17 (95% CI 1.77 to 9.84) Clinical remission: Baseline IMM use: 35% (23/66) OR 2.04 (95% CI 0.89 to 4.71) No baseline IMM use: 35% (19/55) OR 3.46 (95% CI 1.41 to 8.47)
ACT1	IFX combined	Week 54	[50] Clinical response:

Study name	Treatment arm	Time point	Outcome measure
			Baseline IMM use: 45% (56/125) OR 2.26 (95% CI 1.12 to 4.58) No baseline IMM use: 45% (53/118) OR 4.73 (95% CI 2.21 to 10.1) Clinical remission: Baseline IMM use: 34% (42/125) OR 1.93 (95% CI 0.90 to 4.13) No baseline IMM use: 36% (42/118) OR 3.62 (95% CI 1.63 to 8.03)
ACT2	PBO	Week 2	[50] Partial Mayo score Median (IQR) at baseline , 6.0 (5.0-7.0)) Partial Mayo score Median (IQR) at week 2 , 5.0 (4.0-7.0)
ACT2	PBO	Week 6	[50] Partial Mayo score Median (IQR) at week 6, 5.0 (4.0-7.0),)
ACT2	PBO	Week 8	[50] Refractory to corticosteroid therapy, 37.5 (12/32) Not refractory to corticosteroid therapy, 37.5 (12/32) Partial Mayo score Median (IQR) at week 8 , 5.0 (3.0-7.0),) Daily corticosteroid dose in mg (median, IQR) at baseline. , 20.0 (15.0-30.0), Daily corticosteroid dose in mg (median, IQR) at week 8. , 20.0 (15.0-30.0

Study name	Treatment arm	Time point	Outcome measure
ACT2	PBO	Week 30	[50] Partial Mayo score Median (IQR) at week 30 , 6.0 (3.0-7.0), Daily corticosteroid dose in mg (median, IQR), 20.0 (5.6-30.0 Clinical remission and discontinued use of corticosteroids, 2/60 (3.3) [50] Clinical response: Baseline IMM use: 26% (14/54) No baseline IMM use: 26% (18/69) Clinical remission: Baseline IMM use: 9% (5/54) No baseline IMM use: 12% (8/69)
ACT2	PBO	Week 54	[50] Partial Mayo score Median (IQR) at week 54 , NR, Daily corticosteroid dose in mg (median, IQR) at week 54. , NR, [50] Clinical response at week 54 WITH baseline immunomodulator use OR= 3.09 (1.36, 6.98)
ACT2	IFX 5mg/kg	Week 2	[50] Partial Mayo score Median (IQR) at baseline, 6.0 (5.0-7.0) Partial Mayo score Median (IQR) at week 2, 4.0 (2.0-5.0)
ACT2	IFX 5mg/kg	Week 6	[50] Partial Mayo score Median (IQR) at week 6, 3.0 (1.0-5.0)

Study name	Treatment arm	Time point	Outcome measure
ACT2	IFX 5mg/kg	Week 8	[50] Refractory to corticosteroid therapy,63.3 (19/30) (P=0.053) Not refractory to corticosteroid therapy, 63.3 (19/30) (P=0.053) Partial Mayo score Median (IQR) at week 8, 2.0 (1.0-4.0) Daily corticosteroid dose in mg (median, IQR) at baseline. 5 mg/kg IFX= 20.0 (10.0-30.0 Daily corticosteroid dose in mg (median, IQR) at week 8. 5 mg/kg IFX= 20.0 (10.0-30.0)
ACT2	IFX 5mg/kg	Week 30	[50] Partial Mayo score Median (IQR) at week 30, 4.0 (1.0-6.0) Daily corticosteroid dose in mg (median, IQR), 7.5 (0.0-20.0) Clinical remission and discontinued use of corticosteroids, 11/60 (18.3) (P=0.010) [50] Clinical response: Baseline IMM use: 52% (27/52) OR 3.09 (95% CI 1.36 to 6.98) No baseline IMM use: 26% (18/69) 43% (30/69) 61% (43/70) 53% (73/139) OR 2.18 (95% CI 1.06 to 4.47) Clinical remission: Baseline IMM use: 35% (18/52) OR 5.19 (95% CI 1.76 to 15.3) No baseline IMM use: 19% (13/69) OR 1.77 (95% CI 0.68 to 4.59)
ACT2	IFX 5mg/kg	Week 54	[50] Partial Mayo score Median (IQR) at week 54, NR Daily corticosteroid dose in mg (median, IQR), NR [50] Clinical response WITHOUT baseline immunomodulator use OR= 2.18 (1.06, 4.47)
ACT2	IFX combined	Week 54	[50] Clinical remission WITH baseline immunomodulator use OR= 5.19 (1.76, 15.3)

Study name	Treatment arm	Time point	Outcome measure
			[50] Clinical remission WITHOUT baseline immunomodulator use OR= 1.77 (0.68, 4.59)
ACT 1 and 2 extension studies	Randomised patients in the infliximab group of the ACT-1 or ACT-2 trials who entered the extension studies.	Weeks 0 to 152	[55] Patients with no disease activity (PGA assessment of no disease): Week E0: ACT1, 55.7% (64/115); ACT2, 27.9% (31/111) Week E24: ACT1, 66.4% (73/110); ACT2, 43.3% (42/97) Week E48: ACT1, 70.8% (75/106); ACT2, 52.3% (46/88) Week E72: ACT1, 72.7% (72/99); ACT2, 52.9% (45/85) Week E104: ACT1, 69.5% (57/82); ACT2, 66.2% (51/77) Week E128: ACT1, 79.5% (35/44); ACT2, 66.0% (35/53) Week E152: ACT1, 88.9% (8/9); ACT2, 45.5% (5/11) Patients with no or mild disease activity (PGA assessment of no or mild disease):: Week E0: ACT1, 84.3% (97/115); ACT2, 68.5% (76/111) Week E24: ACT1, 90.9% (100/110); ACT2, 91.8% (89/97) Week E48: ACT1, 96.2% (102/106); ACT2, 92.0% (81/88) Week E72: ACT1, 94.9% (94/99); ACT2, 88.2% (75/85) Week E104: ACT1, 96.3% (79/82); ACT2, 92.2% (71/77) Week E128: ACT1, 95.5% (42/44); ACT2, 90.6% (48/53) Week E152: ACT1, 100.0% (9/9); ACT2, 81.8% (9/11) Randomised patients in the infliximab group of the extension studies with and without a gap in treatment of more than 8 weeks between the last infusion of the main studies and the extension studies week 0 infusions (n=134). Patients with no disease activity: Week E0: Patients without treatment gap (n=134), 52.3% (69/132); Patients with treatment gap (n=95), 27.7% (26/94) Week E8: Patients without treatment gap (n=134), 59.4% (76/128); Patients with treatment gap (n=95), 44.6% (41/92) Week E24: Patients without treatment gap (n=134), 55.6% (69/124); Patients with treatment gap (n=95), 55.4% (46/83) Week E48: Patients without treatment gap (n=134), 64.8% (79/122); Patients with treatment gap (n=95), 58.3% (42/72) Week E72: Patients without treatment gap (n=134), 64.6% (73/113); Patients with treatment gap (n=95), 62.0% (44/71) Week E104: Patients without treatment gap (n=134), 70.5% (67/95); Patients with treatment gap (n=95), 64.1% (41/64) Week E128: Patients without treatment gap (n=134), 72.5% (37/51); Patients with treatment gap (n=95), 71.7% (33/46) Week E152: Patients without treatment gap (n=134), 62.5% (5/8); Patients with treatment gap (n=95), 66.7% (8/12) Patients with no or mild disease activity:

Study name	Treatment arm	Time point	Outcome measure
			Week E0: Patients without treatment gap (n=134), 84.1% (111/132); Patients with treatment gap (n=95), 66.0% (62/94) Week E8: Patients without treatment gap (n=134), 87.5% (112/128); Patients with treatment gap (n=95), 81.5% (75/92) Week E24: Patients without treatment gap (n=134), 92.7% (115/124); Patients with treatment gap (n=95), 89.2% (74/83) Week E48: Patients without treatment gap (n=134), 94.3% (115/122); Patients with treatment gap (n=95), 94.4% (68/72) Week E72: Patients without treatment gap (n=134), 91.2% (103/113); Patients with treatment gap (n=95), 93.0% (66/71) Week E104: Patients without treatment gap (n=134), 90.5% (86/95); Patients with treatment gap (n=95), 100% (64/64) Week E128: Patients without treatment gap (n=134), 94.1% (48/51); Patients with treatment gap (n=95), 91.3% (42/46) Week E152: Patients without treatment gap (n=134), 87.5% (7/8); Patients with treatment gap (n=95), 91.7% (11/12) All randomised patients in the infliximab group who entered the extension studies (n=229): Week E8: Patients with No Disease Activity (PGA assessment of no disease), 46.4% (102/220); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 70.9% (156/220) Week E24: Patients with No Disease Activity (PGA assessment of no disease), 50.7% (105/207); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 80.7% (167/207) Week E48: Patients with No Disease Activity (PGA assessment of no disease), 56.7% (110/194); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 82.0% (159/194) Week E72: Patients with No Disease Activity (PGA assessment of no disease), 58.7% (108/184); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 83.2% (153/184) Week E104: Patients with No Disease Activity (PGA assessment of no disease), 66.0% (105/159); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 88.1% (140/159) Week E128: Patients with No Disease Activity (PGA assessment of no disease), 69.1% (67/97); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 87.6% (85/97) Week E152: Patients with No Disease Activity (PGA assessment of no disease), 65.0% (13/20); Patients with No or Mild Disease Activity (PGA assessment of no or mild disease), 90.0% (18/20) Number of randomised patients in the extension studies (n=229) who used corticosteroids in the past 8 weeks for UC: Week E8: 0 days, 179/223 (80.3%); 1 to 7 days, 4/223 (1.8%); 8 to 30 days, 8/223 (3.6%); > 30 days, 32/223 (14.3%) Week E24: 0 days, 179/208 (86.1%); 1 to 7 days, 1/208 (0.5%); 8 to 30 days, 2/208 (1.0%); > 30 days, 26/208 (12.5%) Week E48: 0 days, 167/194 (86.1%); 1 to 7 days, 2/194 (1.0%); 8 to 30 days, 4/194 (2.1%); > 30 days, 21/194 (10.8%) Week E72: 0 days, 165/188 (87.8%); 1 to 7 days, 3/188 (1.6%); 8 to 30 days, 4/188 (2.1%); > 30 days, 16/188 (8.5%) Week E104: 0 days, 149/161 (92.5%); 1 to 7 days, 0/161 (0.0%); 8 to 30 days, 1/161 (0.6%); > 30 days, 11/161 (6.8%) Week E128: 0 days, 92/99 (92.9%); 1 to 7 days, 0/99 (0.0%); 8 to 30 days, 0/99 (0.0%); > 30 days, 7/99 (7.1%) Week E152: 0 days, 20/20 (100.0%); 1 to 7 days, 0/20 (0.0%); 8 to 30 days, 0/20 (0.0%); > 30 days, 0/20 (0.0%)

Additional efficacy outcomes (paediatric population trial)

Study name	Treatment arm	Time piont	Outcome measure
Hyams	IFX 5mg/q8w	Week 8	[53] Median reduction in Partial Mayo Score 4 points Median corticoteroid use mg/kg/d: 0
Hyams	IFX 5mg/q8w	Week 30	[53] Median reduction in Partial Mayo Score: approx 2.5 points (read from graph) Remission (PUCAI) without corticosteroids: 5/12 (41.7%) EPAR[73] Remission (PUCAI) by age (n=20 evaluable): 6 yrs, 0/0 (0%); 7yrs, 2/2 (100%); 8yrs, 0/0 (0%); 9yrs, 1/1 (100%); 10yrs, 0/1 (0%); 11yrs, 1/1 (100%); 12yrs, 0/1 (0%); 13yrs, 2/4 (50%); 14yrs, 0/0 (0%); 15yrs, 1/3 (33.3%); 16yrs, 0/3 (0%); 17yrs, 1/4 (25%)
Hyams	IFX 5mg/q8w	Week 54	[53] Median reduction in Partial Mayo Score: approx 2.5 points (read from graph) Remission (PUCAI) without corticosteroids: 5/13 (38.5%) Efficacy after step-up decrease of ≥2 points in partial Mayo score - patients with data at week 54: 9/10 (90%). Unclear if this value is for IFX 5mg/q8d group or both groups Median corticoteroid use mg/kg/d: 0.04 EPAR[73] Clinical response: 3/4 patients who had endoscopy at week 54 (optional) Remission (PUCAI) by age (n=20 evaluable): 6 yrs, 0/0 (0%); 7yrs, 1/2 (50%); 8yrs, 0/0 (0%); 9yrs, 1/1 (100%); 10yrs, 0/1 (0%); 11yrs, 1/1 (100%); 12yrs, 0/1 (0%); 13yrs, 3/4 (75%); 14yrs, 0/0 (0%); 15yrs, 1/4 (25.3%); 16yrs, 0/3 (0%); 17yrs, 1/4 (25%)
Hyams	IFX 5mg/q12w	Week 8	[53] Median reduction in Partial Mayo Score: 4 points Median corticoteroid use mg/kg/d: 0.15
Hyams	IFX 5mg/q12w	Week 30	[53] Median reduction in Partial Mayo Score: approx 1 point (read from graph) Remission (PUCAI) without corticosteroids: 1/13 (7.7%) EPAR[73] Remission (PUCAI) by age (n=21 evaluable):

Study name	Treatment arm	Time piont	Outcome measure
			6 yrs, 0/1 (0%); 7yrs, 0/0 (0%); 8yrs, 0/1 (0%); 9yrs, 0/0 (0%); 10yrs, 0/1 (0%); 11yrs, 0/1 (0%); 12yrs, 0/0 (0%); 13yrs, 0/0 (0%); 14yrs, 0/2 (0%); 15yrs, 3/4 (75%); 16yrs, 1/5 50 (20%); 17yrs, 0/5 (0%)
Hyams	IFX 5mg/q12w	Week 54	[53] Median reduction in Partial Mayo Score: approx 1 point (read from graph) Remission (PUCAI) without corticosteroids: 0/13 (0%) Median corticoteroid use mg/kg/d: same as baseline 4 EPAR [73] Clinical response: 3/4 patients who had endoscopy at week 54 (optional) Remission (PUCAI) by age (n=22 evaluable): 6 yrs, 0/1 (0%); 7yrs, 0/0 (0%); 8yrs, 0/1 (0%); 9yrs, 0/0 (0%); 10yrs, 0/1 (0%); 11yrs, 0/1 (0%); 12yrs, 0/2 (0%); 13yrs, 0/0 (0%) 3/4 (75%); 14yrs, 0/2 (0%); 15yrs, 2/4 (50%) 3/8 (37.5%); 16yrs, 1/5 (20%) 1/8 (12.5%); 17yrs, 1/5 (20%) 2/9 (22.2%)
Hyams	All patients (n=60)	Week 8	[53] Disease activity was more severe at the last visit for patients who discontinued after week 8 (no disease, 1 of 10 [10%]; mild, 1 of 10 [10%]; moderate, 6 of 10 [60%]; severe, 2 of 10 [20%]) than for patients who discontinued before week 8 (mild, 4 of 13 [30.8%]; moderate, 4 of 13 [30.8%]; and severe disease, 5 of 13 [38.5%])

Hospitalisation, surgery and mortality data (adult population trials)

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
Adalimumab				
ULTRA1	PBO	From submission[63] Week 8 (n=222, person years=19.6) Physician visits No. events (visits) (events/PY), 21 (0.619) Emergency room visits No. events (visits) (events/PY), 8 (0.236) Hospital admissions No. events (admissions) (events/PY), 7 (0.206) Days in hospital No. events (days) (events/PY), 73 (2.153)	Colectomy, 8/130 (3.6%) during induction, week 8. Elective/emergency NR	0/223 (0%)
ULTRA1	ADA 160/80mg	From submission[63] Week 8 (n=223, person years=34.0) Physician visits No. events (visits) (events/PY), 15 (0.441); p-value 0.559 Emergency room visits No. events (visits) (events/PY), 2 (0.059); p-value NA Hospital admissions No. events (admissions) (events/PY), 5 (0.147); p-value NA Days in hospital No. events (days) (events/PY), 26 (0.764); p-value 0.297	Colectomy, 5/130 (1.4%) during induction, week 8 Elective/emergency NR	0/223 (0%)
ULTRA2	PBO	From submission[63] (n=246, person years=101.6) Physician visits No. events (visits) (events/PY), 169 (1.663) Emergency room visits No. events (visits) (events/PY), 10 (0.098) Hospital admissions No. events (admissions) (events/PY), 13 (0.128) Days in hospital No. events (days) (events/PY), 105 (0.837)	Colectomy, 12/246 (4.9%) during follow-up week 52. Elective/emergency NR	0/260 (0%)

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
ULTRA2	ADA 160/80mg	From submission[63] (n=248, person years=125.5) Physician visits No. events (visits) (events/PY), 169 (1.347); p-value vs. PBO 0.035 Emergency room visits No. events (visits) (events/PY), 12 (0.096); p-value vs. PBO 0.847 Hospital admissions No. events (admissions) (events/PY), 13 (0.104); p-value vs. PBO 0.418 Days in hospital No. events (days) (events/PY), 120 (1.181); p-value vs. PBO 0.467	Colectomy, 10/248 (4%) during follow-up, week 52. Elective/emergency NR	0/257 (0%)
ULTRA 1 and 2	ADA	[64] <u>ADA Hospitalisation and Colectomy Rates in ULTRA 1 and 2: Week 8 ADA Responders:</u> All-cause hospitalisation n/patient years at Risk, 46/260.4 All-cause hospitalisation Incidence rate (n/PYs at Risk), 0.18; All-cause hospitalisation p-value vs. PBO, 0.047 UC-related hospitalisation n/patient years at Risk, 29/266.5 UC-related hospitalisation Incidence rate (n/PYs at Risk), 0.11 UC-related hospitalisation p-value vs. PBO, 0.002 Colectomy n/patient years at Risk, 6/271.9. Elective/emergency NR Colectomy Incidence rate (n/PYs at Risk), 0.02. Colectomy p-value vs. PBO, 0.122 Hospitalisations - all-cause events/patient years, 55/272.7 Hospitalisations - all-cause Incidence rate (events/person years), 0.20 Hospitalisations - all-cause relative risk ADA/PBO, 0.65 p=0.021 Hospitalisations - UC-related events/patient years, 32/272.7 Hospitalisations - UC-related Incidence rate (events/person years), 0.12 Hospitalisations - UC-related relative risk ADA/PBO, 0.48 p<0.001 [66] <u>Non UC-related hospitalisation categories week 52:</u> General disorder; gastrointestinal tract disorder, 3 (0.63%); gynecological disorder and		

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
		pregnancy, 1 (0.21%); musculoskeletal and connective tissue disorder; hepatobiliary disorder, 1 (0.21%); neurological disorder, 1 (0.21%); urogenital tract disorder, 3 (0.63%); cardiovascular disorder, 2 (0.42%); endocrine and metabolic disorder; hematologic disorder, 1 (0.21%); infection, 11 (2.28%) 9 (1.88%); malignancy, 1 (0.21%); skin disorder, 2 (0.42%); trauma and surgical/medical procedure, 3 (0.63%) 66 <u>UC-related hospitalization categories week 52:</u> UC flare, 47 (9.73%) 31 (6.46%); UC leading to colectomy, 19 (3.93%) 15 (3.13%); Extra-intestinal complication of UC, 6 (1.25%); sequelae of colectomy, 1 (0.21%) 66 <u>Hospitalisation and Colectomy Analysis: Induction Period (8 weeks)</u> All cause hospitalisation, 22 (4.6) UC-related hospitalisation, 17 (3.5%) UC- or drug related hospitalisation, 19 (4.0%) Colectomy, 5 (1.0%). Elective/emergency NR 66 <u>Hospitalisation and Colectomy Analysis: 52-Week Period, n/patient years at risk (incident rate); RR (relative risk) (95% CI):</u> All-cause hospitalisation, 69/387.5 (0.18); RR, 0.7 (0.5 to 1.0), p=0.03 UC-related hospitalisation, 47/398.1 (0.12) ; RR, 0.5 (0.4 to 0.8), p=0.002 UC- or drug-related hospitalisation, 55/393.8 (0.14) ; RR, 0.6 (0.4 to 0.9), p=0.005 Colectomy, 15/408.1 (0.04) Sensitivity analysis 1: all events that occurred during the open-label period (during ADA therapy) were excluded for the placebo group. All-cause hospitalisation, 69/387.5 (0.18); RR, 0.6 (0.4 to 0.9), p=0.007 UC-related hospitalisation, 47/398.1 (0.12); RR, 0.5 (0.3 to 0.7) <.001 UC- or drug-related hospitalisation, 55/393.8 (0.14); RR, 0.5 (0.4 to 0.8), p=0.001 Colectomy, 15/408.1 (0.04). Elective/emergency NR Sensitivity analysis 2: all events were attributed to the randomized groups regardless of whether		

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
		patients treated with placebo had switched to open-label ADA therapy All-cause hospitalisation, 69/387.5 (0.18) ; RR, 0.8 (0.6 to 1.0), p=0.08 UC-related hospitalisation, 47/398.1 (0.12) 0.7; RR, (0.5 to 1.0), p=0.03 UC- or drug-related hospitalisation, 55/393.8 (0.14) ; RR, 0.7 (0.5 to 1.0), p=0.045 Colectomy, 15/408.1 (0.04). Elective/emergency NR [344] <u>Incidence rates for All-cause and UC-related hospitalisations for ADA-treated pts by Mayo subscores at wk 8:</u> Mayo subscore 0 (n=433): Stool frequency - all cause, 0.08 UC-related, 0.05; rectal bleeding - all cause, 0.13 UC-related, 0.06; PGA - all cause, 0.11 UC-related, 0.05; endoscopy - all cause, 0.14 UC-related, 0.06 Mayo subscore 1 (n=433): Stool frequency - all cause, 0.11 UC-related, 0.05; rectal bleeding - all cause, 0.13 UC-related, 0.10; PGA - all cause, 0.11 UC-related, 0.06; endoscopy - all cause, 0.11 UC-related, 0.06 Mayo subscore 2 (n=433): Stool frequency - all cause, 0.15 UC-related, 0.11; rectal bleeding - all cause, 0.22 UC-related, 0.13; PGA - all cause, 0.20 UC-related, 0.14; endoscopy - all cause, 0.16 UC-related, 0.10 Mayo subscore 3 (n=422): Stool frequency - all cause, 0.22 UC-related, 0.16; rectal bleeding - all cause, 0.29 UC-related, 0.29; PGA - all cause, 0.23 UC-related, 0.19; endoscopy - all cause, 0.27 UC-related, 0.21		
ULTRA 1 and 2	PBO	[64] <u>PBO Hospitalisation and Colectomy Rates in ULTRA 1 and 2: Week 8</u> ADA Responders: All-cause hospitalisation n/patient years at Risk, 58/222.3 All-cause hospitalisation Incidence rate (n/PYs at Risk), 0.26 UC-related hospitalisation n/patient years at Risk, 49/223.6 UC-related hospitalisation Incidence rate (n/PYs at Risk), 0.22 Colectomy n/patient years at Risk, 11/231.7. Elective/emergency NR Colectomy Incidence rate (n/PYs at Risk), 0.05 Hospitalisations - all-cause events/patient years, 71/232.8 Hospitalisations - all-cause Incidence rate (events/person years), 0.31		

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
		Hospitalisations - UC-related events/patient years, 59/232.8 Hospitalisations - UC-related Incidence rate (events/person years), 0.25 66 <u>Non UC-related hospitalisation categories week 52:</u> General disorder, 1 (0.21%); gastrointestinal tract disorder, 1 (0.21%); gynecological disorder and pregnancy, 2 (0.41%); musculoskeletal and connective tissue disorder, 1 (0.21%); hepatobiliary disorder, 0 (0%); neurological disorder, 0 (0%); urogenital tract disorder, 3 (0.62%); cardiovascular disorder, 1 (0.21%); endocrine and metabolic disorder, 1 (0.21%); hematologic disorder, 0 (0%); infection, 11 (2.28%); malignancy, 1 (0.21%); skin disorder, 1 (0.21%); trauma and surgical/medical procedure, 3 (0.62%) 66 <u>UC-related hospitalisation categories week 52:</u> UC flare, 47 (9.73%); UC leading to colectomy, 19 (3.93%); Extra-intestinal complication of UC, 8 (1.66%); sequelae of colectomy, 1 (0.21%) 66 <u>Hospitalisation and Colectomy Analysis: Induction Period (8weeks)</u> All cause hospitalisation, 37 (7.7%) p=0.46 UC-related hospitalisation, 34 (7.0%) p=0.02 UC- or drug related hospitalisation, 36 (7.5) p=0.02 Colectomy, 6 (1.2%) p=0.77 66 <u>Hospitalisation and Colectomy Analysis: 52-Week Period, n/patient years at risk (incident rate):</u> All-cause hospitalisation, 58/222.3 (0.26) UC-related hospitalisation, 49/223.6 (0.22) UC- or drug-related hospitalisation, 53/223.2 (0.24) Colectomy, 11/231.7 (0.05). Elective/emergency NR Sensitivity analysis 1: all events that occurred during the open-label period (during ADA therapy) were excluded for the placebo group.		

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
		All-cause hospitalisation, 47/159.2 (0.30) UC-related hospitalisation, 40/159.8 (0.25) UC- or drug-related hospitalisation, 43/159.5 (0.27) Colectomy, 7/166.3 (0.04). Elective/emergency NR <u>Sensitivity analysis 2: all events were attributed to the randomized groups regardless of whether patients treated with placebo had switched to open-label ADA therapy</u> All-cause hospitalisation, 89/377.7 (0.24) UC-related hospitalisation, 68/384.2 (0.18) UC- or drug-related hospitalisation, 76/382.2 (0.20) Colectomy, 19/400.0 (0.05). Elective/emergency NR		
Suzuki	PBO	NR	NR	NR
Suzuki	ADA80/40mg	NR	NR	NR
Suzuki	ADA160/80mg	NR	NR	NR
Suzuki	ADA40mg EOW	NR	NR	NR
Suzuki	Rescue arm	NR	NR	NR
PURSUIT-SC	All randomised PBO	NR	NR	NR
PURSUIT-SC	All randomised GOL 200/100 mg	NR	NR	NR
PURSUIT-SC	Phase II PBO	NR	NR	NR
PURSUIT-SC	Phase II GOL 200/100 mg all randomised	NR	NR	NR
PURSUIT-SC	Phase III PBO	NR	NR	NR
PURSUIT-SC	Phase III GOL 200/100 mg phase III	NR	NR	NR

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
PURSUIT-Maintenance	PBO randomised	NR	NR	Deaths reported through week 54. PBO=0 Deaths reported after week 54. PBO SC induction and maintenance = 1 (pneumonia and heart failure)
PURSUIT-Maintenance	GOL 50 mg randomised	NR	From submission In the PURSUIT trial, only 2%-3% of golimumab induction responders re-randomised to golimumab 50mg or 100mg had a colectomy at the end of maintenance.	Deaths reported through week 54. GOL 50 mg= 0 Deaths reported after week 54. GOL SC 100/50 mg induction, 50 mg maintenance = 1 (heart dysfunction in the presence of pronounced athersclerosis and stenosis affecting aorta, large arteries and coronary arteries).

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
PURSUIT-M (Maintenance)	GOL 100 mg randomised	NR	NR	Deaths reported through week 54. GOL 100 mg= 3 (causes= malnutrition and sepsis (GOL 2 mg/kg IV induction), cardiac failure with history of thrombosis (GOL 400/200 mg SC induction), disseminated tuberculosis in patient who tested positive for latent TB on innduction study entry and was receiving isoniazid at time of event (GOL 200/100 mg SC induction) Deaths reported after week 54. PBO SC induction, GOL 100 mg maintenance = 1 (myocardial infarction in patient with history of myocardial infarction). GOL 2 mg/kg IV induction, GOL 100 mg maintenance = 2 (gallbladder adenocarcinoma with liver metastasis), (sepsis). GOL 200/100 mg SC induction, GOL 100 mg maintenance = 1 (accidental nitrous oxide overdose).
ACT1	PBO	From submission[65] UC-related hospitalisation, mean (SD): 0.22 (0.57)	From submission[65] Colectomy n(%), 9 (7.4) Ostomy n (%), 5 (4.1)	NR

Study acronym	Treatment arm	Rates of hospitalisation	Rates of surgical intervention	Death
ACT1	IFX 5 mg/kg	From submission[65] UC-related hospitalisation, mean (SD): 0.11 (0.34); p-value (assume vs. PBO), 0.061	From submission[65] Colectomy n(%), 7 (5.8) Ostomy n (%), 3 (2.5)	1 during ACT2 extension
ACT 2	PBO	From submission[65] UC-related hospitalisation, mean (SD): 0.21 (0.55)	From submission[65] Colectomy n(%), 1 (0.7) Ostomy n (%), 1 (0.7)	NR
ACT 2	IFX 5 mg/kg i	From submission[65] UC-related hospitalisation, mean (SD): 0.07 (0.29); p-value (assume vs. PBO), 0.009	From submission[65] Colectomy n(%), 0 (0.0) Ostomy n (%), 0 (0.0)	NR
UC-SUCCESS	AZA	NR	NR	NR
UC-SUCCESS	IFX	NR	NR	NR
UC-SUCCESS	IFX/AZA	NR	NR	NR
Probert	PBO		One PBO patient underwent colectomy during the intervention period and was recorded as a treatment failure. One patient (unclear which group) refused sigmoidoscopic assessment but by other clinical measures was deemed to be a treatment failure	NR
Probert	IFX	NR	NR	NR

Hospitalisation, surgery and mortality data (paediatric population trial)

Study acronym (copy to each study group row for each study)	Treatment arm	Rates of hospitalisation	Rates of surgical intervention (both elective and emergency)	Death
Infliximab				
Hyams	IFX 5mg/q8w	NR	Patients requiring colectomy in the 54-week period: 1/22 (4.5%)	NR
Hyams	IFX 5mg/q12w	NR	Patients requiring colectomy in the 54-week period: 2/23 (8.7%)	NR
Hyams	All patients to wk8 (n=60)	NR	Patients requiring colectomy in the 54-week period: 5/60 (8%) (2 of 15 nonrandomised)	NR

Appendix 6: Safety data tables

Safety - participants experiencing adverse events, serious adverse events and withdrawal due to adverse events (adult population trials)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
ULTRA1	PBO	NR	Week 8	[45] Any severe (not defined) 17/223 (7.6%); any serious (not defined, 17/223 (7.6%) Submission[63] 17/223 (7.6%); drug related SAE, 4/223 (1.8%)	[45] 108/223 (48.4%) Submission[63] 108/223 (48.1%); possibly drug related, 48/223 (21.5%)	[45] 12/223 (5.4%) Submission[63] 12/223 (5.4%)
ULTRA1	ADA 160/80mg	NR	Week 8	[45] Any severe (not defined) 19/223 (8.5%); any serious (not defined, 9/223 (4.0%)	[45] 112/223 (50.2%)	[45] 12/223 (5.4%)
ULTRA1	ADA 160/80mg	NR	Week 8	Submission[63] 97/223 (4.0%); drug related SAE, 1/223 (0.4%)	Submission[63] 112/223 (50.2%); possibly drug related, 43/223 (19.3%)	Submission[63] 12/223 (5.4%)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
ULTRA2	PBO	NR	Week 52	[46] Any severe (not defined) 37/260 (14.2%); any serious (not defined, 32/260 (12.3%) Submission[63] 32/260 (12.3%)	[46] 218/260 (83.8%); possibly drug-related, 86/260 (33.1%) Submission[63] 218/260 (83.8%)	[46] 34/260 (13.1%) Submission[63] 34/260 (13.1%)
ULTRA2	ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW beginning at week 4	NR	Week 52	[46] Any severe (not defined) 41/257 (16.0%); any serious (not defined, 31/257 (12.1%) Submission[63] 31/257 (12.1%)	[46] 213/257 (82.9%); possibly drug-related, 101/257 (39.3%) Submission[63] 213/257 (82.9%)	[46] 23/257 (8.9%) Submission[63] 23/257 (8.9%)
ULTRA3	ADA 40 mg EOW or EW	NR	Week 52	[56] 76/577 (13.6%) Events, 93 (events per 100 PY, 21.8)	[56] 421/577 (75.6%) Events, 2187 (events per 100 PY, 512.3)	[56] 78/577 (14.0%) Events, 90 (events per 100 PY, 21.1)
ULTRA3	ADA 40 mg EOW or EW	NR	Week 52	Submission[63] ADA 40 mg EOW/EW n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): 414 (17.7)	Submission[63] ADA 40 mg EOW/EW n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): 8057 (344.6)	Submission[63] ADA 40 mg EOW/EW n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): 249 (10.7)
Suzuki	PBO	52 weeks	Week 8	[47]	[47]	[47]

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
				Week 8: 7/96 (7.3%)	Week 8: 45/96 (46.9%)	Week 8: 4/96 (4.2%)
Suzuki	PBO	52 weeks	Week 52	[47] Week 52 (n=96, patient years = 44.8): events, 273 (events/100 patient years, 609.4)	[47] Week 52 (n=96, patient years = 44.8): events, 14 (events/100 patient years, 31.3)	[47] Week 52 (n=96,patient years = 44.8): events, 6 (events/100 patient years, 13.4)
Suzuki	ADA160/80mg	52 weeks	Week 8	[47] Week 8: 4/90 (4.4%)	[47] Week 8: 40/90 (44.4%)	[47] Week 8: 6/90 (6.7%)
Suzuki	ADA80/40mg or ADA160/80 to week 8 then ADA40 EOW	52 weeks	Week 52	[47] Week 52 (n=177, patient years, 98.2): events, 33 (events/100 patient years, 33.6) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 20 (events/100 patient years, 29.1)	[47] Week 52 (n=177, patient years, 98.2): events, 538 (events/100 patient years, 547.9) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 343 (events/100 patient years, 499.3)	[47] Week 52 (n=177, patient years, 98.2): events, 22 (events/100 patient years, 22.4) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 11 (events/100 patient years, 16.0)
PURSUIT-SC	PBO	6.05 weeks Mean 1.98	Week 6	[48] Patients with ≥ SAE (not defined) 20/330 (6.1)	[48] 126/330 (38.2) Headache 17/330 (5.2) Nasophayngitis 11/330 (3.3) Pyrexia 7/330 (2.1) Nausea 7/330 (2.1) exacerbation of UC 13/330 (3.9)	[48] 3/330 (0.9) (viral infection= erythema nodosum, exacerbation of UC)
PURSUIT-SC	GOL 200/100 mg	6.08 weeks Mean 1.99	Week 6	[48] 9/331 (2.7)	[48] 124/331 (37.5) Headache 10/331 (3.0) Nasophayngitis 11/331 (3.3) Pyrexia 6/331 (1.8) Nausea 3/331 (0.9) exacerbation of UC 7/331 (2.1)	[48] 1/331 (0.3) (worsening of UC/clostridia infection)
PURSUIT-	PBO. N=156	32.7 weeks	Week 54	[49]	[49]	[49]

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
M		8.2 Total number of study agent injections. PBO, 3333		≥ SAE. PBO, 12 (7.7)	103 (66.0) (all are treatment-emergent Aes). Exacerbation of UC, 29 (18.6) nasopharyngitis, 11 (7.1) headache, 14 (9.0) arthralgia, 12 (7.7) abdominal pain= 4 (2.6) upper respiratory tract infection= 4 (2.6) rash, 3 (1.9) phayngitis, 4 (2.6) cough, 5 (3.2)	PBO, 10 (6.4)
PURSUIT-M	GOL 50 mg. N=154	44.3 11.1 Total number of study agent injections. GOL 50 mg, 4392	Week 54	[49] GOL 50 mg, 13 (8.4)	[49] 112 (72.7) Exacerbation of UC, 27 (17.5) nasopharyngitis, 14 (9.1) headache, 12 (7.8) arthralgia, 11 (7.1) abdominal pain= 11 (7.1) upper respiratory tract infection= 8 (5.2) rash, 9 (5.8) phayngitis, 8 (5.2) cough, 5 (3.2)	[49] GOL 50 mg, 8 (5.2)
PURSUIT-M	GOL 100 mg. N=154	46.3 11.3 Total number of study agent injections. GOL 100 mg, 4440	Week 54	[49] GOL 100 mg, 22 (14.3)	[49] 113 (73.4) Exacerbation of UC, 24 (15.6) nasopharyngitis, 21 (13.6) headache, 12 (7.8) arthralgia, 8 (5.2) abdominal pain= 11 (7.1) upper respiratory tract infection= 9 (5.8) rash, 7 (4.5) phayngitis, 5 (3.2) cough, 9 (5.8)	[49] GOL 100 mg, 14 (9.1)
UC-SUCCESS	AZA	Week 8 Week 8 to 16	Week 8 and week 8 to 16	[52] Week 8: 6/79 (8%) no definition Week 8 to 16: 0/42 (0%); AFX to IFX/AZA, 1/20 (5%)	[52] Week 8: 41/79 (52%) Week 8 to 16: 11/42 (26%); AFX to IFX/AZA, 7/20 (35%) Week 8: Abdominal pain 4/79 (5%); abdominal pain= upper 4/79 (5%); anaemia 4/79 (5%); fatigue 4/79 (5%);	[52] Week 8: 6/79 (8%) Week 8 to 16: 1/42 (2%); AFX to IFX/AZA, 3/20 (15%)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
					headache 8/79 (10%); nausea 10/79 (13%); pyrexia 3/79 (4%); vomiting 6/79 (8%) Week 8 to 16: Arthralgia 3/42 (7%); Aspergillosis, Chest discomfort, Conjunctival haemorrhage, Drug hypersensitivity, Dyspnea, Leukopenia, Nasopharyngitis, Painful defecation= Pyrexia, Ulcerative colitis - all 0/42 (0%); Pain in extremity 2/42 (5%) Week 8 to 16: AFX to IFX/AZA: Arthralgia, 0/20 (0%); Aspergillosis 1/20 (5); Chest discomfort 1/20 (5); Conjunctival hemorrhage 1/20 (5); Drug hypersensitivity, 1/20 (5); Dyspnea, 1/20 (5); Leukopenia, 1/20 (5); Nasopharyngitis 1/20 (5); Painful defecation= 1/20 (5); Pain in extremity 0 Pyrexia, 1/20 (5); Ulcerative colitis, 1/20 (5)	
UC-SUCCESS	IFX	NR	Week 8 and week 8 to 16	[52] Week 8: 2/78 (3%) Week 8 to 16: 4/74 5	[52] Week 8: 26/78 (33%) Week 8 to 16: 22/30 (29%) Week 8: Abdominal pain 3/78 (4%); abdominal pain= upper 0/78 (0%); anaemia 3/78 (4%); fatigue 0/78 (0%); headache 4/78 (5%); nausea 1/78 (1%); pyrexia 5/78 (6%); vomiting 0/78 (0%) Week 8 to 16: Arthralgia, 2/74 (3%); aspergillosis, chest discomfort, conjunctival haemorrhage, drug	[52] Week 8: 2/79 (3%) Week 8 to 16: 3/74 (4%)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
					hypersensitivity, dyspnea, leukopenia, painful defecation= pain in extremity – all 0/74 (0%); nasopharyngitis 1/74 (1%); pyrexia, 2/74 (3%); ulcerative colitis, 4/74 (5%)	
UC-SUCCESS	IFX/AZA	NR	Week 8 and week 8 to 16	[52] Week 8: 3/80 (4%) Week 8 to 16: 1/72 (1%)	[52] Week 8: 30/80 (38%) Week 8 to 16: 21/72 (29%) Week 8: Abdominal pain 0/80 (0%); abdominal pain= upper 0/80 (0%); anaemia 1/80 (1%); fatigue 1/80 (1%); headache 4/80 (5%); nausea 7/80 (9%); pyrexia 2/80 (3%); vomiting 1/80 (1%) Week 8 to 16: Arthralgia, 2 (3); aspergillosis, chest discomfort, conjunctival haemorrhage, drug hypersensitivity, dyspnea, leukopenia, painful defecation= pain in extremity, pyrexia – all 0/72 (0%); nasopharyngitis, 1/72 (1%): ulcerative colitis, 3/72 (4%)	[52] Week 8: 3/79 (4%) Week 8 to 16: 3/72 (4%)
Probert	PBO	Week 6	Week 6		2/20 (10%); 1 septic complications, 1 colectomy due to toxic exacerbation and spontaneous perforation	
ACT1	PBO	36.2 weeks (all mean= SD NR) 24.2 weeks treatment (Mean= SD NR for all)	Week 54	[50] 31/121 (25.6) States SAEs most commonly related to gastrointestinal system in both studies (no further details)	[50] 103/121 (85.1) AEs occurring in ≥ 10% of any treatment group only reported. Worsening UC, 40/121 (33.1) Abdominal pain= 16/121 (13.2 Nausea, 14/121 (11.6) Upper RTI, 28/121 (23.1)	[50] 11/121 (9.1)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
					Pharyngitis, 10/121 (8.3) Sinusitis, 4/121 (3.3) Pain= 19/121 (15.7) Rash, 16/121 (13.2) Arthralgia, 18/121 (14.9) Headache, 27/121 (22.3) Fever, 10/121 (8.3) Anaemia, 12/121 (9.9) Fatigue, 11/121 (9.1)	
ACT1	PBO	Mean duration of treatment, 23 weeks (no SD reported) mean duration of follow-up, 32 weeks (no SD reported) NR	Week 54	[68] 57/244 (23) patients with long-term follow-up (mean 30 weeks) patients with 1 or more SAEs (%) 6/14 (43) UC 6/14 (43) fever 1/14 (7)	[68] Any AE (%) 196/244 (80) AEs occurring in > 10% of any treatment group: worsening UC 61/244 (25) abdominal pain 31/244 (13) nausea 23/244 (9) upper RTI 43/244 (18) pharyngitis 16/244 (7) sinusitis 12/244 (5) pain 30/244 (12) fatigue 19/244 (8) arthralgia 26/244 (11) fever 22/244 (9) headache 45/244 (18) anaemia 25/244 (10)	[68] 23/244 (9)
ACT1	IFX 5 mg/kg	44.9 weeks 34.8 weeks	Week 54	[50] 26/121 (21.5)	[50] 106/121 (87.6) Worsening UC, 23/121 (19.0) Abdominal pain=11/121 (9.1)) Nausea, 14/121 (11.6) Upper RTI, 20/121 (16.5) Pharyngitis, 12/121 (9.9) Sinusitis, 8/121 (6.6) Pain=14/121 (11.6) Rash, 14/121 (11.6) Arthralgia, 21/121 (17.4)	[50] 10/121 (8.3)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
					Headache, 22/121 (18.2) Fever, 14/121 (11.6) Anaemia, 4/121 (3.3) Fatigue, 14/121 (11.6)	
ACT1	IFX 5 mg/kg	Mean duration of treatment, 33 weeks (no SD reported) mean duration of follow-up, 41 weeks (no SD reported) NR	Week 54	[68] 43/242 (18) patients with long-term follow-up (mean 25 weeks) patients with 1 or more SAEs (%) (Aes included, IFX-related AEs or those requiring hospitalisation for treatment of UC (including colectomy) 5/15 (33) UC 5/15 (33) fever 0	[68] Any AE (%) 208/242 (86) AEs occurring in > 10% of any treatment group: worsening UC 36/242 (15) abdominal pain 22/242 (9) nausea 21 (242 (9) upper RTI 39/16 (16) pharyngitis 23/242 (10) sinusitis 20/242 (8) pain 25/242 (10) fatigue 21/242 (9) arthralgia 40/242 (17) fever 27/242 (11) headache 44/242 (18) anaemia 11/242 (5)	[68] 14/242 (6)
ACT2	PBO	21.9 weeks (Mean for all, SD NR) 14.4 weeks Duration of treatment (Mean for all, SD NR)	Week 54	[50]	[50] AEs selected as for ACT1: PBO 90/123 (73.2) Worsening UC, 20/123 (16.3) Abdominal pain= 14/123 (11.4) Nausea, 9/123 (7.3) Upper RTI, 14/123 (11.4) Pharyngitis, 3/123 (2.4) Sinusitis, 7/123 (5.7) Pain= 11/123 (8.9) Rash, 3/123 (2.4) Arthralgia, 6/123 (4.9) Headache, 18/123 (14.6) Fever, 12/123 (9.8) Anaemia, 13/123 (10.6) Fatigue, 6/123 (4.9)	[50] 12/123 (9.8)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)	Discontinuation due to adverse event(s) nN (%)
ACT2	IFX 5 mg/kg	27.5 weeks 99/121 (81.8)	Week 54	[50]	[50] IFX 5 mg/kg 99/121 (81.8) Worsening UC, 11/121 (9.1) Abdominal pain= IFX5 mg/kg 10/121 (8.3) Nausea, 6/121 (5.0) Upper RTI, 16/121 (13.2) Pharyngitis, 7/121 (5.8) Sinusitis, 11/121 (9.1) Pain= 9/121 (7.4) Rash, 2/121 (1.7) Arthralgia, 16/121 (13.2) Headache, 19/121 (15.7) Fever, 13/121 (10.7) Anaemia, 6/121 (5.0) Fatigue, 6/121 (5.0)	[50] 2/121 (1.7)
ACT1, ACT2 extension studies	IFX combined group N=230	Mean (SD) duration of follow-up of 113 (642) weeks (range, 4–184 weeks; median= 128 weeks; 25–75 interquartile [IQ] range, 96–144 weeks). Mean 2.16 years (no SD) Treatment duration mean (no SD) 1.99 years		[55] 49/230 (21.3%) experienced SAE. Number serious adverse events, 21 per 100 patient-years. SAE experienced by more than 1 patient: UC falre n=11 (4.8%) pneumonia n=5 (2.2%) gastrointestinal bleeding n=4 (1.7%) nausea n=3 (1.3%) bone fracture n=3 (1.3%) abdominal pain n=2 (0.9%) intestinal obstruction n=2 (0.9%) fever n=2 (0.9%).	[55] Number adverse events, 506 per 100 patient-years	[55] 4.63 per 100 patient-years

Safety - participants experiencing adverse events, serious adverse events and withdrawal due to adverse events (paediatric population trial)

Trial name	Treatment arm	Length of safety follow up / mean number of administrations	Time point	Discontinuation due to adverse event(s) nN (%)	Number of patients experiencing 1 or more adverse event, nN= (%)	Number of patients experiencing 1 or more serious adverse event, nN (%) (including definition)
Hyams	IFX 5mg/q8w	Mean weeks, 50.4 Mean exposure weeks, 41.0 Total infusions, 165	Week 8	[53] 22/22 (100%) 22/22 (100%)	[53] 4/22 (18.2%) (1, serious infection; 1, pancreatitis and UC flare during induction plus viral infection after step-up; 1, UC flare after step-up; 1, anaemia during maintenance)	[53] 3/22 (13.6%)
Hyams	IFX 5mg/q12w	Mean weeks, 44.6 Mean exposure weeks, 34.3 Total infusions, 135	Week 8	[53] 23/23 (100%) 23/23 (100%)	[53] 5/23 (21.7%) (1, pharyngitis during induction; 1, urinary tract infection during induction; 1, UC flare during induction [x1] and UC flare during maintenance [x1]; 1, UC flares during induction [x2]; 1, UC flare after step-up [x1])	[53] 6/23 (26.1%)
Hyams	All patienst to wk 8 (n=60)	Mean weeks, 38.0	Week 8	[53] 57/60 95%	[53] 14/60 23.3%	[53] 13/60 (21.7%)

Safety – infections, serious infections, infections requiring treatment, reactivation of TB or hepatitis, injection site reactions, infusion reactions, serious allergic reactions (adult population trials)

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ULTRA1	PBO	Week 8	[45] 35/223 (15.7%)	NR	[45] 3/223 (1.3%) (pneumonia, 1; sepsis, 1; wound infection staphylococcal, 1)	NR	NR	[45] 7/223 (3.1%)	NR	NR
ULTRA1	ADA 80/40mg	Week 8	[45] 26 /130 (20.0%)	NR	[45] 2/130 (1.5%) (abscess rupture, 1; perirectal abscess, 1)	NR	NR	[45] 7/130 (5.4%)	NR	NR
ULTRA1	ADA 160/80mg	Week 8	[45] 32/223 (14.3%); oportunist infection (oesophogeal candidiasis) 1/223 (0.4%)	NR	[45] 0/223 (0%)	NR	NR	[45] 13/223 (5.8%)	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ULTRA2	PBO	Week 52	[46] 103/260 (39.6%); opportunistic infection-related AE (excluding TB) 3/260 (1.2%)	NR	[46]5/260 (1.9%)	NR	NR	[46] 10/260 (3.8%)	NR	NR
ULTRA2	ADA 160/80mg	Week 52	[46]116/257 (45.1%); opportunistic infection-related AE (excluding TB) 5/257 (1.9%)	NR	[46]4/257 (1.6%)	NR	NR	[46]31/257 (12.1%)	NR	NR
ULTRA3	ADA 80/40mg		[56]213/577 (38.2%) Events, 382 (events per 100 PY, 89.5) Opportunistic infection: 5/577 (0.9%) Events, 6	NR	[56]17/577 (3.1%) Events, 17 (events per 100 PY, 4.0)	NR	NR	[56] 8/577 (1.4%) Events, 8 (events per 100 PY, 1.9)	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			(events per 100 PY, 1.4)							
ULTRA3	ADA 160/80	Week 52	NR	NR	Submission[63] ADA 40 mg EOW/EW n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): serious infection= 79 (3.4); opportunistic infection excluding TB, 6 (0.3)	ADA 40 mg EOW/EW n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): 1 (<0.1)	NR	Submission[63] ADA 40 mg EOW/EW n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): 246 (10.5)	NR	NR
Suzuki	PBO	Week 8	[47]Week 8: 15/96 (15.6%) Opportunistic infection (excluding tuberculosis): Week 8: 0/96 (0%)	NR	[47]Week 8: 0/96 (0%)	[47]Week 8: 0/96 (0%)	NR	[47]Week 8: 2/96 (2.1%)	NR	NR
Suzuki	PBO	Week 52	[47]Week 52	NR	[47]Week 52	[47]Week 52	NR	[47]Week 52	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			(n=96, patient years, 44.8): events, 70 (events/100 patient years, 156.3) Opportunistic infection (excluding tuberculosis): Week 52 (n=96, patient years, 44.8): events, 0 (events/100 patient years, 0)		(n=96, patient years, 44.8): events, 2 (events/100 patient years, 4.5)	(n=96, patient years, 44.8): events, 0 (events/100 patient years, 0)		(n=96, patient years, 44.8): events, 4 (events/100 patient years, 8.9)		
Suzuki	ADA80/40	Week 8	Week 8: 11/87 (12.6%) Opportunistic infection (excluding tuberculosis): Week 8:	NR	Week 8: 0/87 (0%)[47]	Week 8: 0/87 (0%)[47]	NR	Week 8: 5/87 (5.7%)[47]	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			0/87 (0%)[47]							
Suzuki	ADA160/80	Week 8	Week 8: 17/90 (18.9%) Opportunistic infection (excluding tuberculosis): Week 8: 1/90 (1.1%)[47]	NR	Week 8: 3/90 (3.3%)[47]	Week 8: 1/90 (1.1%)[47]	NR	Week 8: 7/90 (7.8%)[47]	NR	NR
Suzuki	ADA40 EOW	Week 52	Week 52 (n=177, patient years, 98.2): events, 134 (events/100 patient years, 136.5) ADA week 8	[47]	Week 52 (n=177, patient years, 98.2): events, 8 (events/100 patient years, 8.1) ADA week 8 responders per	Week 52 (n=177, patient years, 98.2): events, 1 (events/100 patient years, 1.0) ADA week 8 responders per		Week 52 (n=177, patient years, 98.2): events, 20 (events/100 patient years, 20.4) ADA week 8 responders per		NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			responders per full Mayo score (n=82, patient years, 68.7): 90 (events/100 patient years, 131.0) Opportunistic infection (excluding tuberculosis): Week 52 (n=177, patient years, 98.2): events, 2 (events/100 patient years, 2.0) ADA week 8 responders per full Mayo score (n=82, patient		full Mayo score (n=82, patient years, 68.7): 6 (events/100 patient years, 8.7) [47]	full Mayo score (n=82, patient years, 68.7): 0 (events/100 patient years, 0) [47]		full Mayo score (n=82, patient years, 68.7): 9 (events/100 patient years, 13.1) [47]		

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			years, 68.7): 2 (events/100 patient years, 2.9) [47]							
PURSUIT-SC	PBO	Week 6	Patients with ≥ 1 infection 40/330 (12.1) 1 opportunistic infection (cytomegalovirus infection) (not reported as serious) [48]	Patients with ≥ 1 infection requiring treatment 23/330 (7.0) [48]	Patients with ≥ 1 serious infection 6/330 (1.8) (1 pneumonia) [48]	NR	NR	Patients with ≥ 1 injection-site reaction 5/330 (1.5) [48]	NR	NR
PURSUIT-SC	GOL 100/50 mg		8/71 (11.3) [48]	[48]0	[48]0	NR	NR	[48]4/71 (5.6)	NR	NR
PURSUIT-SC	GOL 200/100 mg	Week 6	39/331 (11.8)[48]	15/331 (4.5)[48]	1/331 (0.3) (1 pneumonia)[48]	NR	NR	11/331 (3.3)[48]	NR	NR
PURSUIT-SC	GOL 400/200 mg	Week 6	41/332 (12.3) 1 opportunistic infection (oesophageal candidiasis) (not reported as serious)[48]	25/332 (7.5)[48]	3/332 (0.9)[48]	NR	NR	10/332 (3.0)[48]	NR	NR
PURSUI	PBO.	Week 54	. ≥ 1	24 (15.4)[49]	3 (1.9)	TB reported	NR	Injections with	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
T-M	N=156		infection (as assessed by investigator) . 44 (28.2)[49]		Serious opportunistic infection. GOL SC 200/100 mg induction= PBO maintenance, 1 (cytomegalovirus infection apprxo 3 months after last GOL dose).[49]	through week 54. GOL 4 mg/kg IV induction= PBO maintenance, 1[49]		injection-site reactions. 18 (0.5) ≥ 1 injection-site reactions. 3 (1.9)[49]		
PURSUI T-M	GOL 50 mg. N=154	Week 54	60 (39.0)[49]	39 (25.3)[49]	5 (3.2) Serious opportunistic infection. GOL 50 mg maintenance, 0[49]	0	NR	Injections with injection-site reactions. 18 (0.4) ≥ 1 injection-site reactions. 3 (1.9)[49]	NR	NR
PURSUI T-M	GOL 100 mg. N=154	Week 54	60 (39.0)[49]	44 (28.6)[49]	5 (3.2) Serious opportunistic infection. GOL 200/100 mg induction= GOL 100 mg maintenance, 1	3 (1.9) GOL 100 mg maintenance (1 each GOL 400/200 mg SC, 4 mg/kg IV and 200/100 mg	NR	Injections with injection-site reactions. 28 (0.6) ≥ 1 injection-site reactions. 11 (7.1)[49]	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
					(Staphylococcus aureus and Nocardia cultured from a brain abscess).[49]	SC induction) (inlc fatal case reported previously).[49]				
UC-SUCCESS	AZA	Week 8	NR	NR	1/79 (1%)[52]	NR	NR	NR	1/79 (1%)[52]	NR
UC-SUCCESS	IFX	Week 8	NR	NR	1/78 (1%)[52]	NR	NR	NR	0/78 (0%)[52]	NR
UC-SUCCESS	IFX/AZA	Week 8	NR	NR	0/80 (0%)[52]	NR	NR	NR	0/80 (0%)[52]	NR
ACT 1	PBO	Week 54	47/121 (38.8) Fungal dermatits, 8/121 (6.6) Pneumonia, 0 Varicella-zoster virus infection= 1/121 (0.8) Herpes zoster, 0[50]	25/121 (20.7)[50]	Serious infections, 5/121 (4.1) Bacterial infection= 1/121 (0.8) Upper RTI, 1/121 (0.8) Pneumonia, 0 Tuberculosis, 0 Abscess, 1/121 (0.8) Pharyngitis,	NR	NR	NR	Acute infusion reaction (any AE occurring ≤ 2 hr after start of infusion) 13/121 (10.7)	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
					1/121 (0.8) Gastroenteritis, 0 Earache, 0 Fever, 0 Vaginitis, 0 Appendicitis, 0 Colitis, 0 Surgical wound infection= 1/121 (0.8) Pancreatitis, 0 Pleurisy, 0 Sinusitis, 1/121 (0.8)[50]					
ACT 1	PBO	ACT-1 and -2, and ACT-2 extension through 54 weeks	80/244 (33)[68]	NR	Serious infections (%) 6/244 (2) bacterial infection 1/244 (0.4) upper RTI 1/244 (0.4) pneumonia 0, tuberculosis 0, abscess 2/244 (1) pharyngitis 1/244 (0.4)	NR	NR	NR	Possible delayed hypersensitivity reactions (%) 2/242 (1)	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
					gastroenteritis 0, earache 0, fever 0, vaginitis 0, appendicitis 0, colitis 0, infection 1/244 (0.4)(no further details) pancreatitis 0, pericarditis 0, pleurisy 0, pyelonephritis 0, sinusitis 1/244 (0.4)[68]					

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT 1	IFX 5 mg/kg	Week 54	53/121 (43.8) Fungal dermatits,1/121 (0.8) [50] Pneumonia, 2/121 (1.7) Varicella-zoster virus infection= 1/121 (0.8) Herpes zoster, 1/121 (0.8) n=121 All infections baseline IMM: 32／66 (48.5%) All infections no baseline IMM: 21/55 (38.2%)[343]	39/121 (32.2) [50]	Serious infections, 3/121 (2.5) Bacterial infection= 0 Upper RTI, 0 Pneumonia, 0 Tuberculosis, 0 Abscess, 0 Pharyngitis, 0 Gastroenteritis, 1/121 (0.8) Earache, 0 Fever, 0 Vaginitis, 0 Appendicitis, 1/121 (0.8) Colitis, 0 Surgical wound infection= 0 (0.8) Pancreatitis, 1/121 (0.8) Pleurisy, 0 Sinusitis, 0 [50] [50] n=121 Serious infections baseline IMM: 3/66 (4.5%) Serious infections no baseline IMM: 0/55 (0.0%)	NR	NR	NR	/121 (9.9) Possible delayed hypersensitivity reactions 2/121 (1.7) [50] n=121: Infusions n with reactions baseline IMM: 7/423 (1.7%) Infusions n with reactions no baseline IMM: 8/364 (2.2%) Patients n with any infusion reactions baseline IMM: 6/66 (9.1%) Patients n with any infusion reactions no baseline IMM: 6/55 (10.9%) n=121: Infusions n with serious	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT 1	IFX 5 mg/kg	ACT-1 and -2, and ACT-2 extension through 54 weeks	94/242 (39)[68]	[68]	Serious infections (%) 7/242 (3) bacterial infection 0, upper RTI 0, pneumonia 2/242 (1) tuberculosis 0, abscess 0, pharyngitis 0, gastroenteritis 2/242 (1) earache 1/242 (0.4) fever 1/242 (0.4) vaginitis 0, appendicitis 1/242 (0.4) colitis 0, infection 0, pancreatitis 1/242 (0.4) pericarditis 0, pleurisy 0, pyelonephritis 0, sinusitis [68]	NR	NR	NR	Possible delayed hypersensitivity reactions (%) 2/242 (1)[343]	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT 1	IFX 10 mg/kg	Week 54	60/122 (49.2) Fungal dermatits, 3/122 (2.5) Pneumonia, 4/122 (3.3) Varicella-zoster virus infection, IFX 10 mg 0 Herpes zoster, 0 [50] n=122 All infections baseline IMM: 32/59 (54.2%) All infections no baseline IMM: 28/63 (44.4%)[343]	43/122 (35.2) [50]	Serious infections, 8/122 (6.6) Bacterial infection= 0 Upper RTI, 0 Pneumonia, 3/122 (2.5) Tuberculosis, 1 (0.8) Abscess, 2/122 (1.6) Pharyngitis, 1/122 (0.8) Gastroenteritis, 1/122 (0.8) Earache, 0 Fever, 1/122 (0.8) Vaginitis, 0 Appendicitis, 0 Colitis, 1/122 (0.8) Surgical wound infection= 1/122 (0.8) Pancreatitis, 0 Pleurisy, 1/12[50]2 (0.8) Sinusitis, 0 RM#1103 n=122 Serious infections baseline IMM: 6/59 (10.2%) Serious 343	NR	NR	NR	15/122 (12.3) Possible delayed hypersensitivity reactions 2/121 (1.7) [50] n=122: Infusions n with reactions no baseline IMM: 8/403 (2.0%) Infusions n with reactions baseline IMM: 8/367 (2.2%) Patients n with any infusion reactions baseline IMM: 8/59 (13.6%)[343] Patients n with any infusion reactions no baseline IMM: 7/63 (11.1%) n=122: Infusions n with serious	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT 1	IFX comb	Week 54	n=243 All infections baseline IMM: 64/125 (51.2%) All infections no baseline IMM: 49/118 (41.5%) ACT 2 Week 30 n=241 All infections baseline IMM: 30/102 (29.4%) All infections no baseline IMM: 37/139 (26.6%)[343]	[343]	n=243 Serious infections baseline IMM: 9/125 (7.2%) Serious infections no baseline IMM: 2/118 (1.7%) ACT 2 Week 30 n=241 Serious infusion reactions baseline IMM: 0/102 (0.0%) Serious infusion reactions no baseline IMM: 0/139 (0.0%)[343]	NR	NR	NR	n=243: Infusions n with reactions baseline IMM: 15/790 (1.9%) Infusions n with reactions no baseline IMM: 16/767 (2.1%) Patients n any infusion reactions baseline IMM: 14/125 (11.2%) Patients n any infusion reactions no baseline IMM: 13/118 (11.0%)	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
									n=243: Infusions n with serious reactions baseline IMM: 0/790 (0.0%) Infusions n with serious reactions no baseline IMM: 0/767 (0.0%) Serious infusion reactions baseline IMM: 0/125 (0.0%) Serious infusion reactions no baseline IMM: 0/118 (0.0%)[343]	
ACT2	PBO	Week 30	29/123 (23.6) Fungal	15/123 (12.2) [50]	1/123 (0.8) Bacterial infection= 0	NR	NR	NR	Defined as for ACT1: 10/123 (8.1)	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			dermatitis, 0 IFX Pneumonia, 0 IFX Varicella-zoster virus infection= 0 Herpes zoster, 1/123 (0.8) IFX 5 [50]		Upper RTI, 0 Pneumonia, 0 Tuberculosis, 0 Abscess, 1/123 (0.8) Pharyngitis, 0 Gastroenteritis, PBO 0 Earache, 0 Fever, 0 Vaginitis, 0 Appendicitis, 0 Colitis, 0 Surgical wound infection= 0 Pancreatits, 0 Pleurisy, 0 Sinusitis, 0[50]				Possible delayed hypersensitivity 0[50]	

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT2	IFX 5 mg/kg	Week 30	18/121 (14.9) Fungal dermatitis, 0 Pneumonia, 0 Varicella-zoster virus infection= 1/121 (0.8) Herpes zoster, 2/121 (0.8) [50] n=121 All infections baseline IMM: 17/52 (32.7%) All infections no baseline IMM: 16/69 (23.2%)[343]		2/121 (1.7) Bacterial infection= 0 Upper RTI, 0 Pneumonia, 0 Tuberculosis, 0 Abscess, 0 Pharyngitis, 0 Gastroenteritis, 1/121 (0.8) Earache, 1/121 (0.8) Fever, 1/121(0.8) Vaginitis, 0 Appendicitis, 0 Colitis, 0 Surgical wound infection= 0 Pancreatits, 0 Pleurisy, 0 Sinusitis, 0 [50] n=121 Serious infections [343]baseline IMM: 1/52 (1.9%) Serious infections no baseline IMM: 1/69 (1.4%)				14/121 (11.6) Possible delayed hypersensitivity 0[50] n=121: Infusions n with reactions baseline IMM: 4/242 (1.7%) Infusions n with reactions no baseline IMM: 12/316 (3.8%) Patients n with any infusion reactions baseline IMM: 3/52 (5.8%) Patients n with any infusion reactions no baseline IMM: 11/69 (15.9%) n=121: Infusions n with serious	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT2	IFX 10 mg/kg	Week 30	17/120 (14.2) Fungal dermatitis, 10 mg/kg 1 (0.8) Pneumonia, 10 mg/kg 2/120 (1.7) Varicella-zoster virus infection= 0 Herpes zoster, 1/120 (0.8) [50]	[50]	3/120 (2.5) Bacterial infection= 0 Upper RTI, 0 Pneumonia, 0 Tuberculosis, 0 Abscess, 1/120 (0.8) Pharyngitis, 0 Gastroenteritis, 0 Earache, 0 Fever, 0 Vaginitis, 1/120 (0.8) Appendicitis, 0 Colitis, 0 Surgical wound infection, 1/120 (0.8) Pancreatits, 0 Pleurisy, 0 Sinusitis, 0[50]	NR	NR	[50]	14/120 (11.7) Possible delayed hypersensitivity 1/120 (0.8)[343]	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT2	All treated patients, safety at week 30, N=121	Week 30	IFX 5 mg/kg WITH immunomodulators, 17/52 (32.7) IFX 5 mg/kg WITHOUT immunomodulators, 16/69 (23.2)[343] Rn=120 All infections baseline IMM: 13/50 (26.0%) All infections no baseline IMM: 21/70 (30.0%)[343]	NR	IFX 5 mg/kg WITH immunomodulators, 1/52 (1.9)[343] n=120 Serious infections baseline IMM: 1/50 (2.0%) Serious infections no baseline IMM: 2/70 (2.9%)[343]	NR	NR	NR	IFX 5 mg/kg WITH immunomodulators, 3/52 (5.8) IFX 5 mg/kg WITHOUT immunomodulators, 11/69 (15.9) Serious infusion reactions. IFX 5 mg/kg WITH immunomodulators, 0/52 (0) Serious infusion reactions. IFX 5 mg/kg WITHOUT immunomodulators, 0/69 (0)[343] n=120: [343] Infusions n with reactions baseline IMM: 6/224 (2.7%) Infusions n with reactions no baseline IMM:	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
ACT2	IFX comb	Week 30	ACT 1 Week 54 n=243 All infections baseline IMM: 64/125 (51.2%) All infections no baseline IMM: 49/118 (41.5%) ACT 2 Week 30 n=241 All infections baseline IMM: 30/102 (29.4%) All infections no baseline IMM:	NR	ACT 1 Week 54 n=243 Serious infections baseline IMM: 9/125 (7.2%) Serious infections no baseline IMM: 2/118 (1.7%) ACT 2 Week 30 n=241 Serious infusion reactions baseline IMM: 0/102 (0.0%) Serious infusion reactions no baseline IMM: 0/139 (0.0%)[343]	NR	NR	NR	n=241: Infusions n with reactions baseline IMM: 10/466 (2.2%) Infusions n with reactions no baseline IMM: 27/634 (4.3%) Patients n any infusion reactions baseline IMM: 8/102 (7.82%) Patients n any infusion reactions no baseline IMM: 20/139 (14.4%)	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
			37/139 (26.6%)[343]						n=241: Infusions n with reactions no baseline IMM: 27/634 (4.3%) Infusions n with serious reactions baseline IMM: 0/466 (0.0%) Serious infusion reactions baseline IMM: 0/102 (0.0%) Serious infusion reactions no baseline IMM: 0/139 (0.0%)[343]	
ACT1, ACT2 extension	IFX combined group	NR	Number infections, 99 per 100	Number infections requiring	patients (4.3%) had serious infection.	NR	NR	NR	NR	NR

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
studies	N=230		patient-years[55]	antimicrobial treatment, 41 per 100 patient-years[55]	Number serious infections 3.4 per 100 patient-years. During extension studies, no reports of TB or other opportunistic infections.[55]					

Safety – infections, serious infections, infections requiring treatment, reactivation of TB or hepatitis, injection site reactions, infusion reactions, serious allergic reactions (paediatric population trial)

Trial name	Treatment arm	Time point	Infections	Infections requiring treatment	Serious infections	Reactivation of TB	Reactivation of hepatitis B	Injection site reactions (relevant to ADA and GOL)	Infusion reactions (relevant to IFX)	Serious allergic reactions (e.g. anaphylaxis)
Hyams	IFX 5mg/q8w	Week 8	[53] 13/22 (59.1%)	NR	NR	NR	NR	NR	[53] 4/22 (18.2%)	NR
Hyams	IFX 5mg/q12w	Week 8	[53] 14/23 (60.9%)	NR	NR	NR	NR	NR	[53] 3/23 (13.0%)	NR
Hyams	All patients to wk 8 (n=60)	Week 8	[53] 31/60 (51.7%)	NR	NR	NR	NR	NR	[53] 8/60 (13.3%)	NR

Safety – heart failure, malignancies, hepatobiliary events, autoimmune processes, neurological events, haematological reactions (adult population trials)

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
ULTRA1	PBO	Week 8	[45] 0/223 (0%)	[45] 2/223 (0.9%) - basal cell carcinoma, 1; breast cancer, 1	NR	[45] Demyelinating disease 0/223 (0%)	Demyelinating disease, 0/223 (0%)	NR	Submission MedDRA System Organ Class and Preferred Term: any AE, 218/223 (83.8%); Colitis ulcerative, 21/223 (9.4%)
ULTRA1	ADA 160/80mg	Week 8	0/130 (0%)	[45] 0/130 (0%)	NR	[45] Lupus-like syndrome, 0/130 (0%)	[45] Demyelinating disease, 0/130 (0%)	NR	Submission MedDRA System Organ Class and Preferred Term: any AE, 213/223 (82.9%); Colitis ulcerative, 13/223 (5.8%)
ULTRA2	PBO	Week 52	[46] 0/260 (0%)	[46] 0/260 (0%)	NR	[46] Lupus-like syndrome, 0/260 (0%)	[46] Demyelinating disease, 0/260 (0%)	[46] 0/260 (0%)	NR

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
ULTRA2	ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW beginning at week 4	Week 52	[46]1/257 (0.4%)	[46] Malignancies: 2/257 (0.8%)	NR	[46] Lupus-like syndrome, 1/257 (0.4%)	[46] Demyelinating disease, 0/257 (0%)	[46] 5/257 (1.9%)	Submission MedDRA System Organ Class and Preferred Term: any AE, 213/257 (82.9%); Anaemia, 10/257 (3.9%); Iron deficiency anaemia, 7/257 (2.7%); Colitis ulcerative, 58/257 (22.6%); Abdominal pain= 20/257 (7.8%); Nausea, 15/257 (5.8%); Fatigue, 16/257 (6.2%); Pyrexia, 11/257 (4.3%); Gastroenteritis, 9/257 (3.5%); Nasopharyngitis, 45/257 (17.5%); Pharyngitis, 9/257 (3.5%); URTI, 11/257 (4.3%); Arthralgia, 20/257 (7.8%); Headache, 22/257 (8.6%); Oropharyngeal pain= 15/257 (5.8%)

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
ULTRA3	ADA 40 mg EOW or EW	Week 52	1/577 (0.2%) Events, 1 (events per 100 PY, 0.2)	3/577 (0.5%) Events, 3 (events per 100 PY, 0.7)	NR	Demyelinating disease: 1/557 (0.2%); 1 event (events per 100 PY, 0.2)	NR	11/577 (2.0%) Events, 13 (events per 100 PY, 3.0)	NR
ULTRA3	ADA 40 mg EOW or EW	Week 52	Submission n=1010; Patient Years, 2338 Events (Events / 100 Patient Years): 4 (0.2)	Submission Patient Years, 2338 Events (Events/ 100 Patient Years): excluding lymphoma, 23 (1.0); lymphoma, 3 (0.1)	Submission n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): 12 (0.5)	Submission n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): demyelinating disease, 3 (0.1)	NR	NR	Submission n=1010; Patient Years, 2338 Events (Events/ 100 Patient Years): UC worsening, 588 (25.2); flare, 588 (25.2)
Suzuki	PBO	Week 8	NR	[47] Week 8: 0/96 (0%)	[47] Week 8: 1/96 (1.0%)	NR	NR	[47] Week 8: 1/96 (1.0%)	[47] UC worsening/flare: Week 8: 8/96 (8.3%)
Suzuki	PBO	Week 52	NR	[47] Week 52 (n=96, patient years, 44.8): events, 0 (events/100 patient years, 0)	[47] Week 52 (n=96, patient years, 44.8): events, 3 (events/100 patient years,	NR	NR	[47] Week 52 (n=96, patient years, 44.8): events, 4 (events/100 patient years,	[47] Week 52 (n=96, patient years, 44.8): events, 15 (events/100 patient years, 33.5)

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
					6.7)			8.9)	
Suzuki	ADA 160/80mg only	Week 8	NR	[47] 1/90 (1.1%)	[47] 1/90 (1.1%)	NR	NR	[47] 1/90 (1.1%)	[47] UC worsening/flare: Week 8: 2/90 (2.2%)
Suzuki	ADA 80/40mg or ADA160/80 to week 8 then ADA40 EOW	Week 52	NR	Week 52 (n=177, patient years, 98.2): events, 2 (events/100 patient years, 2.0) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 1 (events/100 patient years, 1.5)	Week 52 (n=177, patient years, 98.2): events, 5 (events/100 patient years, 5.1) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 3 (events/100 patient years, 4.4)	NR	NR	Week 52 (n=177, patient years, 98.2): events, 6 (events/100 patient years, 6.1) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 4 (events/100 patient years, 5.8)	UC worsening/flare: Week 52 (n=177, patient years, 98.2): events, 18 (events/100 patient years, 18.3) ADA week 8 responders per full Mayo score (n=82, patient years, 68.7): 7 (events/100 patient years, 10.2)
PURSUIT-SC	Phase III PBO	Week 6	NR	NR	NR	NR	NR	NR	[48] Proportion of patients reporting Aes through week 6, 38.2
PURSUIT-SC	Phase III GOL 200/100 mg phase III	Week 6	NR	NR	NR	NR	NR	NR	[48] Proportion of patients reporting Aes through week 6, 37.5
PURSUIT-	PBO	Week	NR	[49]	NR	NR	NR	NR	NR

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
M		54		Neoplasm benign= malignant and unspecified. PBO, 1 (0.6) Breast cancer was reported in a patient who had received only placebo during induction and maintenance.					
PURSUIT-M	GOL 50mg	Week 54	NR	[49] GOL 50 mg, 4 (2.6)	NR	NR	NR	NR	NR
PURSUIT-M	GOL 100mg	Week 54	NR	[49] GOL 100 mg, 4 (2.6) Three malignancies were reported through week 54 in patients receiving golimumab 100 mg maintenance; 2 of these (rectal cancer and thyroid cancer) presented with symptoms while the patients were receiving SC placebo induction and 1 (lung adenocarcinoma) occurred in a patient with a 40-year smoking history who	NR	NR	NR	0	NR

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
				received golimumab 200/100 mg SC induction therapy					
UC-SUCCESS	AZA	Week 8	NR	NR	[52] 13/79 (16%)	NR	NR	NR	NR
UC-SUCCESS	IFX	Week 8	NR	NR	[52] 3/78 (4%)	NR	NR	NR	NR
UC-SUCCESS	IFX/AZA	Week 8	NR	NR	[52] 5/80 (6%)	NR	NR	NR	NR
ACT 1	PBO	Week 54	NR	[68] 1 (basal cell carcinoma) 1 colonic dysplasia through week 54 in RESULTS-UC	NR	NR	NR	NR	[68] AEs of particular interest (%): fungal dermatitis 8/244 (3) pneumonia 0, varicella zoster virus infection 1/244 (0.4) herpes zoster 1/244 (0.4)

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
ACT 1	IFX 5 mg/kg i.v.	Week 54	NR	[68] N=2. 1 patient with prostatic adenocarcinoma with 2 year history of elevated PSA concn. 1 patient with colonic dysplasia. 2 (prostate adenocarcinoma, rectail adenocarcinoma) through 54 weeks in RESULTS-UC, 1 new cancer (squamous cell skin carcinoma) developed in IFX 5 mg/kg group patient, plus 1 colonic dysplasia	NR	0	[68] 1 patient with optic neuritis	NR	[68] AEs of particular interest (%): fungal dermatitis, pneumonia, varicella zoster virus infection= herpes zoster
ACT 2	PBO	Week 30	NR	[50] N=1. Basal-cell carcinoma	NR	0	0	NR	[50] At week 30 proportion with positive tests for antibodies to IFX who had infusion

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
									reaction 50.0 (6/12)
ACT 2	IFX	Week 30	NR	[50] N=1. Rectal adenocarcinoma.	NR	[50] N=1. 1 patient with lupus-like reaction (considered SAE)	[50] N=1. 1 patient with optic neuritis	NR	NR
ACT 2	IFX	Week 30	NR	0	NR	0	[50] N=1. 1 patient with multifocal motor neuropathy	NR	NR
ACT1, ACT2 extension studies	IFX all		NR	[55] Malignancy 1.01 per 100 patient-years. 5 malignancies reported during extension studies for IFX-treated patients. 19 yr old patient with adenocarcinoma of lung diagnosed (receiving 5 mg/kg IFX) 1 month after E128 infusion. Patient was non-snoker and died approx 18 months after completing extension study. 1 patient each	NR	NR	[55] No cases of optic neuritis or multifocal motir neuropathy were reported during extension studies.	NR	NR

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
				developed breast cancer and prostate cancer, both receiving IFX 5 mg/kg. Breast cancer diagnosied after week E72 infusion in 33 yr old patient with no family history of breast cancer. IFX discontinued and patient treated. Prostate cancer diagnosed approx 2 weeks after E72 infusion in 64 yr old patient with preexisting prostatitis (elevated PSA levels) at week E32. IFX disctontinued and patient treated. 2 patients, each on IFX 10 mg/kg, developed a skin neoplasm. Neither resulted in discontinuation of treatment. 1 patient with extensive disease and 10 yr UC history at main study					

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
				baseline received IFX 5 mg/kg and demonstrated colonic dysplasia during extension studies.					

Safety – heart failure, malignancies, hepatobiliary events, autoimmune processes, neurological events, haematological reactions (paediatric population trial)

Trial name	Treatment arm	Time point	Heart failure	Malignancies and lymphoproliferative disorders	Hepatobiliary events / liver enzyme changes	Autoimmune processes (e.g. lupus-like syndrome)	Neurological events	Haematologic reactions	Other
Hyams	IFX q8w	NR	NR	NR	NR	NR	NR	NR	NR
Hyams	IFX q12w	NR	NR	NR	NR	NR	NR	NR	NR

Appendix 7: Quality of life tables

Quality of life outcomes

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ULTRA1	PBO	[345] Change from baseline at week 4: IBDQ overall, 146 SF-36 mental and physical component summary, 43 [345] Change from baseline value at week 8: IBDQ overall, 152 SF-36 mental and physical component summary, 44 [62] IBDQ mean response (SD) at week 8 (n=130): 75 (57.7)
ULTRA1	ADA 160/80mg	[345] Change from baseline at week 4: IBDQ overall, 149 SF-36 mental and physical component summary, 45; p-value vs. PBO, <0.05 [345] Change from baseline value at week 8: IBDQ overall, 153 SF-36 mental and physical component summary, 46 [62] IBDQ mean response (SD) at week 8 (n=130): 70 (53.8); p-value vs. PBO, 0.532

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ULTRA2	PBO	[61] IBDQ (domain NR) Value at week 8: 20 (36) Value at week 32: 20 (41) Value at week 52: 19 (41) [61] Increase in IBDQ ≥ 16 points from baseline: Value at week 8: 112/246 (45.5%) Value at week 32: 54/246 (22.0%) Value at week 52: 40/246 (16.3%) Value at week 8, 32 and 52: 30/246 (12.2%)
ULTRA2	ADA 160/80mg	[61] IBDQ (domain NR) Value at week 8: 29 (36); p-value vs. PBO <0.05 Value at week 32: 28 (41); p-value vs. PBO <0.05 Value at week 52: 27 (42); p-value vs. PBO <0.05 [61] Increase in IBDQ ≥ 16 points from baseline: Value at week 8: 144/248 (58.1%); p-value vs. PBO, p<0.05 Value at week 32: 86/248 (34.7%); p-value vs. PBO, p<0.05 Value at week 52: 65/248 (26.2%); p-value vs. PBO, p<0.05 Value at week 8, 32 and 52: 58/248 (23.4%); p-value vs. PBO, p<0.05

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ULTRA3		[61] Value at week 12: IBDQ overall, 178.2 (34.60) SF-36 Physical, NA SF-36 Mental, NA [61] Value at week 48: IBDQ overall, 177.2 (34.94) SF-36 Physical, 49.6 (8.24) SF-36 Mental, 46.1 (10.77) [61] Value at week 108: IBDQ overall, 176.3 (37.15) SF-36 Physical, 49.4 (8.13) SF-36 Mental, 46.0 (11.00)
PURSUIT-SC	Phase II PBO	[48] Phase II PBO Change from baseline in IBDQ overall, N =41, Mean (SD) 14.8 (37.16), Median (IQR) 14.0 (- 2.0, 34.0) Read from graph: Randomised in Phase II, 13 Randomised while Phase II data analysed, 12.5 Randomised in Phase III, 12.5
PURSUIT-SC	Phase II GOL 100/50 mg (regimen discontinued after phase II)	[48] Phase II GOL 100/50 Change from baseline in IBDQ overall, N =40, Mean (SD) 26.2 (39.71), Median (IQR) 24.5 (- 5.5, 55.0)

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
PURSUIT-SC	Phase II GOL 200/100 mg all randomised	[48] Phase II GOL 200/100 Change from baseline in IBDQ overall, N =40 Mean (SD) 24.9 (36.89), Median (IQR) 16.0 (-2.5, 49.5) (P=0.287) (P=0.318) Read from graph: Randomised in Phase II, 14 Randomised while Phase II data analysed, 25 Randomised in Phase III, 27
PURSUIT-SC	Phase II GOL 400/200 mg all randomised	[48] Phase II GOL 400/200 Change from baseline in IBDQ, N, 40 Mean (SD) (31.6 (26.21), Median (IQR) 33.0 (9.0, 54.0) (P=0.021) Read from graph: Randomised in Phase II, 32 Randomised while Phase II data analysed,30 Randomised in Phase III, 25
PURSUIT-SC	Phase III PBO	[48] Phase III change from baseline IBDQ PBO N=251 Mean (SD), 14.8 (31.25), Median (IQR)= 11.0 (-3.0, 29.0)
PURSUIT-SC	Phase III GOL 200/100 mg phase III	[48] Phase III change from baseline vGOL 200/100 mg, n= 252/253 , Mean (SD)= 27.0 (33.72), Median (IQR)= 22.5 (0.5, 48.5) (P<0.0001)
PURSUIT-SC	Phase III GOL 400/200 mg phase III	[48] Phase III change from baseline IBDQ GOL 400/200 mg, n= 255/257, Mean (SD)= 26.9 (34.28), Median (IQR)= 21.0 (0.0, 50.0) (P<0.0001)

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
PURSUIT-SC		[346] Compared against PBO, significantly greater improvements experienced in combined GOL-treated group in IBDQ (27.2 vs. 14.6 P<0.001), physical component summary (4.14 vs. 2.46 P<0.01) and mental component summary (4.89 vs. 1.60 P<0.001) at week 6. Mean improvements in IBDQ (27.4 and 27.0), physical component summary (4.51 and 3.78) and mental component summary (4.69 and 5.10) comparable for GOL 200/100 mg and 400/200 mg groups. Distributions of IBDQ score chnged from mean of 129.4 (SD33.9) at baseline to 156.5 (SD 39.8) at week 6 in GOL-treated patients, with 45.2% patients achieving IBDQ remission vs. PBO group mean of 144.2 (SD 37.1) with 28.1% achieiving IBDQ remission (P<0.001 vs. combined GOL group). [346] In cumulative percentage curve vs. PBO, greater proportions of patients in each GOL group achieved "any improvement" to "clinically meaningful improvement" in IBDQ (51.1% vs. 35.2% P<0.001), physical component summary (41.0% vs. 31.6% P=0.01) and mental component summary (42.7% vs. 28.5% P<0.001) at week 6.
PURSUIT-SC and PURSUIT-M		[341] GOL-treated patients achieiving clinical remission at week 6 displayed greater mean improvement in physical component summary, mental component summary, EQ5D and IBDQ than those not achieving remission (physical component summary 8.0 vs. 2.9 P<0.001, mental component summary 10.7 vs. 2.6 P<0.001, EQ5D 21.4 vs. 7.2 P<0.001 and IBDQ 54.7 vs. 17.7 P<0.001). [341] Patients in clinical remission more likely to achieve normalised physical component summary, normalised mental component summary and IBDQ remission than those not achieving clinical remission (physical component summary 53.6% vs. 25.3% P<0.001, mental component summary 63.6% vs. 31.6% P<0.001, IBDQ 85.5% vs. 32.2% P<0.001). Furthermore, GOL-treated patients achieiving clinical remission during induction and maintained clinical remission at wee 54 in maintenane were also more likely to achieve normalised physical component summary, mental component summary and IBDQ remission than those not (physical component summary 73.5% vs. 22.7% P<0.001, mental component summary 63.3% vs. 28.4% P<0.001, IBDQ remission 89.8% vs. 22.7& P<0.001).
UC-SUCCESS	AZA	[52] Change from baseline at week 8: IBQD overall, n=50: 37.84; p-value between IFX, 0.539; IFX/AZA, 0.070 Change from baseline at week8: SF-36 physical function, n=58: 3.45; p-value between IFX, 0.422; IFX/AZA, 0.044 [52] Change from baseline at week 16: IBQD overall, n=53: 32.51; p-value between IFX, 0.482; IFX/AZA, <0.001 Change from baseline at week 16: SF-36 physical function, n=54: 4.13; p-value between IFX, 0.522; IFX/AZA, 0.052

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
UC-SUCCESS	IFX	[52] Change from baseline at week 8: IBQD, n=53: 33.42; p-value between IFX/AZA, 0.003 Change from baseline at week 8: SF-36 physical function, n=63: 3.24; p-value between IFX/AZA, 0.010 [52] Change from baseline at week 16: IBQD, n=58: 38.55; p-value between IFX/AZA, 0.004 Change from baseline at week 16: SF-36 physical function, n=59: 4.10; p-value between IFX/AZA, 0.022
UC-SUCCESS	IFX/AZA	[52] Change from baseline at week 8: IBQD, n=53: 49.83 Change from baseline at week 8: SF-36 physical function, n=59: 6.42 [52] Change from baseline at week 16: IBQD, n=57: 57.70 Value at week 16: SF-36 physical function, n=59: 7.70
Probert	PBO	[51] Change from baseline at week 6: IBQD (domain NR) 25 (28) Change from baseline at week 6: EuroQOL (domain NR) 4 (16)
Probert	IFX	[51] Change from baseline at week 6: IBQD (domain NR) 36 (49); p-value vs. PBO, 0.22 Change from baseline at week 6: EuroQOL (domain NR) 7 (17); p-value vs. PBO, 0. 3
ACT1	PBO	[347] Change from baseline at week 8: SF-36 physical component summary, 4.5 (6.8) SF-36 mental component summary, 3.1 (9.7) [347] Change from baseline at week 8: SF-36 physical component summary, 2.9 (6.0) SF-36 mental component summary, 3.1 (9.7)

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1	PBO	[348] Mean IBDQ scores (read from graph): Week 8, 20.70 Week 30, 17.83 Week 54, 12.33
ACT1	IFX 5mg/kg	[347] Change from baseline at week 8: SF-36 physical component summary, 6.8 (7.8); p-value vs. PBO <0.05 SF-36 mental component summary, 5.6 (10.2); p-value vs. PBO <0.05 [347] Change from baseline at week 8: SF-36 physical component summary, 6.8 (7.4); p-value vs. PBO <0.05 SF-36 mental component summary, 6.1 (10.8); p-value vs. PBO <0.05
ACT1	IFX 5mg/kg	[348] Mean IBDQ scores (read from graph): Week 8, 41.72 Week 30, 33.31 Week 54, 32.38
ACT1	IFX 10mg/kg	[348] Mean IBDQ scores (read from graph): Week 8, 34.71 Week 30, 35.01 Week 54, 31.42

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1	IFX 10mg/kg	[347] Change from baseline at week 8: SF-36 physical component summary, 5.6 (7.8) SF-36 mental component summary, 6.6 (12.0); p-value vs. PBO <0.05 [347] Change from baseline at week 8: SF-36 physical component summary, 6.2 (7.9); p-value vs. PBO <0.05 SF-36 mental component summary, 6.2 (10.7); p-value vs. PBO <0.05
ACT1	IFX combined	[347] Change from baseline at week 8: SF-36 physical component summary, 6.2 (7.8); p-value vs. PBO <0.05 SF-36 mental component summary, 6.1 (11.1); p-value vs. PBO <0.05 [347] Change from baseline at week 8: SF-36 physical component summary, 6.5 (7.7); p-value vs. PBO <0.05 SF-36 mental component summary, 6.2 (10.7); p-value vs. PBO <0.05
ACT1	IBDQ: responders who were not in remission (n=150)	[349] Change from baseline at week 8 IBDQ (domain NR), 47 (P<0.001 vs. nonresponders)
ACT1	IBDQ: responders who were in remission (n=206)	[349] Change from baseline at week 8 IBDQ (domain NR), 65

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1	Patients who had not discontinued corticosteroids at week 30 (n=21)	[349] Change from baseline at week 30: IBDQ 55.2 [58.0] SF-36 physical component summary 5.9 [7.4] SF-36 mental component summary 11.4 [9.6]
ACT1	Patients who had discontinued corticosteroids at week 30 (n=70)	[349] Change from baseline at week 30: IBDQ (domain NR), 64.7 [65.5] SF-36 physical component summary 9.8 [10.4] SF-36 mental component summary 11.0 [9.2]
ACT 2	PBO	[348] Mean IBDQ scores (read from graph): Week 8, 19.81 Week 30, 17.87
ACT 2	IFX 5mg/kg	[348] Mean IBDQ scores (read from graph): Week 8, 38.65 Week 30, 31.64
ACT 2	IFX 10mg/kg	[348] Mean IBDQ scores (read from graph): Week 8, 35.75 Week 30, 35.99
ACT1&2 combined	IBDQ: non-responders (n=137)	[349] Mean change from baseline IBDQ (no SD for all), 12 [349] Significantly greater proportions of patients in responder and remission subgroups achieved at least a 16 point increase (87% and 96%) respectively ot a 32 point increase (68% and 87%) in total IBDQ score vs. patients classed as nonresponders (39% and 26% respectively, P<0.001 for all comparisons).

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1&2 combined	PBO	[347] Change from baseline at week 8 n=244: IBDQ Bowel, 7.9 (9.7) IBDQ Emotional, 6.2 (10.6) IBDQ Systemic, 3.0 (4.8) IBDQ Social, 3.8 (6.0) SF-36 Physical functioning, 6.0 (17.3) SF-36 Role-physical, 22.4 (39.7) SF-36 Bodily pain, 13.1 (24.7) SF-36 General health, 5.6 (15.8) SF-36 Vitality, 11.5 (20.7) SF-36 Social functioning, 15.8 (24.8) SF-36 Role-emotional, 12.4 (47.6) SF-36 Mental health, 5.0 (18.4) SF-36 physical component summary, 3.7 (6.5) SF-36 mental component summary, 3.0 (9.6) [347] Percentage of Patients Who Achieved Clinically Meaningful Improvement at Value at week 8: IBDQ change ≥ 16, 49.6% IBDQ change ≥ 32, 32.6% SF-36 physical component summary change ≥ 3, 40.6% SF-36 mental component summary change ≥ 3, 32.4% SF-36 physical component summary change ≥ 5, 34.0% SF-36 mental component summary change ≥ 5, 29.2%

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1&2 combined	IFX 5mg/kg	[347] Change from baseline week to 8 n=242: IBDQ Bowel, 14.5 (11.7); p-value vs. PBO <0.05 IBDQ Emotional, 12.7 (12.6); p-value vs. PBO <0.05 IBDQ Systemic, 5.7 (5.9); p-value vs. PBO <0.05 IBDQ Social, 7.4 (8.0); p-value vs. PBO <0.05 SF-36 Physical functioning, 12.8 (19.3); p-value vs. PBO <0.05 SF-36 Role-physical, 29.6 (41.0) SF-36 Bodily pain, 20.2 (22.5); p-value vs. PBO <0.05 SF-36 General health, 10.0 (16.9); p-value vs. PBO <0.05 SF-36 Vitality, 11.5 (20.7 16.6 (22.0); p-value vs. PBO <0.05 SF-36 Social functioning, 21.2 (24.8); p-value vs. PBO <0.05 SF-36 Role-emotional, 15.5 (46.1) SF-36 Mental health, 10.6 (17.5); p-value vs. PBO <0.05 SF-36 physical component summary, 6.8 (7.6); p-value vs. PBO <0.05 SF-36 mental component summary, 5.9 (10.5); p-value vs. PBO <0.05 [347] Percentage of Patients Who Achieved Clinically Meaningful Improvement at Value at week 8: IBDQ change ≥ 16, 69.7%; p-value vs. PBO <0.05 IBDQ change ≥ 32, 56.8%; p-value vs. PBO <0.05 SF-36 physical component summary change ≥ 3, 62.0%; p-value vs. PBO <0.05 SF-36 mental component summary change ≥ 3, 52.1%; p-value vs. PBO <0.05 SF-36 physical component summary change ≥ 5, 48.8%; p-value vs. PBO <0.05 SF-36 mental component summary change ≥ 5, 40.9%; p-value vs. PBO <0.05

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1&2 combined	IFX combined	[347] Change from baseline week to 8 n=484: IBDQ Bowel, 13.7 (11.8); p-value vs. PBO <0.05 IBDQ Emotional, 12.0 (12.6); p-value vs. PBO <0.05 IBDQ Systemic, 5.4 (5.9); p-value vs. PBO <0.05 IBDQ Social, 6.8 (7.5); p-value vs. PBO <0.05 SF-36 Physical functioning, 11.0 (18.9); p-value vs. PBO <0.05 SF-36 Role-physical, 31.1 (42.5); p-value vs. PBO <0.05 SF-36 Bodily pain, 20.0 (23.4); p-value vs. PBO <0.05 SF-36 General health, 10.4 (18.1); p-value vs. PBO <0.05 SF-36 Vitality, 18.3 (22.3); p-value vs. PBO <0.05 SF-36 Social functioning, 21.0 (25.9); p-value vs. PBO <0.05 SF-36 Role-emotional, 18.2 (45.4) SF-36 Mental health, 10.5 (18.2); p-value vs. PBO <0.05 SF-36 physical component summary, 6.4 (7.7); p-value vs. PBO <0.05 SF-36 mental component summary, 6.1 (10.9); p-value vs. PBO <0.05 [347] Percentage of Patients Who Achieved Clinically Meaningful Improvement at Value at week 8: IBDQ change ≥ 16, 68.7%; p-value vs. PBO <0.05 IBDQ change ≥ 32, 54.7%; p-value vs. PBO <0.05 SF-36 physical component summary change ≥ 3, 59.1%; p-value vs. PBO <0.05 SF-36 mental component summary change ≥ 3, 49.0%; p-value vs. PBO <0.05 SF-36 physical component summary change ≥ 5, 50.0%; p-value vs. PBO <0.05 SF-36 mental component summary change ≥ 5, 43.0%; p-value vs. PBO <0.05

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
ACT1&2 combined	IFX combined	[349] Mean change in PCS and MCS and individual SF-36 scale scores from baseline to week 30 by response status (read from graph): Phys comp sum: Non responders (n=137), 40.00; responders who were not in remission (n=150), 44.78; remission (n=206), 49.77 Phys Func: Non responders (n=137), 44.18; responders who were not in remission (n=150), 48.54; remission (n=206), 51.66 Role Phys: Non responders (n=137), 39.24; responders who were not in remission (n=150), 44.85; remission (n=206), 50.76 Bod pain : Non responders (n=137), 42.07; responders who were not in remission (n=150), 47.05; remission (n=206), 52.35 Gen health: Non responders (n=137), 36.19; responders who were not in remission (n=150), 40.24; remission (n=206), 45.54 Vitality: Non responders (n=137), 41.20; responders who were not in remission (n=150), 47.11; remission (n=206), 51.48 Soc funct: Non responders (n=137), 40.61; responders who were not in remission (n=150), 48.39; remission (n=206), 52.13 Role emot: Non responders (n=137), 44.06; responders who were not in remission (n=150), 48.42; remission (n=206), 51.58 Mental health: Non responders (n=137), 45.02; responders who were not in remission (n=150), 48.76; remission (n=206), 50.95 Mental comp sum: Non responders (n=137), 44.12; responders who were not in remission (n=150), 49.1; remission (n=206), 51.6
ACT1&2 combined	PBO IFX combined	[348] Mean baseline and change at week 8 scores per question for each IBDQ dimension (read from graph): Bowel: baseline, 3.33; week 8, 4.48 Emotional: baseline, 3.62; week 8, 4.44 Systemic: baseline, 2.89; week 8, 3.83 Social: baseline, 3.41; week 8, 4.57 Mean baseline and change at week 8 in norm-based SF-36 scale scores (read from graph): Phys Funct: baseline, 42.78; week 8, 45.97 Role phys: baseline, 33.89; week 8, 41.23 Body pain: baseline, 39.03; week 8, 45.10 Gen health: baseline, 34.28; week 8, 37.16 Vitality: baseline, 36.88; week 8, 42.30 Soc Funt: baseline, 34.68; week 8, 41.07 Role emot: baseline, 41.42; week 8, 45.26 Mental health: baseline, 41.14; week 8, 44.34 Mean baseline and change at week 8 scores per question for each IBDQ dimension (read from graph):

Study name	Treatment group	HRQoL instrument and domain, time point, mean (SD) [median] values, n/N (%) reporting improvement
		Bowel: baseline, 3.34; week 8, 5.12 Emotional: baseline, 3.80; week 8, 5.00 Systemic: baseline, 3.00; week 8, 4.22 Social: baseline, 3.74; week 8, 5.47 Mean baseline and change at week 8 in norm-based SF-36 scale scores (read from graph): Phys Funct: baseline, 44.07; week 8, 49.49 Role phys: baseline, 35.81; week 8, 45.07 Body pain: baseline, 39.68; week 8, 48.30 Gen health: baseline, 35.89; week 8, 40.68 Vitality: baseline, 37.84; week 8, 46.14 Soc Funt: baseline, 36.93; week 8, 46.82 Role emot: baseline, 42.39; week 8, 48.14 Mental health: baseline, 42.11; week 8, 47.85

Appendix 8: Network meta-analysis tables

Induction phase

	Treatments		Treatment 1				Treatment 2			
	1	**2**	**No Response**	**Response**	**Remission**	**Total**	**No Response**	**Response**	**Remission**	**Total**
Base case data										
ULTRA 1	Placebo	Adalimumab	72	46	12	130	59	47	24	130
ULTRA 2 (anti-TNF naïve)	Placebo	Adalimumab	89	40	16	145	61	57	32	150
PURSUIT-SC Phase II+III	Placebo	Golimumab	218	79	23	320	162	104	58	324
ACT 1	Placebo	Infliximab	76	27	18	121	37	37	47	121
ACT 2	Placebo	Infliximab	87	29	7	123	43	37	41	121
Sensitivity Analysis 1										
ULTRA 1	Placebo	Adalimumab	72	46	12	130	59	47	24	130
ULTRA 2 (ITT)	Placebo	Adalimumab	161	62	23248	246	123	84	41	248
PURSUIT-SC Phase II+III	Placebo	Golimumab	218	79	23	320	162	104	58	324
ACT 1	Placebo	Infliximab	76	27	18	121	37	37	47	121
ACT 2	Placebo	Infliximab	87	29	7	123	43	37	41	121
Sensitivity Analysis 2										
ULTRA 1	Placebo	Adalimumab	72	46	12	130	59	47	24	130
ULTRA 2 (anti-TNF naïve)	Placebo	Adalimumab	89	40	16	145	61	57	32	150
PURSUIT-	Placebo	Golimumab	218	79	23	320	162	104	58	324

SC Phase II+III										
ACT 1	Placebo	Infliximab	76	27	18	121	37	37	47	121
ACT 2	Placebo	Infliximab	87	29	7	123	43	37	41	121
SUZUKI	Placebo	Adalimumab	62	23	11	96	45	36	9	90
Sensitivity Analysis 3										
ULTRA 1	Placebo	Adalimumab	72	46	12	130	59	47	24	130
ULTRA 2 (ITT)	Placebo	Adalimumab	161	62	23	246	123	84	41	248
PURSUIT-SC Phase II+III	Placebo	Golimumab	218	79	23	320	162	104	58	324
ACT 1	Placebo	Infliximab	76	27	18	121	37	37	47	121
ACT 2	Placebo	Infliximab	87	29	7	123	43	37	41	121
SUZUKI	Placebo	Adalimumab	62	23	11	96	45	36	9	90

[redacted]

	[redacted]			[redacted]				[redacted]				[redacted]			
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

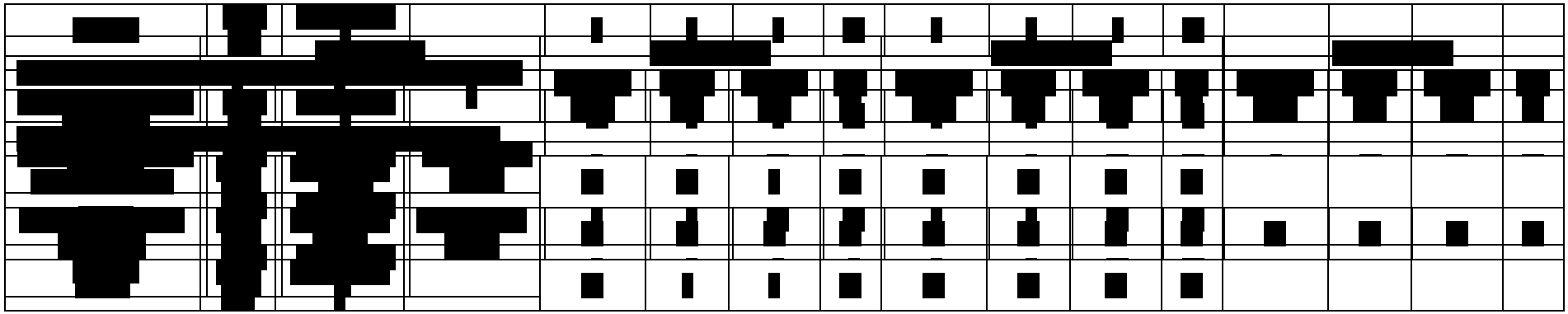
[redacted]

[redacted]

[redacted]

[redacted]

	[redacted]			[redacted]				[redacted]				[redacted]			
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]															
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]				



[redacted]

[redacted]	[redacted]	[redacted]		■	■	▮	■	■	■	■	■				
[redacted]	[redacted]	[redacted]		■	▮	▮	■	■	■	■	■				
[redacted]															
[redacted]	[redacted]	[redacted]		▮	▮	■	■	■	▮	■	■				
[redacted]	[redacted]	[redacted]	[redacted]	■	▮	■	■	■	▮	■	■	■	▮	■	■
[redacted]	[redacted]	[redacted]		▮	▮	■	■	■	▮	■	■				
[redacted]	[redacted]	[redacted]		▮	▮	▮	▮	■	■	■	■				
[redacted]	[redacted]	[redacted]		▮	▮	▮	■	▮	▮	▮	▮				
[redacted]															
[redacted]	[redacted]	[redacted]		■	■	▮	■	▮	■	▮	■				
[redacted]	[redacted]	[redacted]	[redacted]	▮	■	▮	■	▮	■	■	■	■	■	▮	■
[redacted]	[redacted]	[redacted]		▮	▮	▮	■	▮	▮	▮	■				
[redacted]	[redacted]	[redacted]		▮	▮	▮	■	▮	▮	▮	■				
[redacted]															
[redacted]	[redacted]	[redacted]		■	▮	■	■	▮	▮	■	■				
[redacted]	[redacted]	[redacted]	[redacted]	▮	▮	■	■	■	▮	■	■	▮	■	■	■
[redacted]	[redacted]	[redacted]		▮	▮	■	■	▮	▮	■	■				
[redacted]	[redacted]	[redacted]		▮	▮	▮	▮	▮	▮	■	■				

Appendix 9: Network meta-analysis figures

Results when using conventional reference prior for the between study standard deviation

Figure A9.1: Base case – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the induction phase (SD~U(0,2))

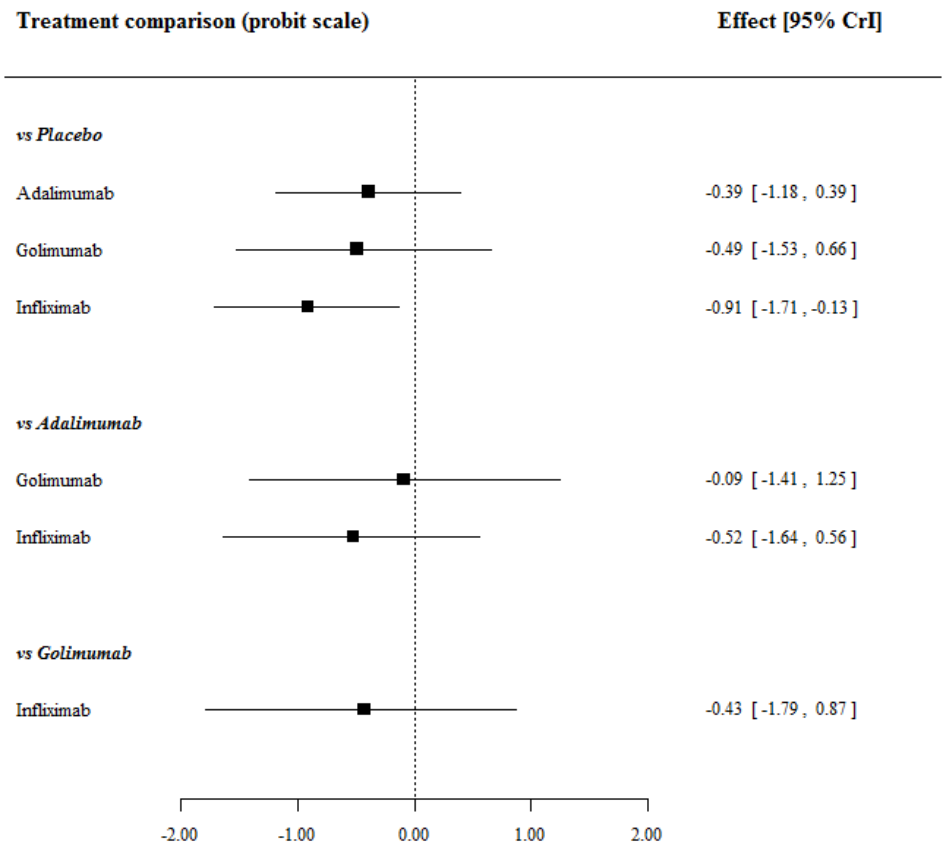


Figure A9.2: Base case – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~U(0,2))

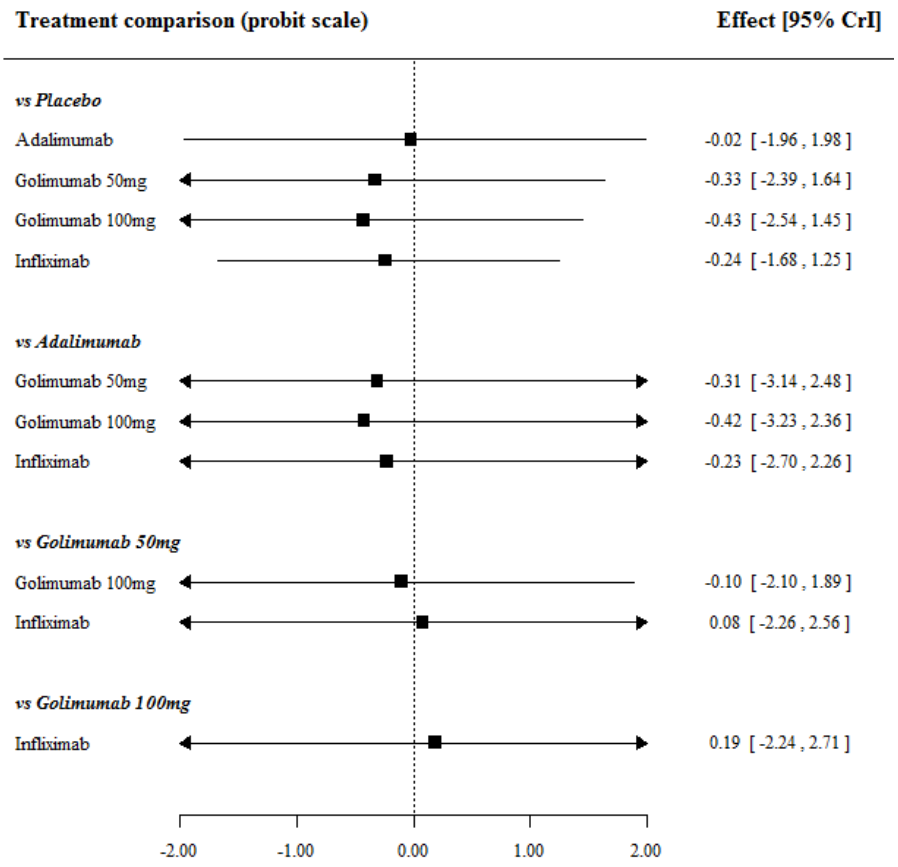


Figure A9.3: Base case – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~U(0,2))

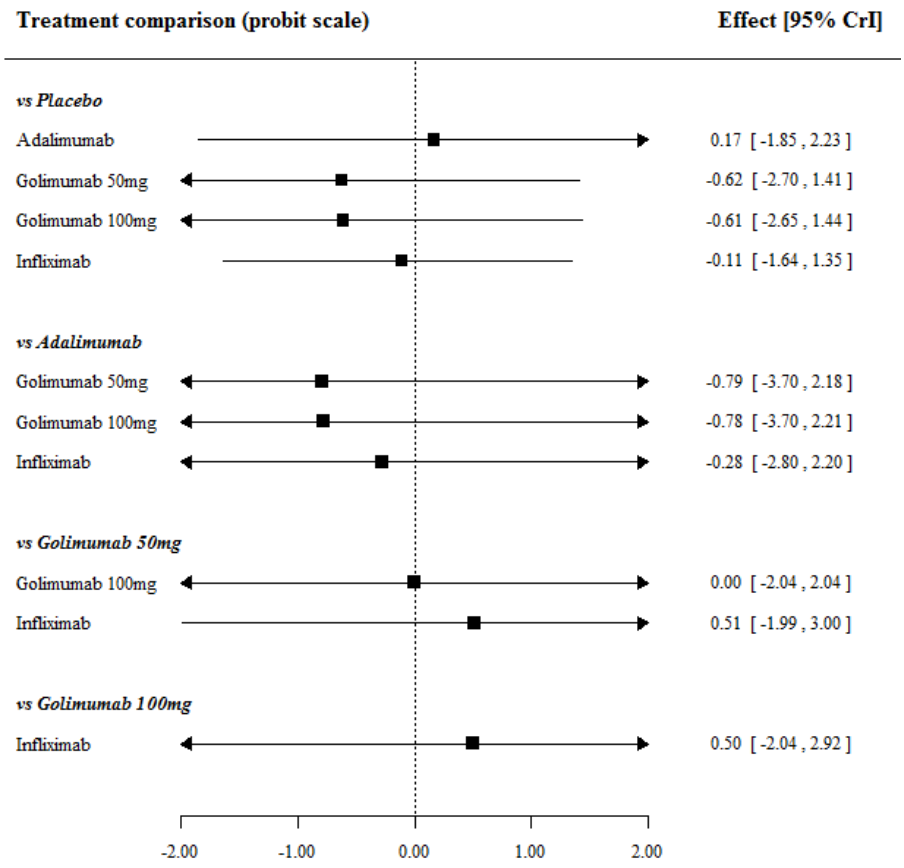


Figure A9.4: Base case – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~U(0,2))

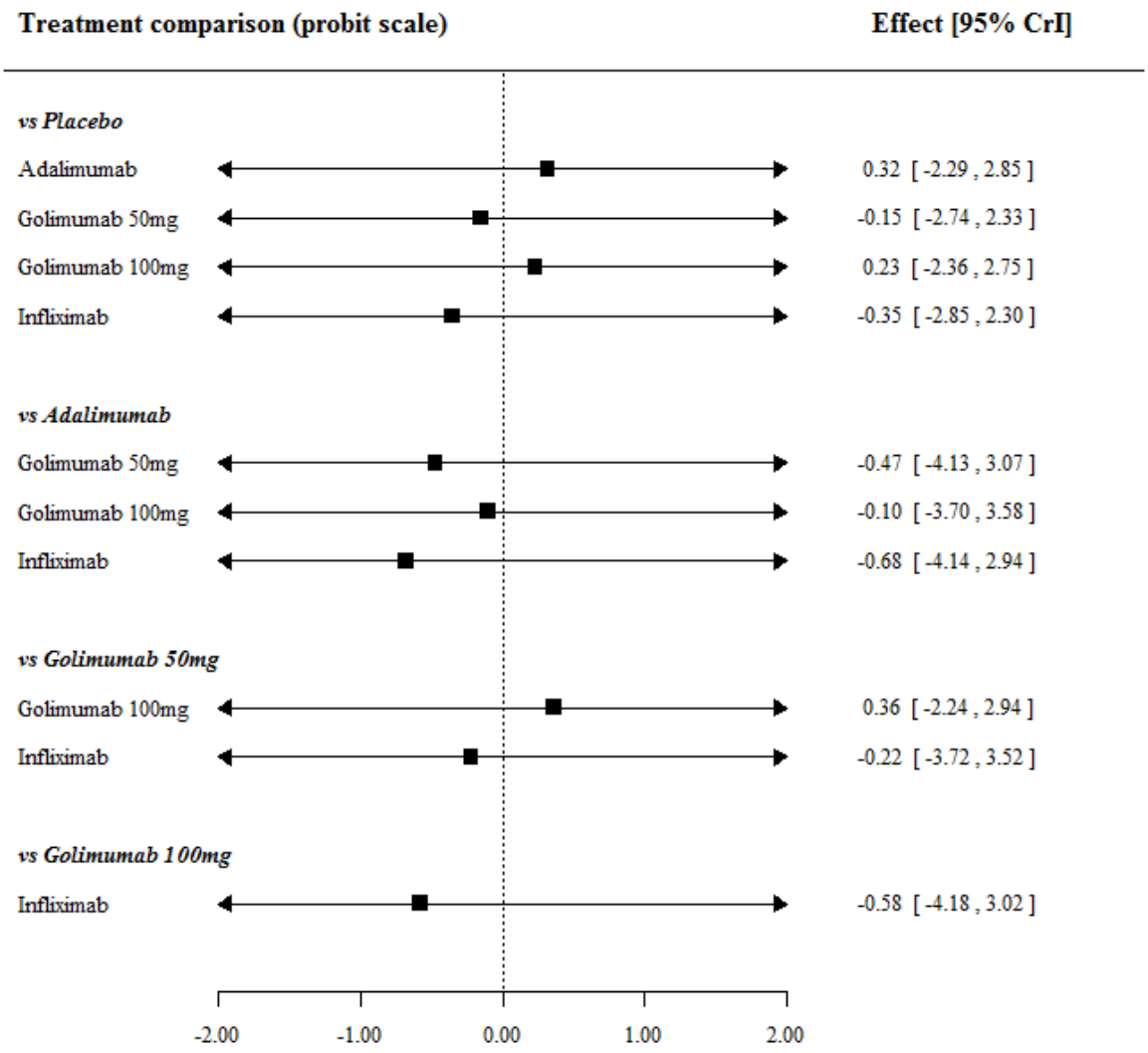


Figure A9.5: Base case – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~U(0,2))

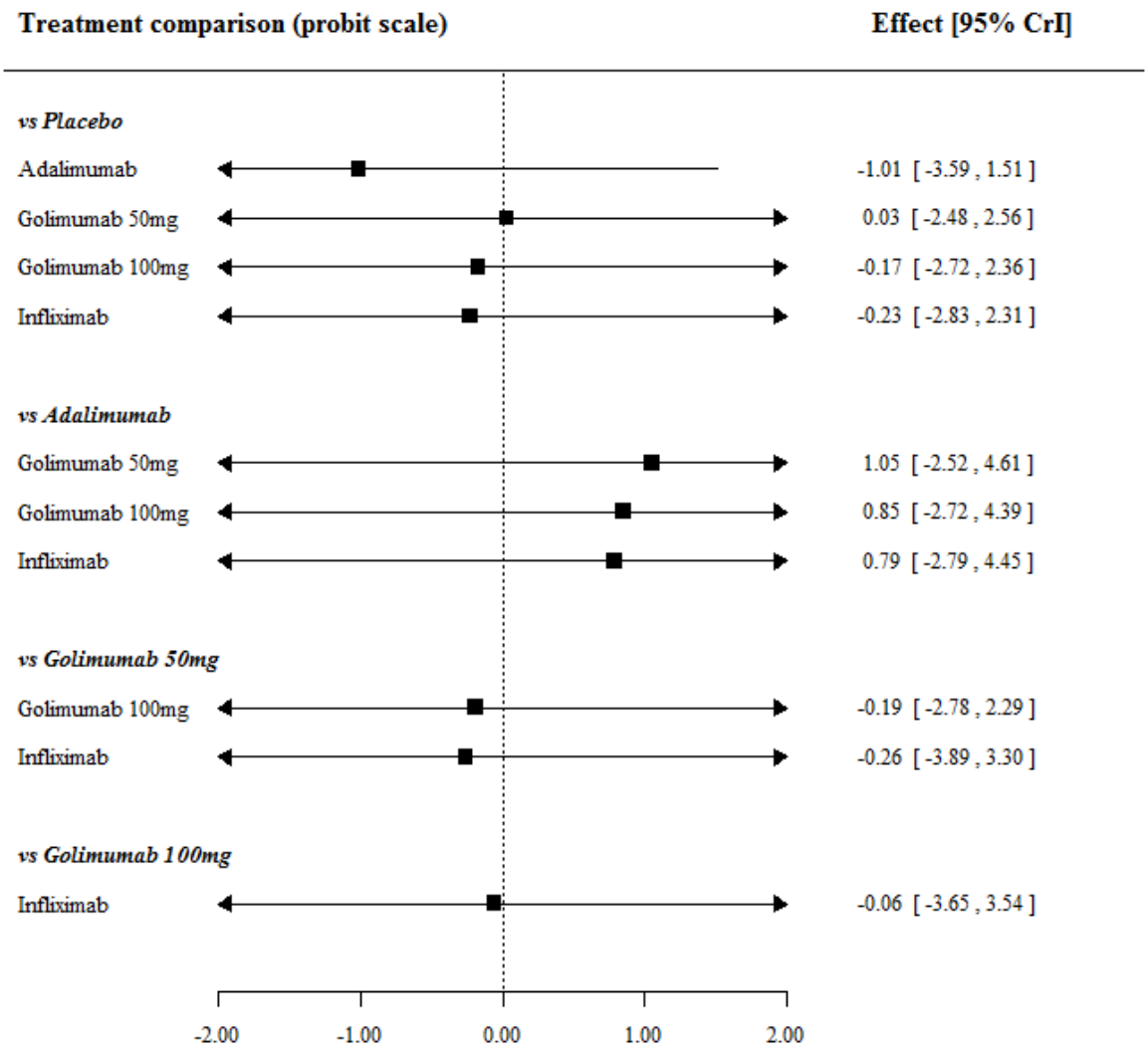


Figure A9.6: Sensitivity Analysis 1 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the induction phase (SD~U(0,2))

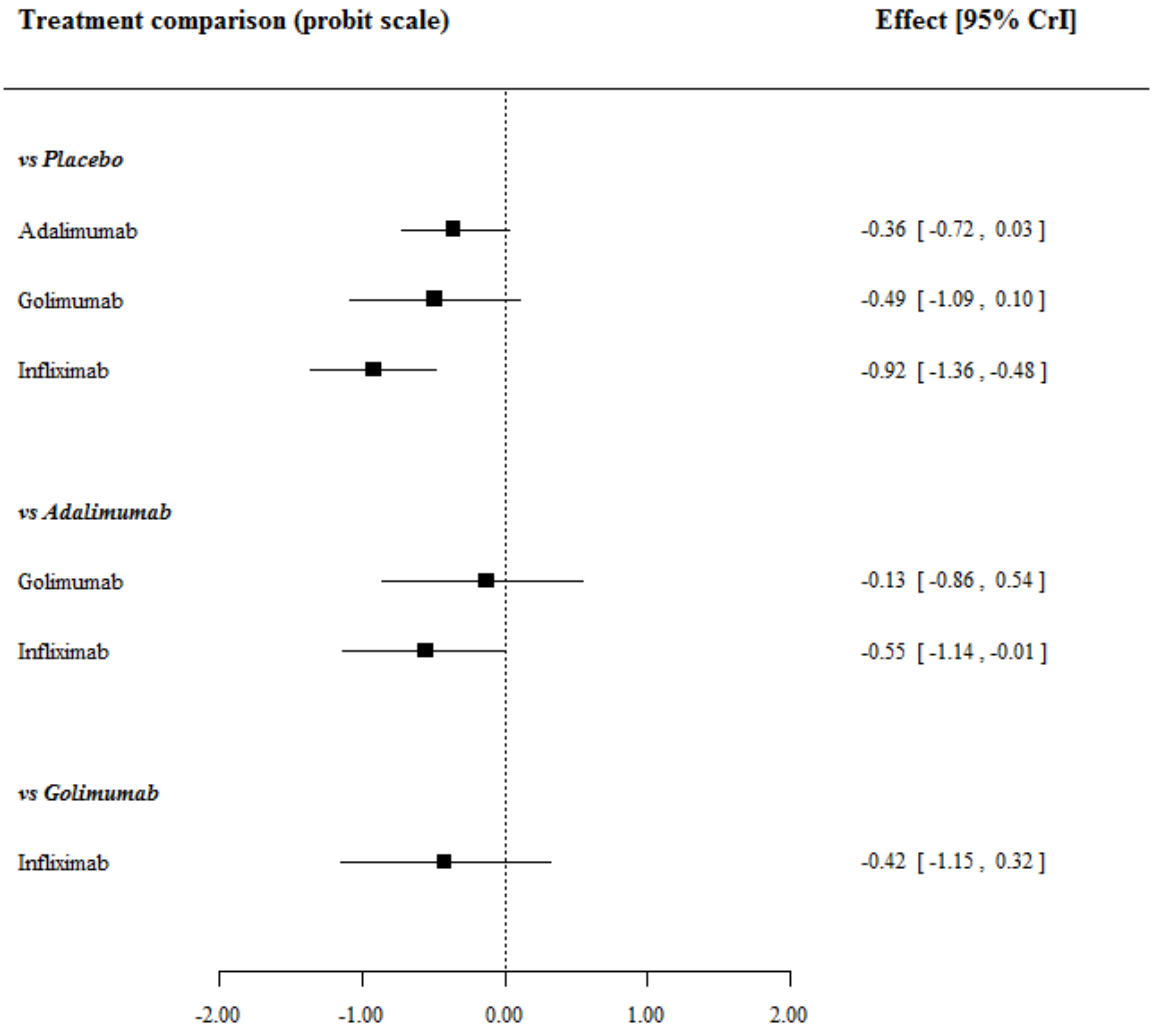


Figure A9.7: Sensitivity Analysis 1 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~U(0,2))

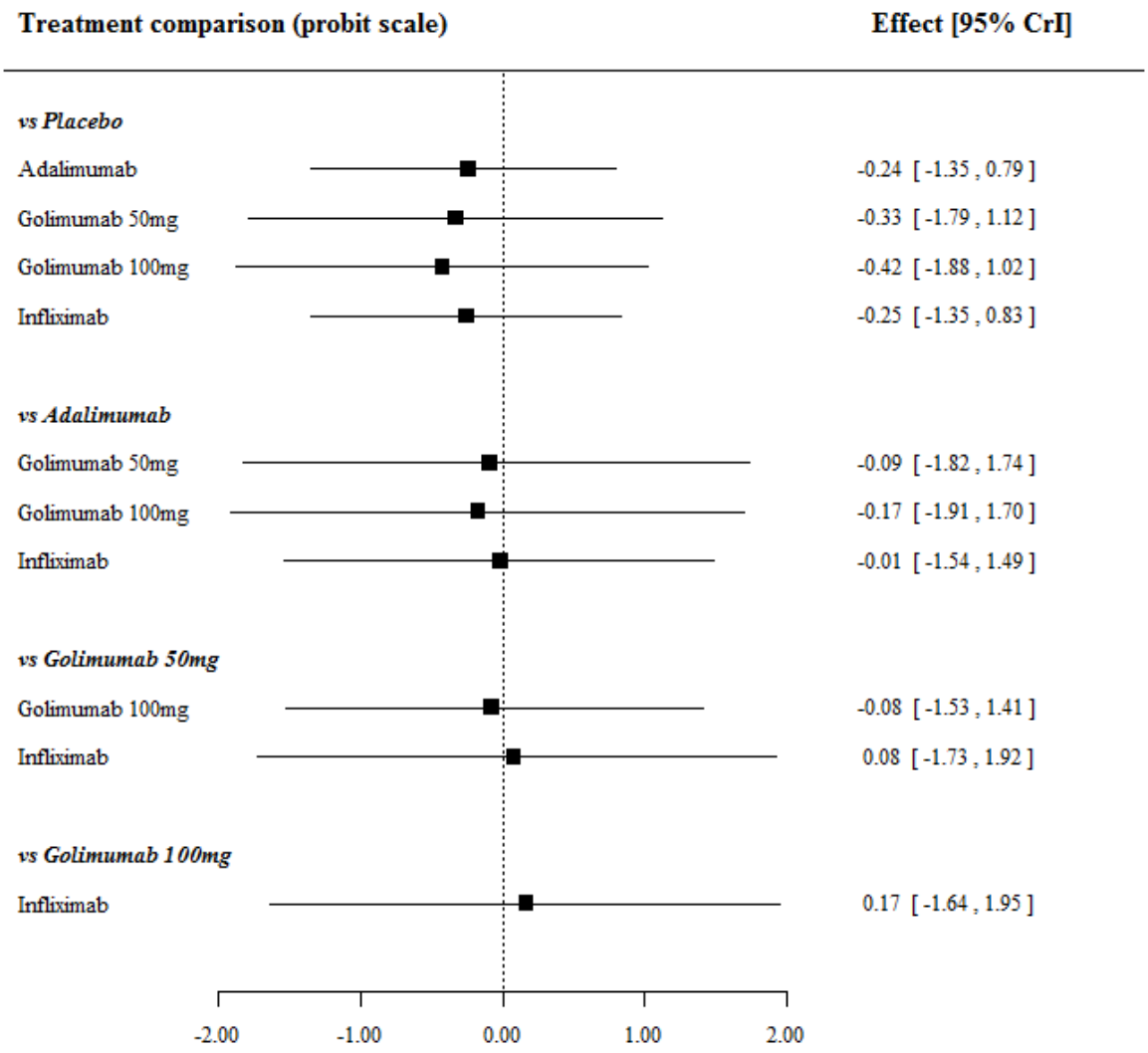


Figure A9.8: Sensitivity Analysis 1 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~U(0,2))

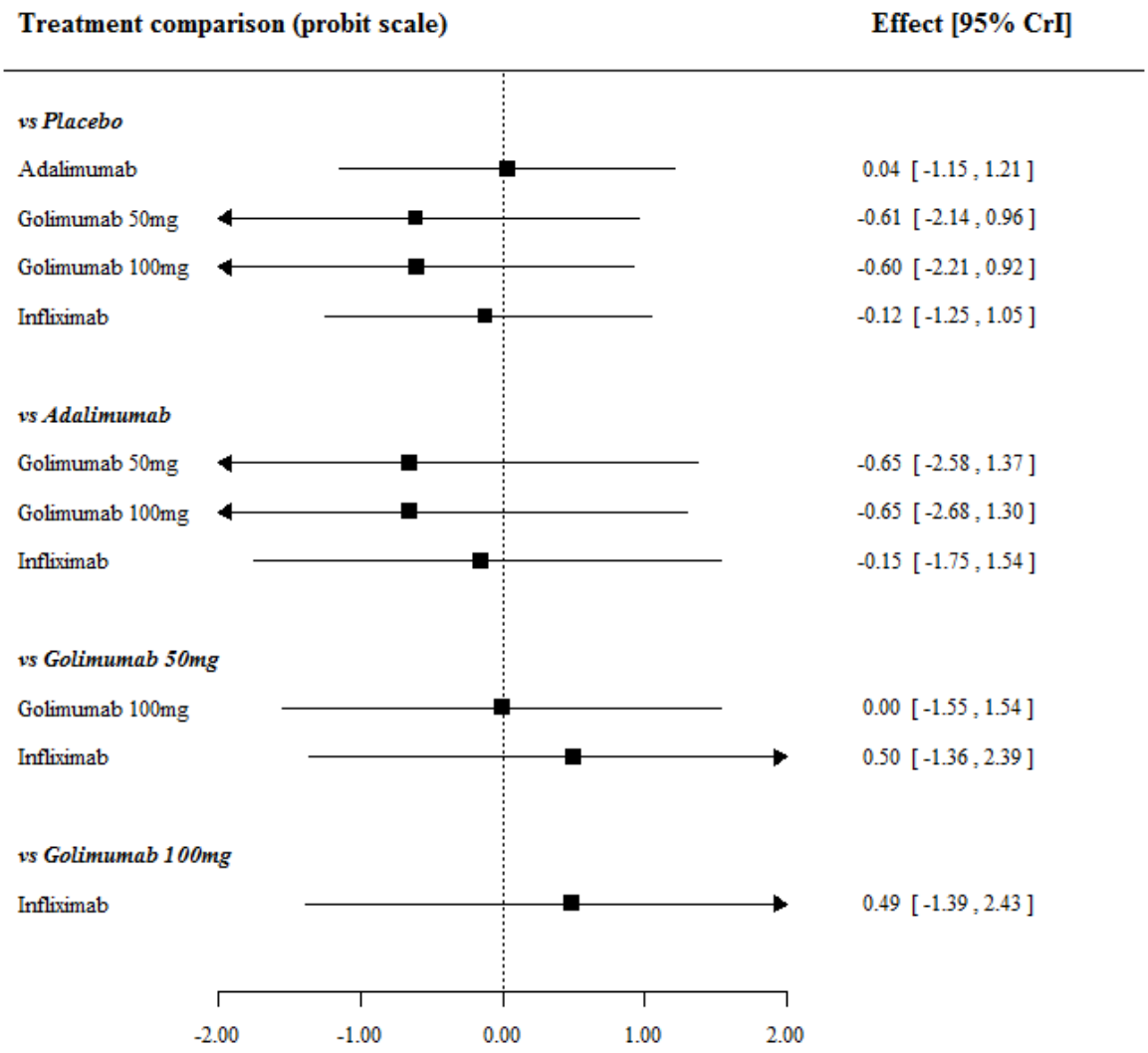


Figure A9.9: Sensitivity Analysis 1 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~U(0,2))

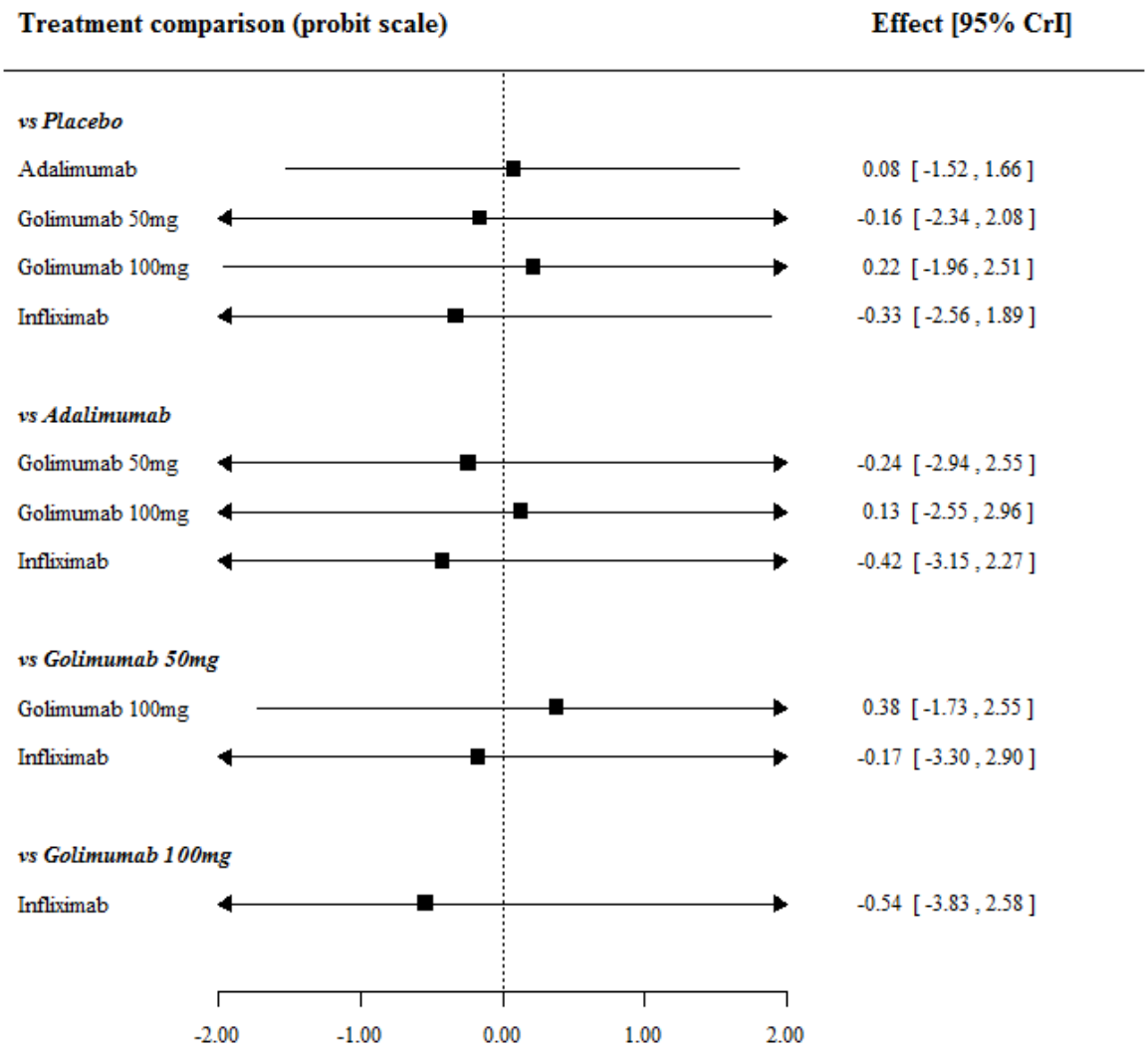


Figure A9.10: Sensitivity Analysis 1 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~U(0,2))

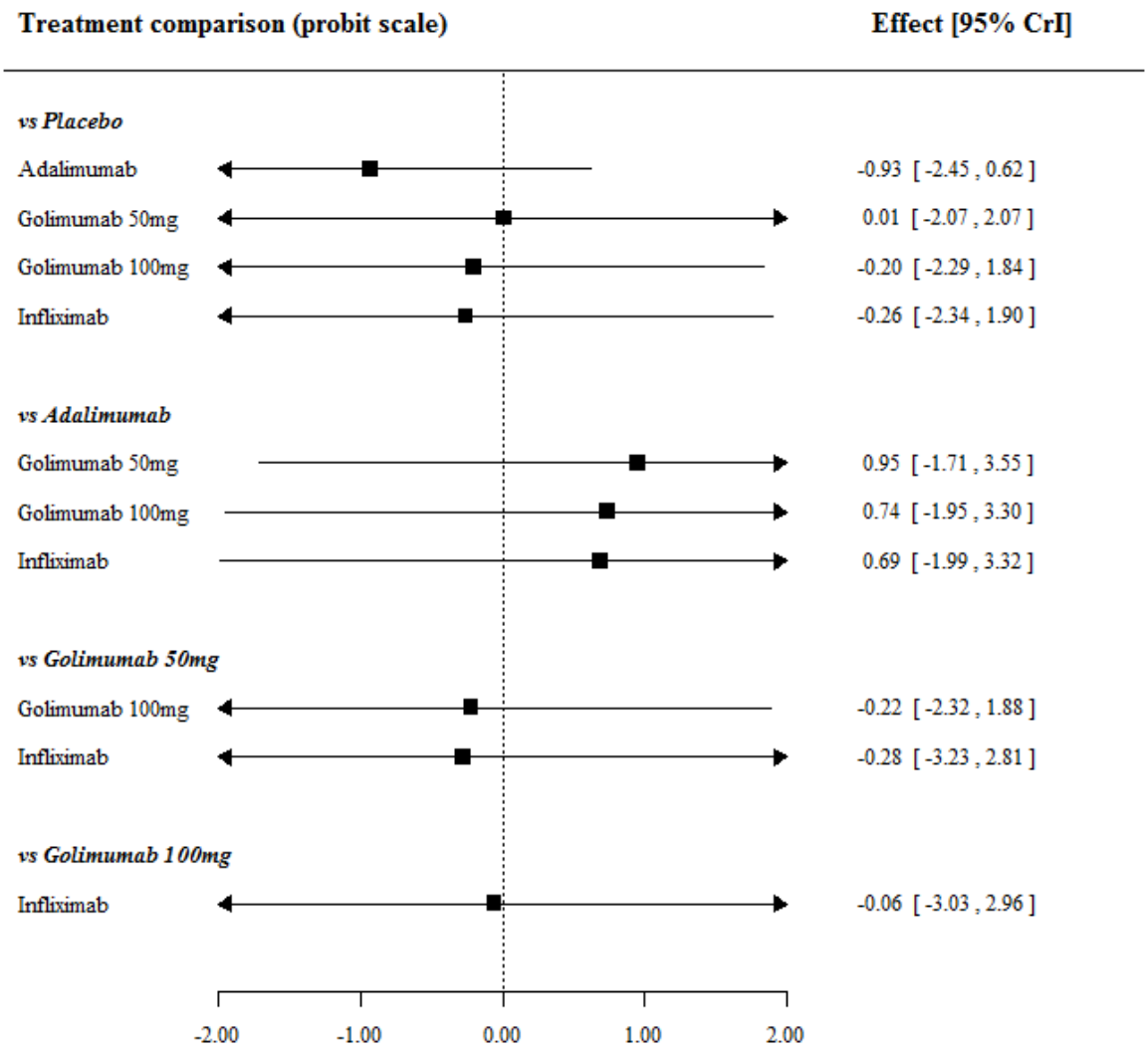


Figure A9.11: Sensitivity Analysis 2 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the induction phase (SD~U(0,2))

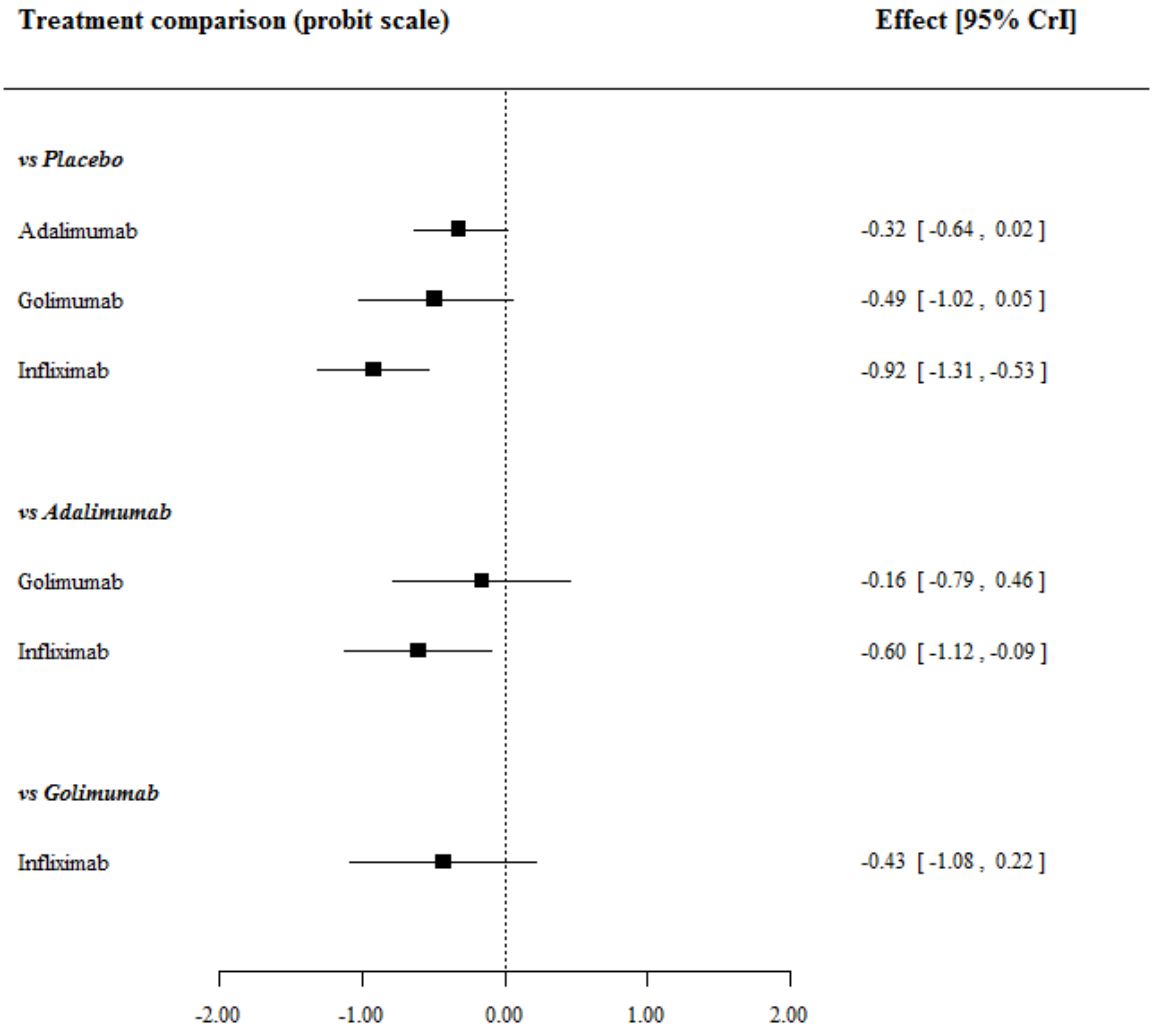


Figure A9.12: Sensitivity Analysis 2 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~U(0,2))

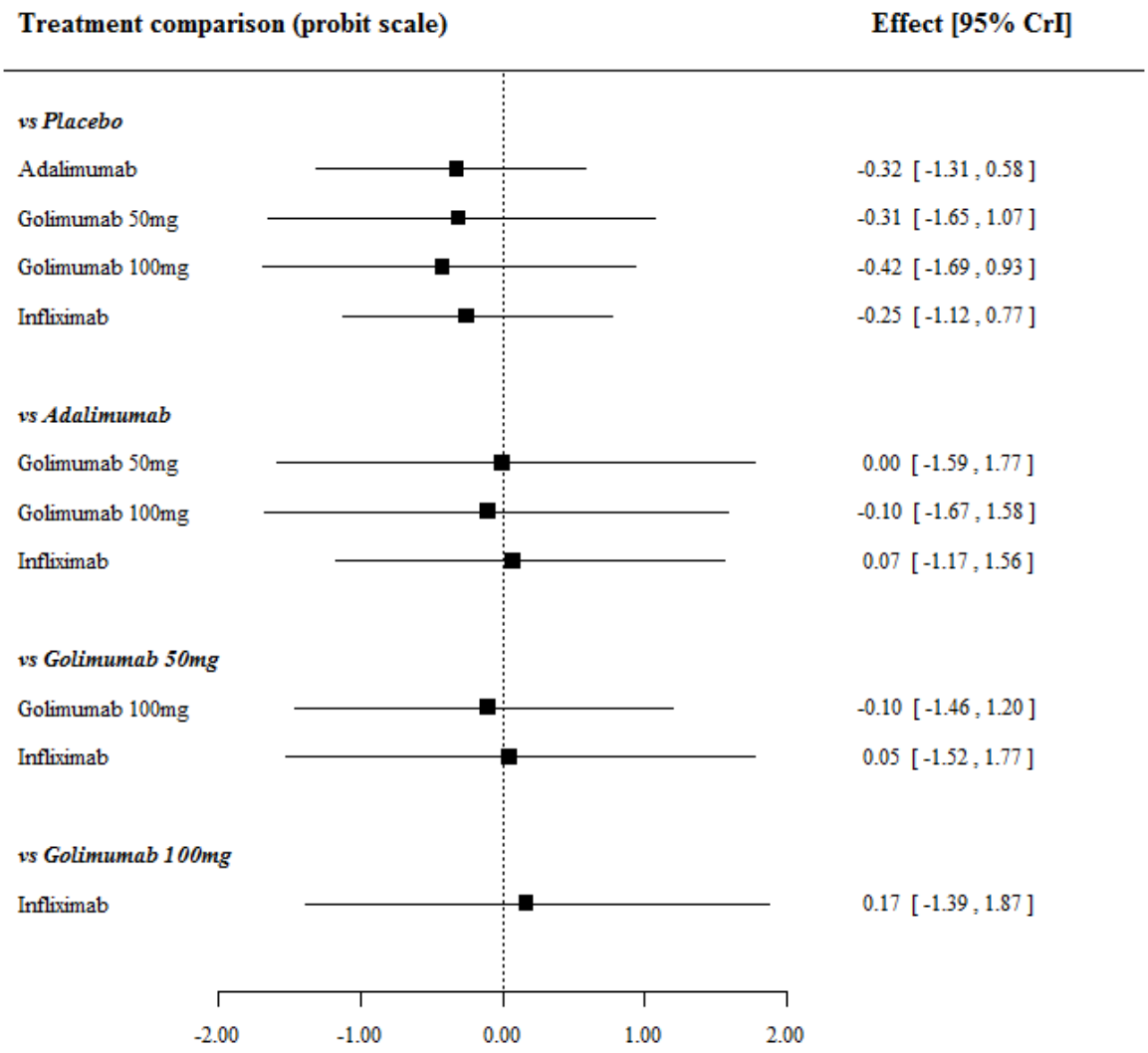


Figure A9.13: Sensitivity Analysis 2 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~U(0,2))

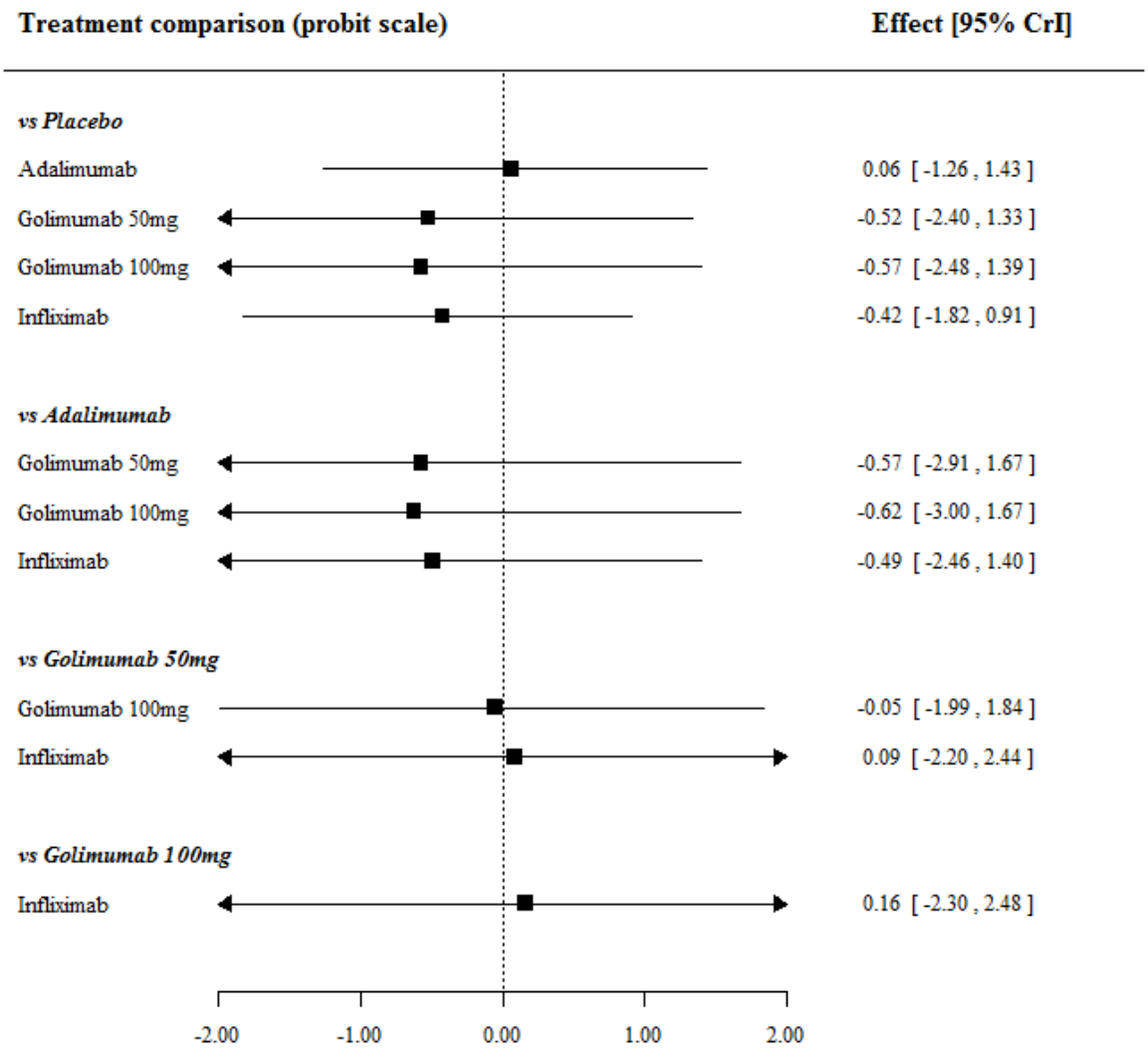


Figure A9.14: Sensitivity Analysis 2 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~U(0,2))

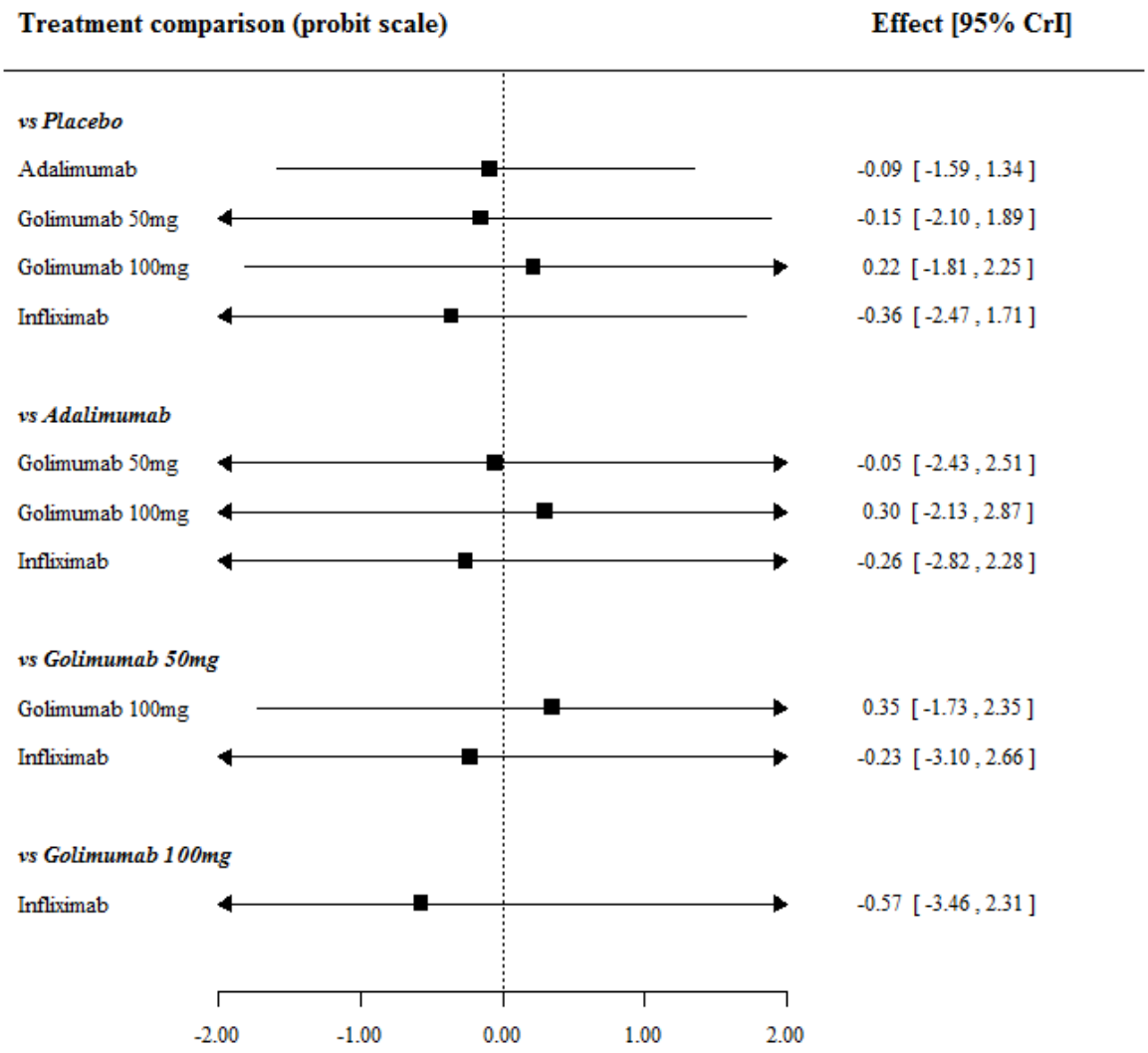


Figure A9.15: Sensitivity Analysis 2 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~U(0,2))

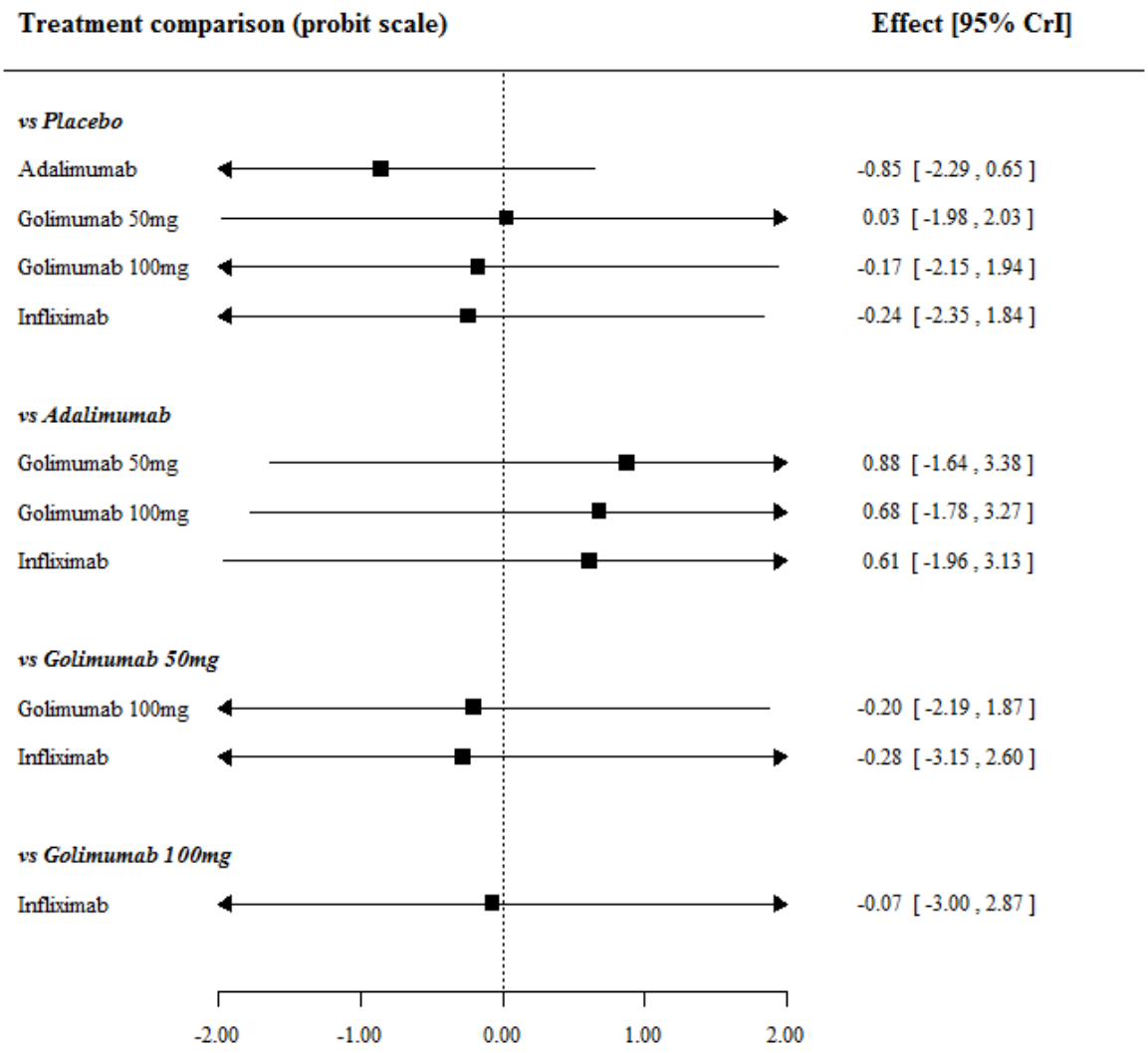


Figure A9.16: Sensitivity Analysis 3 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the induction phase (SD~U(0,2))

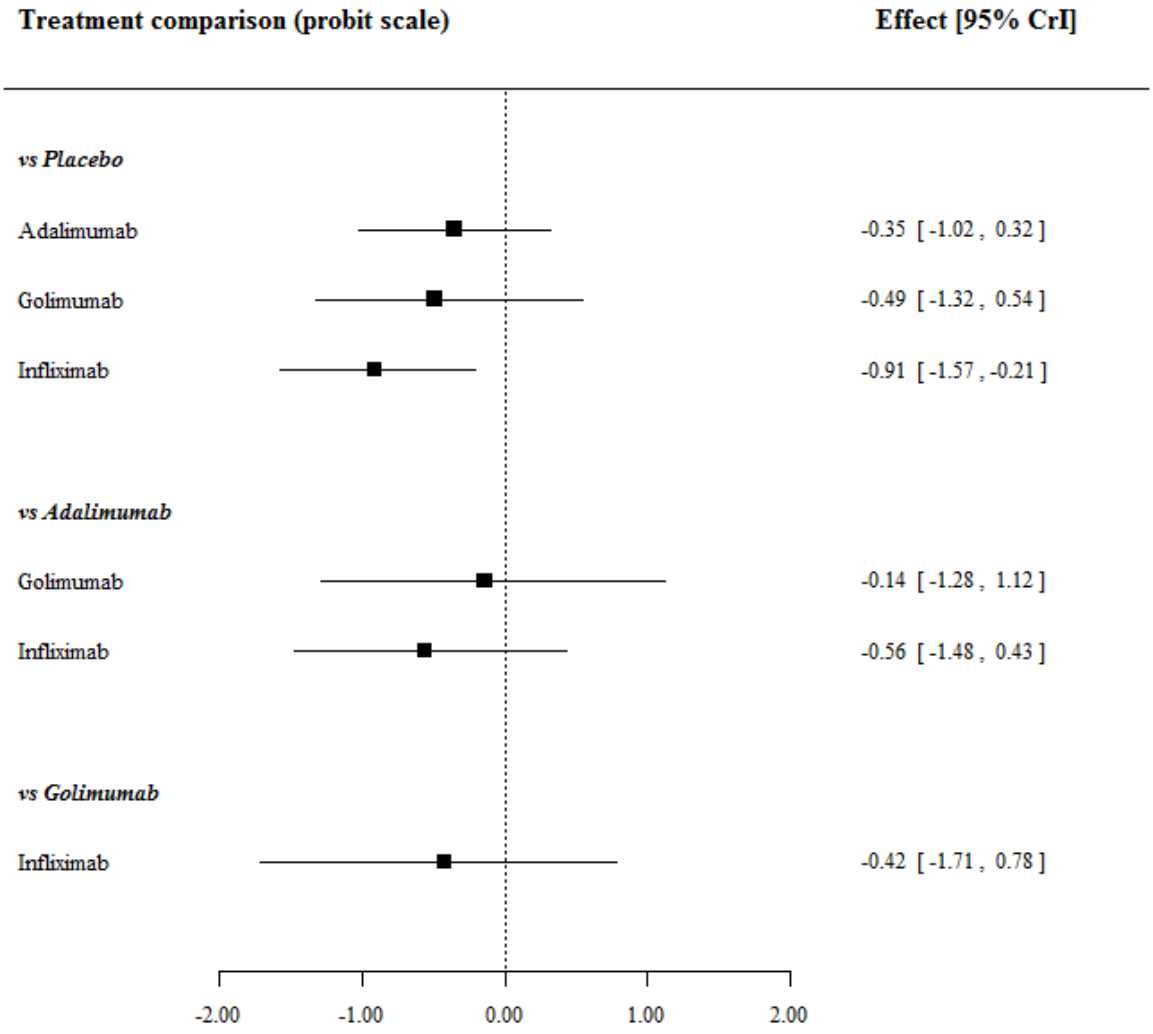


Figure A9.17: Sensitivity Analysis 3 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in response (SD~U(0,2))

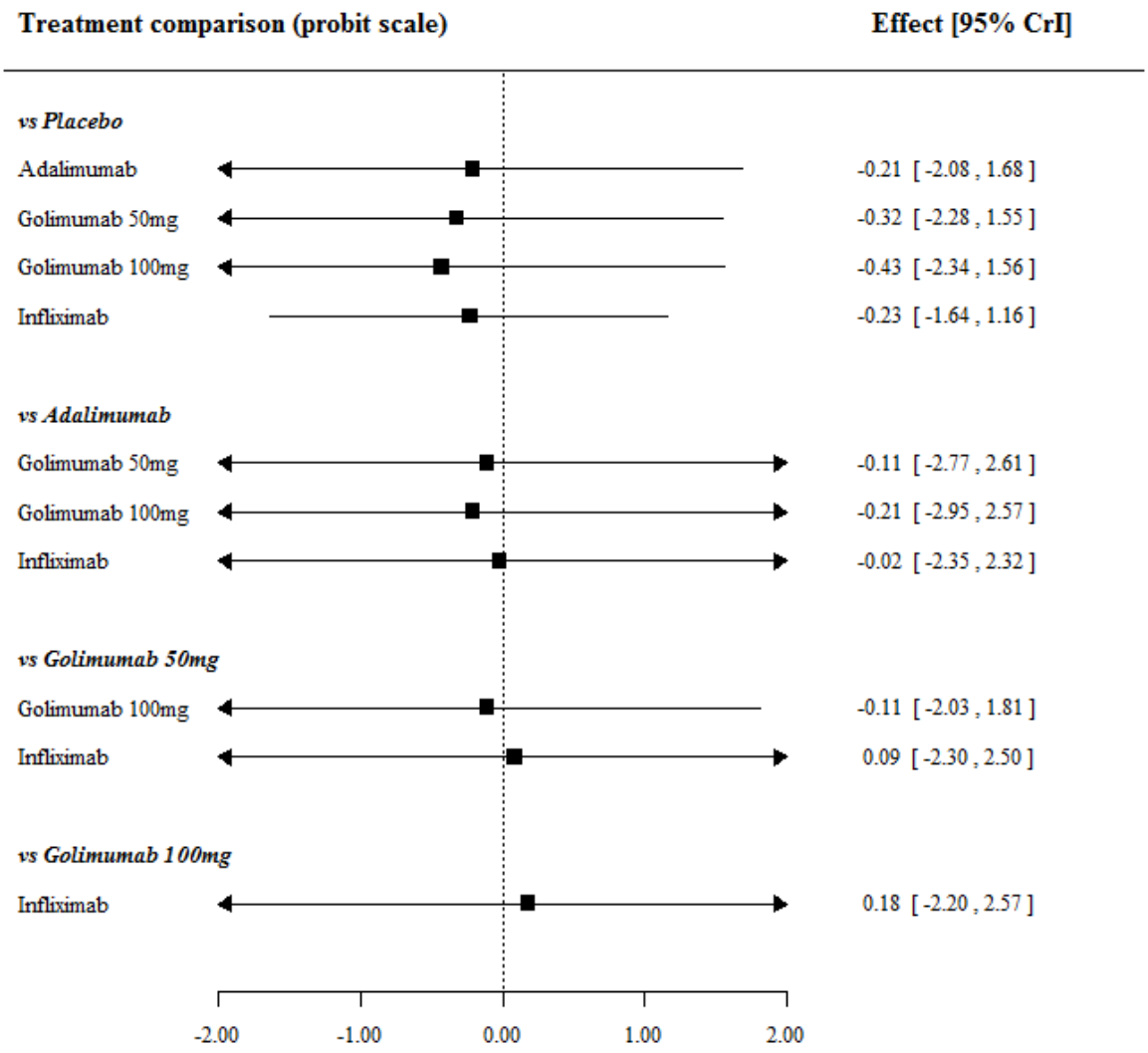


Figure A9.18: Sensitivity Analysis 3 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 8-32 weeks for patients starting in remission (SD~U(0,2))

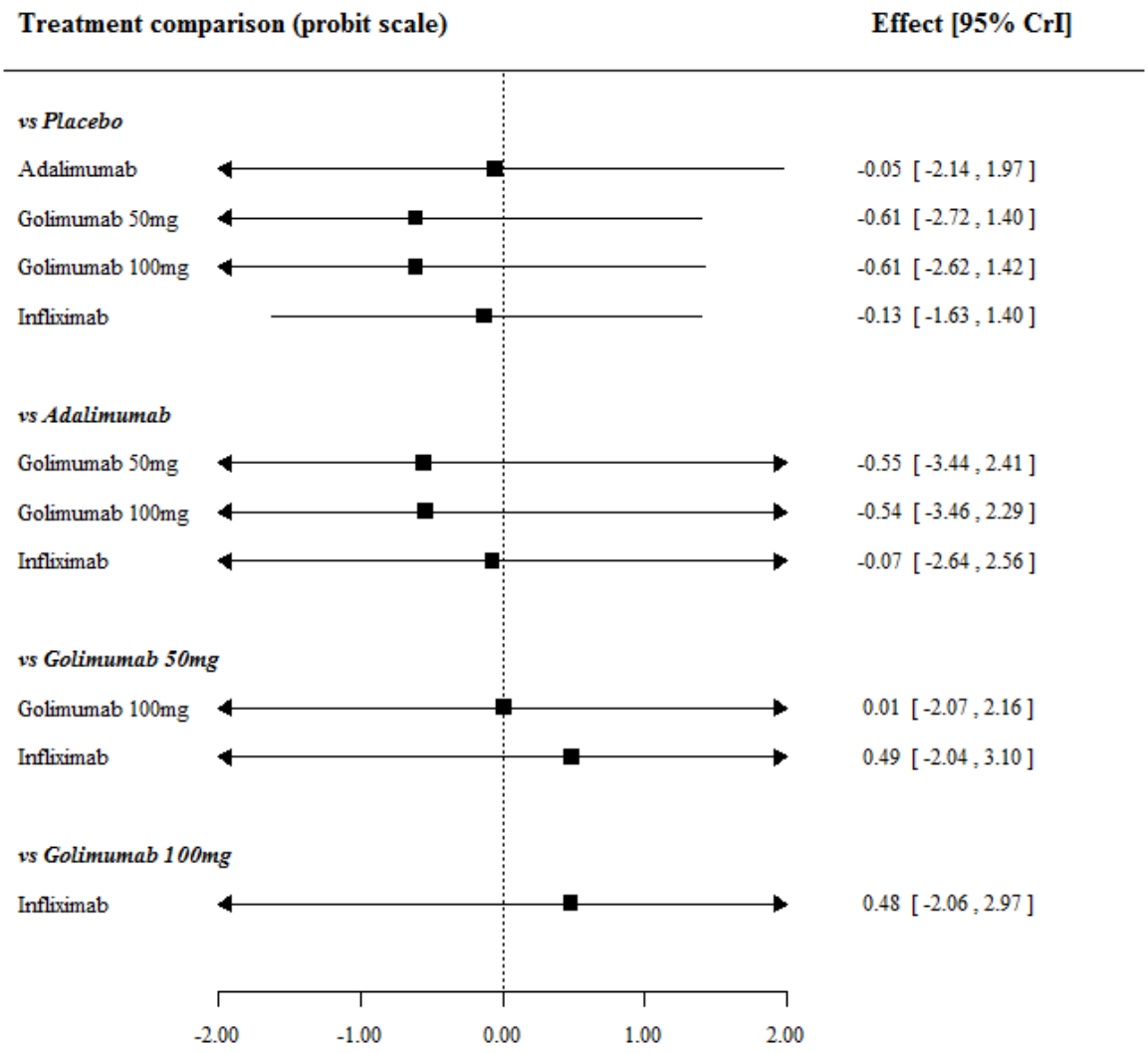


Figure A9.19: Sensitivity Analysis 3 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in response (SD~U(0,2))

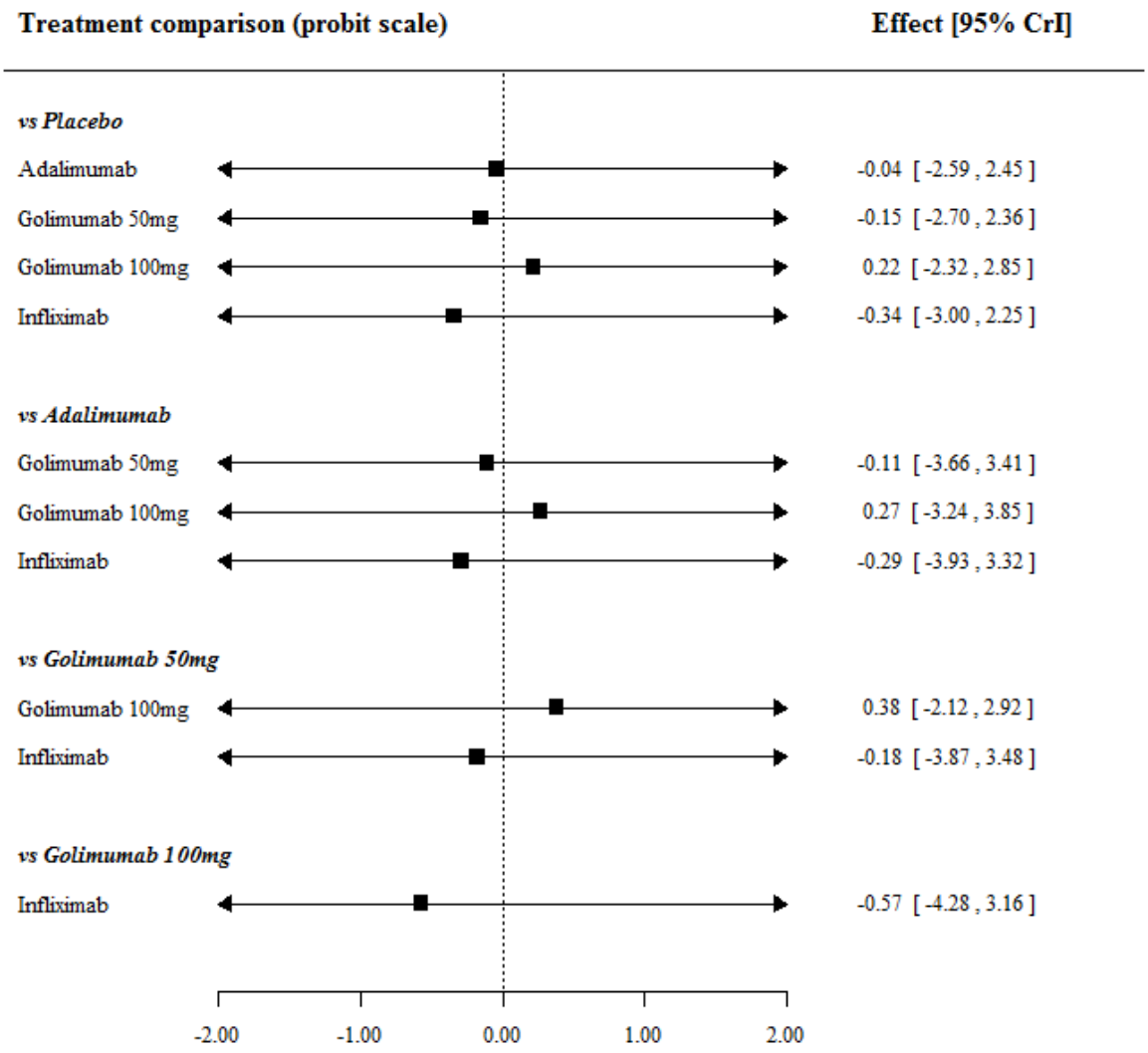
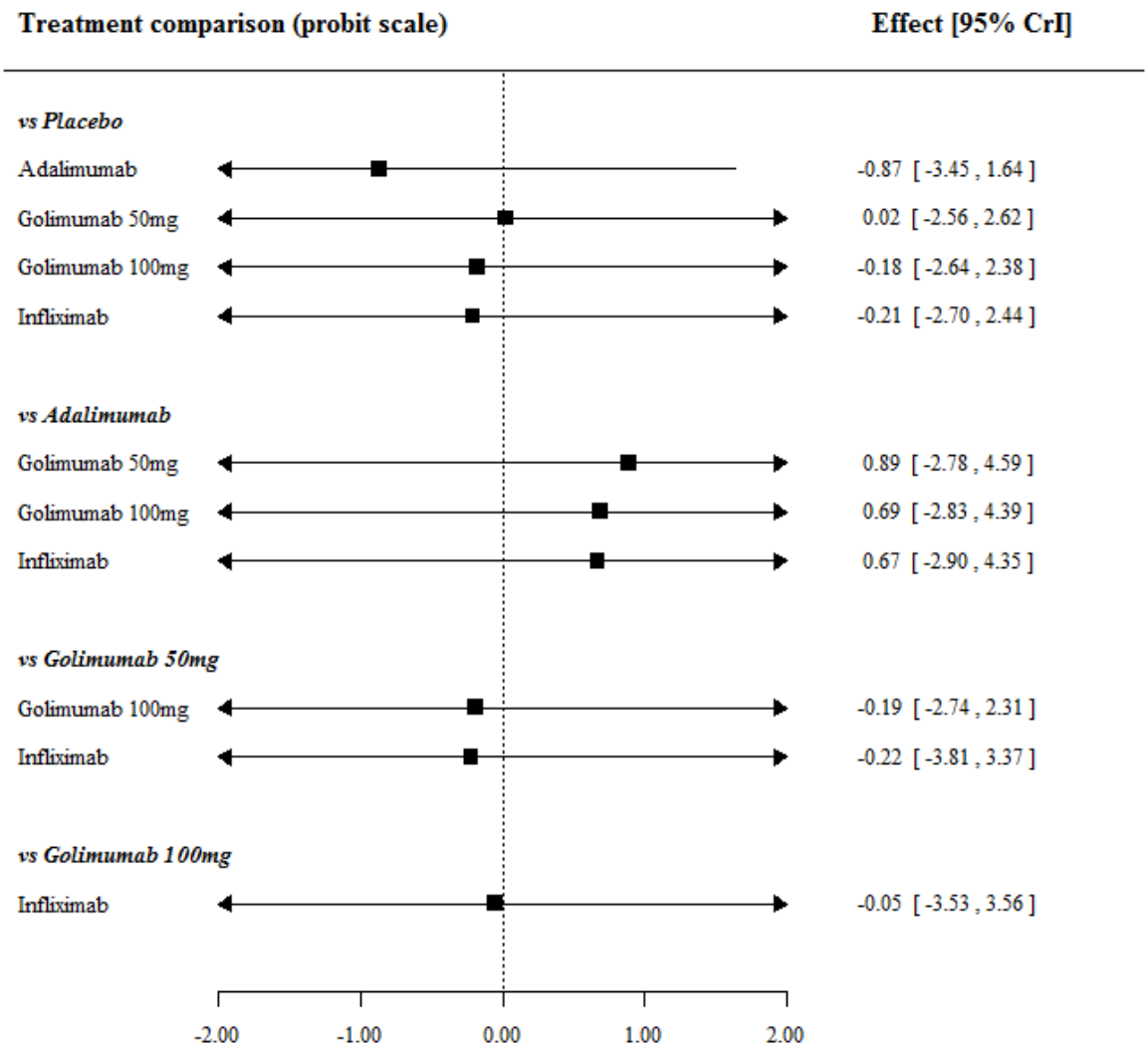


Figure A9.20: Sensitivity Analysis 3 – Comparative effect of anti-TNF-alpha treatment on clinical response/remission in the maintenance phase at 32-52 weeks for patients starting in remission (SD~U(0,2))



Appendix 10: Searches for cost-effectiveness searches

Database: Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R) <1946 to Present>
Search Strategy:
---

46. Colitis, Ulcerative/
47. ulcerative colitis.tw.
48. colitis ulcerosa.tw.
49. uc.tw.
50. colitis ulcerative.tw.
51. Colitis/
52. colitis.tw.
53. colitides.tw.
54. Inflammatory Bowel Diseases/
55. inflammatory bowel disease$.tw.
56. ibd.tw.
57. (col* and ulcer*).tw.
58. colitis gravis.tw.
59. proctocolitis.tw.
60. or/1-14
61. adalimumab.af.
62. humira.af.
63. d 2e7.af.
64. d2e7.af.
65. 331731-18-1.rn.
66. infliximab.af.
67. remicade.af.
68. 170277-31-3.rn.
69. ta650.af.
70. ta 650.af.
71. inx.af.
72. remsima.af.
73. inflectra.af.
74. ct p13.af.
75. ctp13.af.
76. golimumab.af.
77. simponi.af.
78. cnto148.af.
79. cnto 148.af.
80. 476181-74-5.rn.
81. tnf inhibitor$.tw.
82. anti tnf.tw.
83. antitnf.tw.
84. tnf antagonist$.tw.
85. tnf-alpha blocker$.tw.
86. antitumo?r necrosis factor.tw.
87. Biosimilar Pharmaceuticals/

88. (biosimilar$ or biologic$).tw.
89. or/16-43
90. 15 and 44

Terms 1-14 are terms for the condition (ulcerative colitis) which are then combined using OR in term 15. Terms 16-43 are terms for the interventions (infliximab, adalimumab and golimumab) which are then combined using OR in term 44. Terms 15 and 44 are then combined using AND to find studies on the condition and interventions in term 45.
To retrieve Economic evaluations specially designed highly sensitive search filter were combined with term 45. Economics filter below.
Economic filter

37. exp "costs and cost analysis"/
38. economics/
39. exp economics, hospital/
40. exp economics, medical/
41. economics, nursing/
42. exp models, economic/
43. economics, pharmaceutical/
44. exp "fees and charges"/
45. exp budgets/
46. budget$.tw.
47. ec.fs.
48. cost$.ti.
49. (cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab.
50. (economic$ or pharmacoeconomic$ or pharmaco-economic$).ti.
51. (price$ or pricing$).tw.
52. (financial or finance or finances or financed).tw.
53. (fee or fees).tw.
54. (value adj2 (money or monetary)).tw.
55. quality-adjusted life years/
56. (qaly or qalys).af.
57. (quality adjusted life year or quality adjusted life years).af.
58. or/1-22

Appendix 11: EQ-5D search

Database: Ovid MEDLINE(R), Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid OLDMEDLINE(R) <1946 to Present>
Search Strategy:
---

1. Colitis, Ulcerative/
2. ulcerative colitis.tw.
3. colitis ulcerosa.tw.
4. uc.tw.
5. colitis ulcerative.tw.
6. Colitis/
7. colitis.tw.
8. colitides.tw.
9. Inflammatory Bowel Diseases/
10. inflammatory bowel disease$.tw.
11. ibd.tw.
12. (col* and ulcer*).tw.
13. colitis gravis.tw.
14. proctocolitis.tw.
15. or/1-14
16. (euroqol or euro qol or eq5d or eq 5d).tw.
17. 15 and 16

11. REFERENCES

1. National Institute of Health and Care Excellence (NICE). Ulcerative colitis: Management in adults, children and young people. 2013; CG166. NICE.

2. Danese, S., Fiocchi, C. Ulcerative colitis. *New England Journal of Medicine* 2011; 365(18):03.

3. NICE. Ulcerative colitis: Management in adults, children and young people. 2013; CG166.

4. Ford, A.C., Moayyedi, P., Hanauer, S.B., Kirsner, J.B. Ulcerative colitis. *BMJ (Online)* 2013; 346(7894):f432.

5. Bhagat S, Das KM. A shared and unique epitope in the human colon, eye, and joint detected by a monoclonal antibody. *Gastroenterology* 1994; 107(103):108.

6. Dignass, A., Eliakim, R., Magro, F., Maaser, C., Chowers, Y., Geboes, K. et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. *J Crohns Colitis* 2012; 6(10):965-990.

7. Roth, L.S., Chande, N., Ponich, T., Roth, M.L., Gregor, J. Predictors of disease severity in ulcerative colitis patients from Southwestern Ontario. *World Journal of Gastroenterology* 2010; 16(2):232-236.

8. Lee, J.H., Cheon, J.H., Moon CM, Park, J.J., Hong SP, Kim TI et al. Do patients with ulcerative colitis diagnosed at a young age have more severe disease activity than patients diagnosed when older? *Digestion* 2010; 81:237-243.

9. Gibson, P.R., Vaizey, C., Black, C.M., Nicholls, R.J., Weston, A.R., Bampton, P. et al. Relationship between disease severity and quality of life and assessment of health care utilization and cost for ulcerative colitis in Australia: A cross-sectional, observational study. *J Crohns Colitis* 2013; http://dx.doi.org/10.1016/j.crohns.2013.11.017:Online ahead of print.

10. Ng, S.C., Kamm, M.A. Therapeutic strategies for the management of ulcerative colitis. *Inflammatory Bowel Diseases* 2009; 15(6):935-950.

11. Solberg IC, Lygren I, Jahnsen J, Aadland E, Høie O, Cvancarova M et al. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN study). *Scandinavian Journal of Gastroenterology - Supplement* 2009; 44(4):431-440.

12. Aratari, A., Papi, C., Clemente, V., Moretti, A., Luchetti, R., Koch, M. et al. Colectomy rate in acute severe ulcerative colitis in the infliximab era. *Digestive & Liver Disease* 2008; 40(10):821-826.

13. Hoie, O., Wolters, F., Riis, L., Aamodt, G., Solberg, C., Bernklev, T. et al. Ulcerative colitis: Patient characteristics may predict 10-yr disease recurrence in a European-wide population-based cohort. *American Journal of Gastroenterology* 2007; 102(8):1692-1701.

14. Gustavsson, A., Halfvarson, J., Magnuson, A., Sandberg-Gertzen, H., Jarnerot, G. Long-term colectomy rate after intensive intravenous corticosteroid therapy for ulcerative colitis prior to the immunosuppressive treatment era. *American Journal of Gastroenterology* 2007; 102(11):2513-2519.

15. Jess, T., Gamborg, M., Munkholm, P., Sorensen, T. Overall and cause-specific mortality in ulcerative colitis: meta-analysis of population-based inception cohort studies. *American Journal of Gastroenterology* 2007; 102[3], 609-617.

16. Casellas, F., Lopez, V.J., Casado A, Malagelada, J.R. Factors affecting health related quality of life of patients with inflammatory bowel disease. *Quality of Life Research* 2002; 11:775-781.

17. Lix LM, Graff LA, Walker JR, Clara I, Rawsthorne P, Rogala L et al. Longitudinal study of quality of life and psychological functioning for active, fluctuating, and inactive disease patterns in inflammatory bowel disease. *Inflammatory Bowel Disease: Genetics, Barrier Function, Immunologic Mechanisms, and Microbial Pathways* 2008; 14:1575-1584.

18. Waljee, A.K., Joyce J, Wren, P.A., Khan TM, Higgins, P.D. Patient reported symptoms during an ulcerative colitis flare: a Qualitative Focus Group Study. *European Journal of Gastroenterology & Hepatology* 2009; 21:558-564.

19. Kane, S.V. Systematic review: adherence issues in the treatment of ulcerative colitis. *Aliment Pharmacol Ther* 2006; 23:577-585.

20. Higgins, P.D.R., Rubin, D., Kaulback K, Schoenfield P, Kane, S.V. Systematic review: impact of non-adherence to 5-aminosalicylic acid products on the frequency and cost of ulcerative colitis flares. *Aliment Pharmacol Ther* 2009; 29:247-257.

21. Vaizey, C., Gibson, P., Black, C.M., et al. Disease status, patient quality of life, and health care resource use for ulcerative colitis in the United Kingdom: an observational study . *Frontline Gastroenterology* 2014; Published Online First: doi:10.1136/flgastro-2013-100409 http://fg.bmj.com/content/early/2014/01/24/flgastro-2013-100409.full.pdf .

22. Bassi, A., Dodd, S., Williamson, P., Bodger, K. Cost of illness of inflammatory bowel disease in the UK: a single centre retrospective study. *Gut* 2004; 53(10):1471-1478.

23. Cooney, R.M., Warren, B.F., Altman, D.G., Abreu, M.T., Travis, S.P.L. Outcome measurement in clinical trials for Ulcerative Colitis: towards standardisation. *Trials* 2007; 8.

24. Truelove, S.C., Witts, L.J. Cortisone in ulcerative colitis; final report on a therapeutic trial. *British Medical Journal* 1955; 2(4947):1041-1048.

25. Schroeder, K., Tremaine, W., Ilstrup, D. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis-a randomized study. *New England Journal of Medicine* 1987; 317:1625-1629.

26. Turner, D., Otley, A.R., Mack, D., Hyams, J., de, B.J., Uusoue, K. et al. Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study. *Gastroenterology* 2007; 133(2):423-432.

27. Ha, C. Clinical Measures of Disease Activity. In: Bayless T.M., Hanauer S.B., eds. *Advanced Therapy in Inflammatory Bwel Disease*. People's Medical Publishing House - USA; Shelton, Connecticut, USA: 2011; 125-130.

28. Lewis, J.D., Chuai, S., Nessel, L., Lichtenstein, G.R., Aberra, F.N., Ellenberg, J.H. Use of the Noninvasive Components of the Mayo Score to Assess Clinical Response in Ulcerative Colitis. *Inflammatory Bowel Diseases* 2008; 14(12):1660-1666.

29. NICE. Infliximab for subacute manifestations of ulcerative colitis. 2011; TA140.

30. NICE. Adalimumab for the treatment of moderate to severe ulcerative colitis (terminated appraisal). 2013; TA262.

31. NICE. Ulcerative colitis (acute exacerbations) - infliximab: guidance. 2008; TA163.

32. Cohen, R.D., Yu, A.P., Wu, E.Q., Xie, J., Mulani, P.M., Chao, J. Systematic review: the costs of ulcerative colitis in Western countries. *Aliment Pharmacol Ther* 2010; 31(7):693-707.

33. Shih, D.Q., Nguyen, M., Zheng, L., Ibanez, P., Mei, L., Kwan, L.Y. et al. Split-dose administration of thiopurine drugs: a novel and effective strategy for managing preferential 6-MMP metabolism. *Alimentary Pharmacology & Therapeutics* 2012; 36(5):449-458.

34. EMC. Summary of Product Characteristics: Remicade. 2013; available from https://www.medicines.org.uk

35. EMC. Summary of Product Characteristics: Humira. 2014; available from https://www.medicines.org.uk/

36. EMC. Summary of Product Characteristics: Simponi. 2013; available from https://www.medicines.org.uk

37. NICE. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262) - Final Scope. 2013.

38. British National Formulary (BNF). 2014; available from http://www.bnf.org/bnf/index.htm

39. PRISMA. Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Statement. 2009; available from http://www.prisma-statement.org/index.htm

40. Archer, R.J., et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy: protocol record CRD42013006883. 2013; available from http://www.crd.york.ac.uk/PROSPERO

41. National Institute for Health and Care Excellence (NICE). Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): Final Scope. 2013;

42. Engauge (v4.1) 2011. Available from http://sourceforge.net/projects/digitizer/files/Engauge%20Digitizer/digitizer-4.1/ [date accessed 10/03/2014]

43. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons Ltd, England; 2008.

44. Wright, C.C., Sim, J. Intention-to-treat approach to data from randomized controlled trials: a sensitivity analysis. *Journal of Clinical Epidemiology* 2003; 56(9):833-842.

45. Reinisch, W., Sandborn, W.J., Hommes, D.W., D'Haens, G., Hanauer, S., Schreiber, S. et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. *Gut* 2011; 60(6):780-787.

46. Sandborn, W.J., Van, A.G., Reinisch, W., Colombel, J.F., D'Haens, G., Wolf, D.C. et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2012; 142(2):257-265.

47. Suzuki, Y., Motoya, S., Hanai, H., Matsumoto, T., Hibi, T., Robinson, A. et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. *Journal of Gastroenterology* 2014; 49(2):283-294.

48. Sandborn, W.J., Feagan, B.G., Marano, C., Zhang, H., Strauss, R., Johanns, J. et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2014; 146(1):85-95.

49. Sandborn, W.J., Feagan, B.G., Marano, C., Zhang, H., Strauss, R., Johanns, J. et al. Subcutaneous Golimumab Maintains Clinical Response in Patients With Moderate-to-Severe Ulcerative Colitis. *Gastroenterology* 2014; 146[1], 96-109.

50. Rutgeerts, P., Sandborn, W.J., Feagan, B.G., Reinisch, W., Olson, A., Johanns, J. et al. Infliximab for induction and maintenance therapy for ulcerative colitis.[Erratum appears in N Engl J Med. 2006 May 18;354(20):2200]. *New England Journal of Medicine* 2005; 353(23):2462-2476.

51. Probert, C.S., Hearing, S.D., Schreiber, S., Kuhbacher, T., Ghosh, S., Arnott, I.D. et al. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: a randomised controlled trial. *Gut* 2003; 52(7):998-1002.

52. Panaccione, R., Ghosh, S., Middleton, S., M+írquez, J.R., Scott, B.B., Flint, L. et al. Combination Therapy With Infliximab and Azathioprine Is Superior to Monotherapy With Either Agent in Ulcerative Colitis. *Gastroenterology* 2014; 146[2], 392-400.

53. Hyams, J., Damaraju, L., Blank, M., Johanns, J., Guzzo, C., Winter, H.S. et al. Induction and maintenance therapy with infliximab for children with moderate to severe ulcerative colitis. *Clinical Gastroenterology & Hepatology* 2012; 10(4):391-399.

54. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6(7):e1000097.

55. Reinisch, W., Sandborn, W.J., Rutgeerts, P., Feagan, B.G., Rachmilewitz, D., Hanauer, S.B. et al. Long-term infliximab maintenance therapy for ulcerative colitis:

the ACT-1 and -2 extension studies. *Inflammatory Bowel Diseases* 2012; 18(2):201-211.

56. Reinisch, W., Sandborn, W.J., Panaccione, R., Huang, B., Pollack, P.F., Lazar, A. et al. 52-week efficacy of adalimumab in patients with moderately to severely active ulcerative colitis who failed corticosteroids and/or immunosuppressants. *Inflammatory Bowel Diseases* 2013; 19(8):1700-1709.

57. Committee for Medicinal Products for Human Use (CHMP), E.M.A.E. Guideline on the development of new medicinal products for the treatment of ulcerative colitis. 2008. London, UK.

58. Ghosh, S., Wolf, D.C., Sandborn, W., Colombel, J.-F., Zhou, Q., Lazar, A. et al. Sustained efficacy in patients with ulcerative colitis treated with adalimumab: Results from ultra 2. *Gastroenterology Conference: Digestive Disease Week 2013, DDW 2013 Orlando, FL United States Conference Start: 20130518 Conference End: 20130521 Conference Publication: (Var Pagings)* 2013; 144(5 SUPPL.#1):May.

59. Van, A.G., Wolf, D., D'Haens, G., Sandborn, W., Colombel, J.-F., Lazar, A. et al. Reduced steroid usage in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ultra 2. *Inflammatory Bowel Diseases Conference: 2012 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20121213 Conference End: 20121215 Conference P* 2012; 18(pp S21-S22):December.

60. Sandborn, W., Thakkar, R., Lazar, A., Skup, M., Yang, M., Chao, J. et al. Time-in-remission analysis of adalimumab vs. placebo for the treatment of ulcerative colitis. *American Journal of Gastroenterology Conference: 77th Annual Scientific Meeting of the American College of Gastroenterology Las Vegas, NV United States Conference Start: 20121019 Conference End: 20121024 Conference Publication: (Var Pagings)* 2012; 107(pp S641-S642):October.

61. Sandborn, W.J., Van, A.G., Thakkar, R.B., Lazar, A., Kron, M., Yang, M. et al. Adalimumab improves health-related quality of life for 52 weeks in patients with ulcerative colitis. *Gut Conference: Digestive Disorders Federation Meeting 2012 Liverpool United Kingdom Conference Start: 20120617 Conference End: 20120620 Conference Publication: (Var Pagings)* 2012; 61(pp A400-A401):July.

62. Reinisch, W., Sandborn, W., Hommes, D., Thakkar, R., Pollack, P., Kumar, A. et al. Effect of adalimumab induction therapy on clinical laboratory parameters suggesting improved nutrition and inflammation status in patients with moderately to severely active ulcerative colitis. *American Journal of Gastroenterology Conference: 75th Annual Scientific Meeting of the American College of Gastroenterology San Antonio, TX United States Conference Start: 20101015 Conference End: 20101020 Conference Publication: (Var Pagings)* 2010; 105(pp S428-S429):October.

63. Abbvie. Adalimumab, golimumab and infliximab, for the treatment of ulcerative colitis (subacute). Submission to National Institute for Health and Care Excellence. 2014.

64. Feagan, B., Sandborn, W., Yang, M., Skup, M., Thakkar, R., Lazar, A. et al. Adalimumab therapy reduces hospitalization and colectomy rates in patients with ulcerative colitis among initial responders. *American Journal of Gastroenterology Conference: 77th Annual Scientific Meeting of the American College of*

Gastroenterology Las Vegas, NV United States Conference Start: 20121019 Conference End: 20121024 Conference Publication: (Var Pagings) 2012; 107(pp S642-S643):October.

65. Merck Sharp Dohme (MSD). Manufacturer submission of evidence: Infliximab (Remicade). 2014.

66. Feagan, B.G., Sandborn, W.J., Lazar, A., Thakkar, R.B., Huang, B., Reilly, N. et al. Adalimumab therapy is associated with reduced risk of hospitalization in patients with ulcerative colitis. *Gastroenterology* 2014; 146(1):January.

67. Merck Sharp Dohme. Manufacturer submission of evidence: Golimumab (Simponi). 2014.

68. Sandborn, W.J., Rutgeerts, P., Feagan, B.G., Reinisch, W., Olson, A., Johanns, J. et al. Colectomy rate comparison after treatment of ulcerative colitis with placebo or infliximab. *Gastroenterology* 2009; 137(4):1250-1260.

69. Centocor Research & Development, I. Briefing document for Food and Drug Administration Gastrointestinal Drugs Advisory Committee. Remicade (infliximab) for paediatric ulcerative colitis. 2011.

70. Dias, S., Welton, N., Sutton, A., Ades, A. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011.

71. European Medicines Agency. Questions and answers on biosimilar medicines (similar biological medicinal products). 2012. London, UK, European Medicines Agency.

72. Celltrion Healthcare. Infliximab biosimilar (Remsima®). Submission to NICE. 2014.

73. European Medicines Agency. Assessment Report: Remsima. 2013. London, UK.

74. European Medicines Agency. Assessment Report: Inflectra. 2013. London, UK.

75. Danese, S., Gomollon, F. ECCO position statement: the use of biosimilar medicines in the treatment of inflammatory bowel disease (IBD). *J Crohns Colitis* 2013; 7(7):586-589.

76. Kurki, P., Bielsky, M.C. ECCO position challenged by European drug regulators. *J Crohns Colitis* 2014; 8(3):258.

77. Drummond, M.E., Sculpher, M.J., Torrance, G.W., O'Brien, B.J., Stoddart, G.L. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. Oxford University Press, 2005.

78. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. 2013; available from http://www.nice.org.uk/media/D45/1E/GuideToMethodsTechnologyAppraisal2013.pdf

79. National Institute for Health and Care Excellence (NICE). Infliximab for subacute manifestations of ulcerative colitis. *NICE Technology Appraisal Guidance 140* 2014.

80. Schering Plough Limited. Remicade in the treatment of ankylosing spondylitis in England and Wales. A Submission to the National Institute for Health and Clinical Excellence. 2006.

81. Hyde, C., Bryan, S., Juarez-Garcia, A., Andronis, L., Fry-Smith, A. Infliximab for the treatment of ulcerative colitis. [Review] [2 refs]. *Health Technology Assessment (Winchester, England)* 2009; 13:Suppl-11.

82. Tsai, H.H., Punekar, Y.S., Morris, J., Fortun, P. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 2008; 28(10):1230-1239.

83. Park, K.T., Tsai, R., Perez, F., Cipriano, L.E., Bass, D., Garber, A.M. Cost-effectiveness of early colectomy with ileal pouch-anal anastamosis versus standard medical therapy in severe ulcerative colitis. *Annals of Surgery* 2012; 256(1):117-124.

84. Xie, F., Blackhouse, G., Assasi, N., Gaebel, K., Robertson, D., Goeree, R. Cost-utility analysis of infliximab and adalimumab for refractory ulcerative colitis. *Cost Effectiveness & Resource Allocation* 2009; 7:20.

85. Gold, M.R., Siegel, J.E., Russell, L.B., Weinstein, M.C. Cost-Effectiveness in Health and Medicine. Oxford University Press, New York; 1996.

86. Walmsley, R.S., Ayres, R.C., Pounder, R.E., Allan, R.N. A simple clinical colitis activity index. *Gut* 1998; 43(1):29-32.

87. Barabino, A., Tegaldo, L., Castellano, E., Gandullia, P., Torrente, F., Marino, C. et al. Severe attack of ulcerative colitis in children: retrospective clinical survey. *Dig Liver Dis* 2002; 34(1):44-49.

88. Hanauer, S., Schwartz, J., Robinson, M., Roufail, W., Arora, S., Cello, J. et al. Mesalamine capsules for treatment of active ulcerative colitis: results of a controlled trial. Pentasa Study Group. *Am J Gastroenterol* 1993; 88(8):1188-1197.

89. Dignass, A.U., Bokemeyer, B., Adamek, H., Mross, M., Vinter-Jensen, L., Borner, N. et al. Mesalamine once daily is more effective than twice daily in patients with quiescent ulcerative colitis. *Clin Gastroenterol Hepatol* 2009; 7(7):762-769.

90. Kamm, M.A., Lichtenstein, G.R., Sandborn, W.J., Schreiber, S., Lees, K., Barrett, K. et al. Randomised trial of once- or twice-daily MMX mesalazine for maintenance of remission in ulcerative colitis. *Gut* 2008; 57(7):893-902.

91. Ardizzone, S., Maconi, G., Russo, A., Imbesi, V., Colombo, E., Porro, G.B. Randomised controlled trial of azathioprine and 5-aminosalicylic acid for treatment of steroid dependent ulcerative colitis. *Gut* 2006; 55(1):47-53.

92. Hawthorne, A.B., Logan, R.F.A., Hawkey, C.J., Foster, P.N., Axon, A.T.R., Swarbrick, E.T. et al. Randomised Controlled Trial Of Azathioprine Withdrawal In Ulcerative Colitis. *BMJ* 1992; 305(20):22.

93. Chebli, L.A., Chaves, L.D., Pimentel, F.F., Guerra, D.M., Barros, R.M., Gaburri, P.D. et al. Azathioprine maintains long-term steroid-free remission through 3 years in patients with steroid-dependent ulcerative colitis. *Inflamm Bowel Dis* 2010; 16(4):613-619.

94. Watson, S., Pensabene, L., Mitchell, P., Bousvaros, A. Outcomes and Adverse Events in Children and Young Adults Undergoing Tacrolimus Therapy for Steroid-refractory Colitis. *Inflammatory Bowel Diseases* 2011; 17(1):22-29.

95. Chew, S.S., Kerdic, R.I., Yang, J.L., Shi, E.C., Newstead, G.L., Douglas, P.R. Functional outcome and quality of life after ileal pouch-anal anastomosis in children and adults. *ANZ J Surg* 2003; 73(12):983-987.

96. Hahnloser, D., Pemberton, J.H., Wolff, B.G., Larson, D.R., Crownhart, B.S., Dozois, R.R. Results at up to 20 years after ileal pouch-anal anastomosis for chronic ulcerative colitis. *Br J Surg* 2007; 94(3):333-340.

97. Waljee, A., Waljee, J., Morris, A.M., Higgins, P.D.R. Threefold increased risk of infertility: a meta-analysis of infertility after ileal pouch anal anastomosis in ulcerative colitis. *Gut* 2006; 55(11):1575-1580.

98. Stahlberg, D., Gullberg, K., Liljeqvist, L., Hellers, G., Lofberg, R. Pouchitis following pelvic pouch operation for ulcerative colitis. Incidence, cumulative risk, and risk factors. *Dis Colon Rectum* 1996; 39(9):1012-1018.

99. Mimura, T., Rizzello, F., Helwig, U., Poggioli, G., Schreiber, S., Talbot, I.C. et al. Four-week open-label trial of metronidazole and ciprofloxacin for the treatment of recurrent or refractory pouchitis. *Aliment Pharmacol Ther* 2002; 16(5):909-917.

100. Soderlund, S., Brandt, L., Lapidus, A., Karlen, P., Brostrom, O., Lofberg, R. et al. Decreasing time-trends of colorectal cancer in a large cohort of patients with inflammatory bowel disease. *Gastroenterology* 2009; 136(5):1561-1567.

101. Tengs, T.O., Wallace, A. One thousand health-related quality-of-life estimates. *Med Care* 2000; 38(6):583-637.

102. Muir, A.J., Edwards, L.J., Sanders, L.L., Bollinger, R.R., Koruda, M.J., Bachwich, D.R. et al. A prospective evaluation of health-related quality of life after ileal pouch anal anastomosis for ulcerative colitis. *American Journal of Gastroenterology* 2001; 96(5):1480-1485.

103. Provenzale, D. Cost-effectiveness of screening the average-risk population for colorectal cancer. *Gastrointest Endosc Clin N Am* 2002; 12(1):93-109.

104. Gift, T.L., Owens, C.J. The direct medical cost of epididymitis and orchitis: evidence from a study of insurance claims. *Sex Transm Dis* 2006; 33(10 Suppl):S84-S88.

105. Bayless, T.M., Hanauer, S.B. Advanced Therapy of Inflammatory Bowel Disease: Ulcerative Colitis. PMPH-USA, 2011.

106. Penna, C., Dozois, R., Tremaine, W., Sandborn, W., LaRusso, N., Schleck, C. et al. Pouchitis after ileal pouch-anal anastomosis for ulcerative colitis occurs with increased frequency in patients with associated primary sclerosing cholangitis. *Gut* 1996; 38(2):234-239.

107. Hoie, O., Wolters, F.L., Riis, L., Bernklev, T., Aamodt, G., Clofent, J. et al. Low colectomy rates in ulcerative colitis in an unselected European cohort followed for 10 years. *Gastroenterology* 2007; 132(2):507-515.

108. Arai, K., Koganei, K., Kimura, H., Akatani, M., Kitoh, F., Sugita, A. et al. Incidence and outcome of complications following restorative proctocolectomy. *Am J Surg* 2005; 190(1):39-42.

109. Woehl, A., Hawthorne, A.B., Morgan, C.L., Punekar, Y., McEwan, P. The epidemiology and healthcare resource use in patients with Crohn's disease: a population based UK study. Value Health 2007; 10: A355.

110. Arseneau, K.O., Sultan, S., Provenzale, D.T., Onken, J., Bickston, S.J., Foley, E. et al. Do patient preferences influence decisions on treatment for patients with steroid-refractory ulcerative colitis? *Clinical Gastroenterology and Hepatology* 2006; 4(9):1135-1142.

111. Jarnerot, G., Hertervig, E., Friis-Liby, I., Blomquist, L., Karlen, P., Granno, C. et al. Infliximab as rescue therapy in severe to moderately severe ulcerative colitis: a randomized, placebo-controlled study. *Gastroenterology* 2005; 128(7):1805-1811.

112. Fazio, V.W., Ziv, Y., Church, J.M., Oakley, J.R., Lavery, I.C., Milsom, J.W. et al. Ileal pouch-anal anastomoses complications and function in 1005 patients. *Ann Surg* 1995; 222(2):120-127.

113. Raval, M.J., Schnitzler, M., O'Connor, B.I., Cohen, Z., McLeod, R.S. Improved outcome due to increased experience and individualized management of leaks after ileal pouch-anal anastomosis. *Ann Surg* 2007; 246(5):763-770.

114. Swenson, B.R., Hollenbeak, C.S., Koltun, W.A. Factors affecting cost and length of stay associated with the ileal pouch-anal anastomosis. *Dis Colon Rectum* 2003; 46(6):754-761.

115. Chaudhary, M.A., Fan, T. Cost-Effectiveness of Infliximab for the Treatment of Acute Exacerbations of Ulcerative Colitis in the Netherlands. *Biologics in Therapy* 2012; 3(1):June.

116. Brereton, N., Bodger, K., Kamm, M.A., Hodgkins, P., Yan, S., Akehurst, R. A cost-effectiveness analysis of MMX mesalazine compared with mesalazine in the treatment of mild-to-moderate ulcerative colitis from a UK perspective. *J Med Econ* 2010; 13(1):148-161.

117. Connolly, M.P., Nielsen, S.K., Currie, C.J., Poole, C.D., Travis, S.P.L. An economic evaluation comparing once daily with twice daily mesalazine for maintaining remission based on results from a randomised controlled clinical trial. *Journal of Crohn's and Colitis* 2009; 3(1):February.

118. Saini, S.D., Waljee, A.K., Higgins, P.D.R. Cost Utility of Inflammation-Targeted Therapy for Patients With Ulcerative Colitis. *Clinical Gastroenterology and Hepatology* 2012; 10(10):1143-1151.

119. Punekar, Y.S., Hawkins, N. Cost-effectiveness of infliximab for the treatment of acute exacerbations of ulcerative colitis. *European Journal of Health Economics* 2010; 11(1):67-76.

120. Odes, S., Vardi, H., Friger, M., Esser, D., Wolters, F., Moum, B. et al. Clinical and economic outcomes in a population-based European cohort of 948 ulcerative colitis and Crohn's disease patients by Markov analysis. *Aliment Pharmacol Ther* 2010; 31(7):735-744.

121. Buckland, A., Bodger, K. The cost-utility of high dose oral mesalazine for moderately active ulcerative colitis. *Aliment Pharmacol Ther* 2008; 28(11-12):1287-1296.

122. Connolly, M.P., Boersma, C., Oldenburg, B. The economics of mesalazine in active ulcerative colitis and maintenance in the Netherlands. *Netherlands Journal of Medicine* 2013; 70(6):August.

123. Prenzler, A., Yen, L., Mittendorf, T., von der Schulenburg, J.M. Cost effectiveness of ulcerative colitis treatment in Germany: a comparison of two oral formulations of mesalazine. *BMC Health Serv Res* 2011; 11:157.

124. Yen, E.F., Kane, S.V., Ladabaum, U. Cost-effectiveness of 5-aminosalicylic acid therapy for maintenance of remission in ulcerative colitis. *American Journal of Gastroenterology* 2008; 103(12):3094-3105.

125. Reinisch, W.w.r.a.a., Sandborn, W.J., Hommes, D.W., D'Haens, G., Hanauer, S., Schreiber, S. et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. *Gut* 2011; 60(6):780-787.

126. Sandborn, W.J., Van Assche, G., Reinisch, W., Colombel, J.F., D'Haens, G., Wolf, D.C. et al. Adalimumab Induces and Maintains Clinical Remission in Patients With Moderate-to-Severe Ulcerative Colitis. *Gastroenterology* 2012; 142(2):257.

127. Turner, D., Walsh, C.M., Steinhart, A.H., Griffiths, A.M. Response to Corticosteroids in Severe Ulcerative Colitis: A Systematic Review of the Literature and a Meta-Regression. *Clinical Gastroenterology and Hepatology* 2007; 5(1):January.

128. Grijalva, C.G., Chen, L., Delzell, E., Baddley, J.W., Beukelman, T., Winthrop, K.L. et al. Initiation of tumor necrosis factor- antagonists and the risk of hospitalization for infection in patients with autoimmune diseases. *JAMA* 2011; 306(21):2331-2339.

129. Curtis, L. Unit costs of health and social care 2012. 2012. Canterbury, PSSRU. University of Kent.

130. Department of Health. NHS reference costs 2011-2012. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012

131. Department of Health. NHS reference costs 2012-2013. *https://www gov uk/government/uploads/system/uploads/attachment_data/file/261154/nhs_reference_costs_2012-13_acc pdf* 2013; Available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/261154/nhs_reference_costs_2012-13_acc.pdf

132. Turner RM, Spiegelhalter, D. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society* 2009; 172(1):21-47.

133. European Medicines Agency. Humira: EPAR product information. 2012; available from http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000481/WC500050870.pdf

134. Kane, S., Huo, D., Aikens, J., Hanauer, S. Medication nonadherence and the outcomes of patients with quiescent ulcerative colitis. *Am J Med* 2003; 114(1):39-43.

135. Hillson, E., Dybicz, S., Waters, H.C., Stuart, B., Schaneman, J., Dabbous, O. et al. Health care expenditures in ulcerative colitis: the perspective of a self-insured employer. *J Occup Environ Med* 2008; 50(8):969-977.

136. Johnson, P., Richard, C., Ravid, A., Spencer, L., Pinto, E., Hanna, M. et al. Female infertility after ileal pouch-anal anastomosis for ulcerative colitis. *Dis Colon Rectum* 2004; 47(7):1119-1126.

137. Krausz, M.M., Duek, S.D. Restorative proctocolectomy with ileal pouch-anal anastomosis for ulcerative colitis and familial adenomatous polyposis: twenty years follow-up in 174 patients. *Isr Med Assoc J* 2005; 7(1):23-27.

138. Abdelrazeq, A.S., Kandiyil, N., Botterill, I.D., Lund, J.N., Reynolds, J.R., Holdsworth, P.J. et al. Predictors for acute and chronic pouchitis following restorative proctocolectomy for ulcerative colitis. *Colorectal Dis* 2008; 10(8):805-813.

139. Roberts, S.E., Williams, J.G., Yeates, D., Goldacre, M.J. Mortality in patients with and without colectomy admitted to hospital for ulcerative colitis and Crohn's disease: record linkage studies. *BMJ* 2007; 335(7628):1033.

140. Swinburn, P., Elwick, H., Bean, K., Curry, A., Patel, S., Bodger, K. et al. The impact of surgery on health related quality of life in ulcerative colitis. *Gut Conference: Digestive Disorders Federation Meeting 2012 Liverpool United Kingdom Conference Start: 20120617 Conference End: 20120620 Conference Publication: (Var Pagings)* 2012; 61(Suppl. 2):A237.

141. Hu, D., Grossman, D., Levin, C., Blanchard, K., Goldie, S.J. Cost-effectiveness analysis of alternative first-trimester pregnancy termination strategies in Mexico City. *BJOG* 2009; 116(6):768-779.

142. Smith, K.J., Roberts, M.S. Quality-of-Life Utility Values for Erectile Function and Sildenafil Treatment. *Clin Drug Investig* 2005; 25(2):99-105.

143. Buchanan, J., Wordsworth, S., Ahmad, T., Perrin, A., Vermeire, S., Sans, M. et al. Managing the long term care of inflammatory bowel disease patients: The cost to European health care providers. *Journal of Crohns & Colitis* 2011; 5(4):301-316.

144. Craig, B.A., Sendi, P.P. Estimation of the transition matrix of a discrete-time Markov chain. *Health Econ* 2002; 11(1):33-42.

145. Lynch, R.W., Down, C., Roughton, M. No increase in surgical complication in patients treated with rescue therapy for acute severe ulcerative colitis: data from the UK IBD audit. *Gut* 20112; 61(Suppl 2):A71.

146. Office for National Statistics. Interim life tables, England and Wales, 2009-11. 2011; available from http://www.ons.gov.uk/ons/rel/lifetables/interim-life-tables/2009-2011/stb-2009-2011.html

147. European Medicines Agency. Remicade. EPAR Product Information. 2014; available from http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000240/WC500050888.pdf

148. EMA. Assessment Report For Simponi. 2014; available from https://www.google.co.uk/search?q=%3A+http%3A%2F%2Fwww.ema.europa.eu%2Fdocs%2Fen_GB%2Fdocument_library%2FEPAR_-

_Public_assessment_report%2Fhuman%2F000992%2FWC500052372.pdf&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a&channel=fflb&gfe_rd=cr&ei=gTh_U6_HDojY8gfE1oC4Cg

149. Actis, G.C., Fadda, M., David, E., Sapino, A. Colectomy rate in steroid-refractory colitis initially responsive to cyclosporin: a long-term retrospective cohort study. *BMC Gastroenterology* 2007; 7.

150. Gower-Rousseau, C., Dauchet, L., Vernier-Massouille, G., Tilloy, E., Brazier, F., Merle, V. et al. The natural history of pediatric ulcerative colitis: a population-based cohort study. *Am J Gastroenterol* 2009; 104(8):2080-2088.

151. Molnar, T., Farkas, K., Nyari, T., Szepes, Z., Nagy, F., Wittmann, T. Response to first intravenous steroid therapy determines the subsequent risk of colectomy in ulcerative colitis patients. *J Gastrointestin Liver Dis* 2011; 20(4):359-363.

152. Mocciaro, F., Renna, S., Orlando, A., Rizzuto, G., Sinagra, E., Orlando, E. et al. Cyclosporine or infliximab as rescue therapy in severe refractory ulcerative colitis: early and long-term data from a retrospective observational study. *Journal of Crohn's & Colitis* 2012; 6(6):681-686.

153. Briggs, A.H., Claxton, K., Sculpher, M.J. Decision Modelling for Health Economic Evaluation. Oxford University Press, New York; 2006.

154. Casellas, F., Lopez, V.J., Malagela, J.R. Previous experience and quiality of life in patients with inflammatory bowel disease during relapse. *Revista Espanola De Enfermedades Digestivas* 471; 95(7):476-479.

155. Leidl, R., Reitmeir, P., Konig, H.H., Stark, R. The performance of a value set for the EQ-5D based on experienced health states in patients with inflammatory bowel disease. *Value in Health* 2012; 15(1):151-157.

156. Vaizey, C., Gibson, P.R., Black, C.M., Nicholls, R.J., Weston, A.R., Gaya, D.R. et al. The relationship between disease severity, quality of life and health care resource utilization among United Kingdom patients with ulcerative colitis. *Journal of Crohn's and Colitis Conference: 8th Congress of the European Crohn's and Colitis Organisation, ECCO 2013 Vienna Austria Conference Start: 20130214 Conference End: 20130216 Conference Publication: (Var Pagings)* 2013; 7(pp S259):February.

157. van der Valk, M.E., Mangen, M.J., Dijkstra, G., van Bodegraven, A.A., Fidder, H., de Jong, D.J. et al. Is there a difference in quality of life and costs between ulcerative colitis patients with a pouch or an ileostomy? *Gastroenterology Conference: Digestive Diease Week 2012, DDW 2012 San Diego, CA United States Conference Start: 20120519 Conference End: 20120522 Conference Publication: (Var Pagings)* 2012; 142(5 SUPPL.#1):S263.

158. Richards, D.M., Hughes, S.A., Irving, M.H., Scott, N.A. Patient quality of life after successful restorative proctocolectomy is normal. *Colorectal Disease* 2001; 3(4):223-226.

159. Kuruvilla, K., Osler, T., Hyman, N.H. A comparison of the quality of life of ulcerative colitis patients after IPAA vs ileostomy. *Diseases of the Colon & Rectum* 2012; 55(11):1131-1137.

160. Jose Casanova, M., Chaparro, M., Iglesias Flores, E., Rodrigo, L., Domenech, E., Calvet, X. et al. Safety of Immunomodulators and Anti-TNF Drugs for the Treatment of Inflammatory Bowel Disease (IBD) During Pregnancy. *Gastroenterology* 2011; 140(5, Suppl. 1):S266.

161. Arseneau, K.O., Sultan, S., Provenzale, D.T., Onken, J., Bickston, S.J., Foley, E. et al. Do patient preferences influence decisions on treatment for patients with steroid-refractory ulcerative colitis? *Clinical Gastroenterology & Hepatology* 2006; 4(9):1135-1142.

162. Woehl, A., Hawthorne, A., McEwan, P. The relation between disease activity, quality of life and health utility in patients with ulcerative colitis. *Gut* 2008; 57(Suppl. 1):A153.

163. Actis, G.C., Bruno, M., Pinna-Pintor, M., Rossini, F.P., Rizzetto, M. Infliximab for treatment of steroid-refractory ulcerative colitis. *Digestive & Liver Disease* 2002; 34(9):631-634.

164. Actis, G.C. Infliximab for ulcerative colitis. *American Journal of Gastroenterology* 2003; 98(3):709.

165. Afif, W., Leighton, J.A., Hanauer, S.B., Loftus, E.V., Jr., Faubion, W.A., Pardi, D.S. et al. Open-label study of adalimumab in patients with ulcerative colitis including those with prior loss of response or intolerance to infliximab. *Inflammatory Bowel Diseases* 2009; 15(9):1302-1307.

166. Allez, M., Karmiris, K., Louis, E., Van, A.G., Ben-Horin, S., Klein, A. et al. Report of the ECCO pathogenesis workshop on anti-TNF therapy failures in inflammatory bowel diseases: definitions, frequency and pharmacological aspects. *Journal of Crohn's & Colitis* 2010; 4(4):355-366.

167. Infliximab (ulcerative colitis): Caution is needed due to long-term risks. *Prescrire International* 2007; 16(91):October.

168. Armuzzi, A., De, P.B., Lupascu, A., Fedeli, P., Leo, D., Mentella, M.C. et al. Infliximab in the treatment of steroid-dependent ulcerative colitis. *European Review for Medical & Pharmacological Sciences* 2004; 8(5):231-233.

169. Baert, F., de, V.M., Louis, E., Vermeire, S. Immunogenicity of infliximab: How to handle the problem? *Acta Gastro-Enterologica Belgica* 2007; 70(2):April/June.

170. Barbato, M., Curione, M., Viola, F., Versacci, P., Parisi, F., Amato, S. et al. Cardiac involvement in children with IBD during infliximab therapy. *Inflammatory Bowel Diseases* 2006; 12(8):828-829.

171. Barreiro-de, A.M., Lorenzo, A., Mera, J., Dominguez-Munoz, J.E. Mucosal healing and steroid-sparing associated with infliximab for steroid-dependent ulcerative colitis. *Journal of Crohn's and Colitis* 2009; 3(4):December.

172. Baumgart, D.C. How many lives does an ulcerative colitis patient have? *Lancet* 2010; 376(9744):928.

173. Ben-Horin, S. Randomised placebo-controlled trial of rituximab (anti-CD20) in active ulcerative colitis. *Gut* 2012; 61(2):327.

174. Bengi, G., Akpinar, H. What is the importance of infliximab and cyclosporine in the treatment of corticosteroid-refractory severe ulcerative colitis? *Turkish Journal of Gastroenterology* 2012; 23(SUPPL.2):2012.

175. Biancone, L., Petruzziello, C., Calabrese, E., Zorzi, F., Naccarato, P., Onali, S. et al. Long-term safety of Infliximab for the treatment of inflammatory bowel disease: does blocking TNFalpha reduce colitis-associated colorectal carcinogenesis? *Gut* 2009; 58(12):1703.

176. Ciclosporin no better than infliximab for acute refractory ulcerative colitis. *BMJ* 2012; 345:e6923.

177. Bordeianou, L. In flux on infliximab: Conflicting studies on surgical outcomes. *Inflammatory Bowel Diseases* 2009; 15(10):2009.

178. Borruel, N., Navarro, E., Robles, V., Torrejon, A., Casellas, F. Factors influencing adherence to treatment with anti-TNF agents in ulcerative and Crohn's disease patients: A prospective study. *Journal of Crohn's and Colitis Conference: 8th Congress of the European Crohn's and Colitis Organisation, ECCO 2013 Vienna Austria Conference Start: 20130214 Conference End: 20130216 Conference Publication: (Var Pagings)* 2013; 7(pp S212):February.

179. Brooklyn, T.N., Dunnill, M.G., Shetty, A., Bowden, J.J., Williams, J.D., Griffiths, C.E. et al. Infliximab for the treatment of pyoderma gangrenosum: a randomised, double blind, placebo controlled trial. *Gut* 2006; 55(4):505-509.

180. Bujanover, Y., Weiss, B. Anti-tumor necrosis factor therapy for pediatric inflammatory bowel diseases. *Israel Medical Association Journal* 2008; 10(9):August/September.

181. Busquets, D., Aldeguer, X. Clinical experience with adalimumab in anti-TNF-naive patients with ulcerative colitis. *Journal of Crohn's & Colitis* 2013; 7(5):e195.

182. Carbone, J., Perez-Rojas, J., Sarmiento, E. Infectious pulmonary complications in patients treated with anti-TNF-alpha monoclonal antibodies and soluble TNF receptor. *Current Infectious Disease Reports* 2009; 11(3):May.

183. Carinanos, I., Acosta, M.B., Domenech, E. Adalimumab for pyoderma gangrenosum associated with inflammatory bowel disease. *Inflammatory Bowel Diseases* 2011; 17(12):E153-E154.

184. Casteele, V.N., Compernolle, G., Ballet, V., Van, A.G., Gils, A., Vermeire, S. et al. Individualised infliximab treatment using therapeutic drug monitoring: A prospective controlled Trough level Adapted infliXImab Treatment (TAXIT) trial. *Journal of Crohn's and Colitis Conference: 7th Congress of the European Crohn's and Colitis Organisation, ECCO Barcelona Spain Conference Start: 20120216 Conference End: 20120218 Conference Publication: (Var Pagings)* 2012; 6(pp S6):February.

185. Casteele, N.V., Compernolle, G., Ballet, V., Van, A.G., Gils, A., Vermeire, S. et al. Results on the optimisation phase of the prospective controlled trough level adapted infliximab treatment (TAXIT) trial. *Gastroenterology Conference: Digestive Diease Week 2012, DDW 2012 San Diego, CA United States Conference Start: 20120519 Conference End: 20120522 Conference Publication: (Var Pagings)* 2012; 142(5 SUPPL.#1):May.

186. Charles, P., Ackermann, F., Burdy, G., Bletry, O., Leport, J., Kahn, J.E. Cytomegalovirus colitis complicating ulcerative colitis treated with adalimumab. *Scandinavian Journal of Gastroenterology* 2010; 45(4):509-510.

187. Chen, M., Black, C.M., Gurunath, S., Jansen, J.P., Chaudhary, M.A., Fan, T. Network meta-analysis of approved biologic interventions for the maintenance of response in ulcerative colitis. *Value in Health Conference: ISPOR 16th Annual European Congress Dublin Ireland Conference Start: 20131102 Conference End: 20131106 Conference Publication: (Var Pagings)* 2013; 16(7):November.

188. Chey, W.Y., Shah, A. Infliximab for ulcerative colitis. *Journal of Clinical Gastroenterology* 2005; 39(10):920-Dec.

189. Chowers, Y., Sturm, A., Sans, M., Papadakis, K., Gazouli, M., Harbord, M. et al. Report of the ECCO workshop on anti-TNF therapy failures in inflammatory bowel diseases: biological roles and effects of TNF and TNF antagonists. *Journal of Crohn's & Colitis* 2010; 4(4):367-376.

190. Chuang, M.H., Singh, J., Ashouri, N., Katz, M.H., Arrieta, A.C. Listeria meningitis after infliximab treatment of ulcerative colitis. *Journal of Pediatric Gastroenterology & Nutrition* 2010; 50(3):337-339.

191. Cohen, R.D. Infliximab in ulcerative colitis: is there a placebo (effect) in the house? *Gastroenterology* 2003; 124(7):1990-1991.

192. Colombel, J.F., Rutgeerts, P., Reinisch, W., Esser, D., Wang, Y., Lang, Y. et al. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. *Gastroenterology* 2011; 141(4):1194-1201.

193. Colombel, J.-F., Sandborn, W., Van, A.G., D'Haens, G., Wolf, D., Kron, M. et al. Benefit-risk assessment of adalimumab as maintenance treatment for ulcerative colitis: Near analysis of week 8 responders in ultra 2. *American Journal of Gastroenterology Conference: 77th Annual Scientific Meeting of the American College of Gastroenterology Las Vegas, NV United States Conference Start: 20121019 Conference End: 20121024 Conference Publication: (Var Pagings)* 2012; 107(pp S643-S644):October.

194. Cottone, M., Orlando, A., Mocciaro, F. Improving patients' QoL: how the success of treatment can improve workability. *Digestive and Liver Disease Supplements* 2009; 2(1):March.

195. Croft, A., Walsh, A., Doecke, J., Cooley, R., Howlett, M., Radford-Smith, G. Outcomes of salvage therapy for steroid-refractory acute severe ulcerative colitis: Ciclosporin vs. infliximab. *Alimentary Pharmacology and Therapeutics* 2013; 38(3):August.

196. Cross, R.K., Lapshin, O., Finkelstein, J. Patient subjective assessment of drug side effects in inflammatory bowel disease. *Journal of Clinical Gastroenterology* 2008; 42(3):March.

197. Danese, S. IBD: Golimumab in ulcerative colitis: a 'menage a trois' of drugs. *Nature Reviews Gastroenterology & Hepatology* 2013; 10(9):511-512.

198. de, V.M., Dewit, O., D'Haens, G., Baert, F., Fontaine, F., Vermeire, S. et al. Fast and sharp decrease in calprotectin predicts remission by infliximab in anti-TNF naive patients with ulcerative colitis. *Journal of Crohn's & Colitis* 2012; 6(5):557-562.

199. de Vries, H.S., van Oijen, M.G., de Jong, D.J. Safety of infliximab in inflammatory bowel disease needs to be debated. *Clinical Gastroenterology & Hepatology* 2009; 7(5):603-604.

200. D'Haens, G. Infliximab for ulcerative colitis: finally some answers. *Gastroenterology* 2005; 128(7):2161-2164.

201. Dean, K.E., Hikaka, J., Huakau, J.T., Walmsley, R.S. Infliximab or cyclosporine for acute severe ulcerative colitis: a retrospective analysis. *Journal of Gastroenterology & Hepatology* 2012; 27(3):487-492.

202. Dignass, A., Lindsay, J.O., Sturm, A., Windsor, A., Colombel, J.-F., Allez, M. et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 2: Current management. *Journal of Crohn's and Colitis* 2012; 6(10):December.

203. Domenech, E., Zabana, Y., Manosa, M., Garcia-Planella, E., Cabre, E., Gassull, M.A. Infliximab reintroduction is not associated to a higher rate of immune-related adverse effects in patients with inflammatory bowel disease initially treated with a three-infusion induction regimen. *Journal of Clinical Gastroenterology* 2010; 44(1):34-37.

204. Dranitsaris, G., Dorward, K., Hatzimichael, E., Amir, E. Clinical trial design in biosimilar drug development. *Investigational New Drugs* 2012; 31(2):April.

205. Eidelwein, A.P., Cuffari, C., Abadom, V., Oliva-Hemker, M. Infliximab efficacy in pediatric ulcerative colitis. *Inflammatory Bowel Diseases* 2005; 11(3):213-218.

206. Eriksson, U., Janson, U.A., Strid, H., Ohman, L., Bajor, A. Combination therapy with infliximab and thiopurines might reduce colectomy rates compared to infliximab monotherapy in patients with ulcerative colitis (UC). *Journal of Crohn's and Colitis Conference: 7th Congress of the European Crohn's and Colitis Organisation, ECCO Barcelona Spain Conference Start: 20120216 Conference End: 20120218 Conference Publication: (Var Pagings)* 2012; 6(pp S108-S109):February.

207. Esteve, M., Mahadevan, U., Sainz, E., Rodriguez, E., Salas, A., Fernandez-Banares, F. Efficacy of anti-TNF therapies in refractory severe microscopic colitis. *Journal of Crohn's & Colitis* 2011; 5(6):612-618.

208. WHO ICTRP record. 2014 available from: http://apps.who.int/trialsearch/ [date accessed 10/03/2014]

209. Fanjiang, G., Russell, G.H., Katz, A.J. Short- and long-term response to and weaning from infliximab therapy in pediatric ulcerative colitis. *Journal of Pediatric Gastroenterology & Nutrition* 2007; 44(3):312-317.

210. Fasanmade, A.A., Adedokun, O.J., Ford, J., Hernandez, D., Johanns, J., Hu, C. et al. Population pharmacokinetic analysis of infliximab in patients with ulcerative colitis. *European Journal of Clinical Pharmacology* 2009; 65(12):1211-1228.

211. Fasanmade, A.A., Adedokun, O.J., Olson, A., Strauss, R., Davis, H.M. Serum albumin concentration: a predictive factor of infliximab pharmacokinetics and clinical

response in patients with ulcerative colitis. *International Journal of Clinical Pharmacology & Therapeutics* 2010; 48(5):297-308.

212. Feagan, B.G., Greenberg, G.R., Wild, G., Fedorak, R.N., Pare, P., McDonald, J.W.D. et al. Treatment of ulcerative colitis with a humanized antibody to the alpha(4)beta(7) integrin. *New England Journal of Medicine* 2005; 352(24):2499-2507.

213. Feagan, B. Infliximab, adalimumab and Certolizumab: Beneficial and adverse effects (past). *Inflammatory Bowel Diseases* 2006; 12:S6.

214. Florholmen, J., Overland, G., Olsen, T., Rismo, R., Cui, G., Christiansen, I. et al. Short-and long-term clinical outcomes of infliximab in fulminant ulcerative colitis. *Ulcers , 2011 Article Number: 156407 Date of Publication: 2011* 2011;156407.

215. Gao, J., Jiang, X.L. Low-dose infliximab for corticosteroid-refractory ulcerative colitis: Impact of number of infusions on efficacy and safety. *World Chinese Journal of Digestology* 2013; 21:1453-1457.

216. Gavalas, E., Kountouras, J., Stergiopoulos, C., Zavos, C., Gisakis, D., Nikolaidis, N. et al. Efficacy and safety of infliximab in steroid-dependent ulcerative colitis patients. *Hepato-Gastroenterology* 2007; 54(76):1074-1079.

217. Gearry, R.B., Falvey, J.D. Rescue therapy for steroid refractory acute severe ulcerative colitis: more choice, better outcomes? *Journal of Gastroenterology & Hepatology* 2012; 27(3):417-418.

218. Gies, N., Kroeker, K.I., Wong, K., Fedorak, R.N. Treatment of ulcerative colitis with adalimumab or infliximab: long-term follow-up of a single-centre cohort. *Alimentary Pharmacology & Therapeutics* 2010; 32(4):522-528.

219. Ginard, D., Khorrami, S., Vanrell, M. Efficacy and effectiveness of biological therapy in inflammatory bowel disease (IBD): Comparative study between clinical practice and pivotal studies. *Journal of Crohn's and Colitis Conference: 6th Congress of the European Crohn's and Colitis Organisation, ECCO Dublin Ireland Conference Start: 20110224 Conference End: 20110226 Conference Publication: (Var Pagings)* 2008; 5(1):February.

220. Grosen, A., Julsgaard, M., Christensen, L.A. Serum sickness-like reaction due to Infliximab reintroduction during pregnancy. *Journal of Crohn's & Colitis* 2013; 7(5):e191.

221. Gustavsson, A., Jarnerot, G., Hertervig, E., Friis-Liby, I., Blomquist, L., Karlen, P. et al. Clinical trial: colectomy after rescue therapy in ulcerative colitis - 3-year follow-up of the Swedish-Danish controlled infliximab study. *Alimentary Pharmacology & Therapeutics* 2010; 32(8):984-989.

222. Ha, C.Y., Ullman, T.A., Siegel, C.A., Kornbluth, A. Are patients enrolled in randomized controlled trials (RCTs) representative of the general Inflammatory Bowel Disease (IBD) patient population? *Gastroenterology Conference: Digestive Disease Week, DDW 2011 Chicago, IL United States Conference Start: 20110507 Conference End: 20110510 Conference Publication: (Var Pagings)* 2009; 140(5 SUPPL.#1):May.

223. Halpin, S.J., Hamlin, P.J., Ford, A.C. Infliximab rescue therapy in ulcerative colitis, and the effect on subsequent colectomy rates. *Alimentary Pharmacology & Therapeutics* 2010; 32(10):1294-1295.

224. Halpin, S., Hamlin, P.J. Commentary: outcomes after escalation of infliximab in moderately active ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 958; 35(8):957-958.

225. Hamalainen, A., Sipponen, T., Kolho, K.L. Infliximab in pediatric inflammatory bowel disease rapidly decreases fecal calprotectin levels. *World Journal of Gastroenterology* 2011; 17(47):5166-5171.

226. Hanauer, S.B. Infliximab or cyclosporine for severe ulcerative colitis. *Gastroenterology* 2005; 129(4):1358-1359.

227. Hanauer, S.B., Abreu, M.T., Kornbluth, A.A., Present, D.H., Sandborn, W.J. Charging ahead in IBD: Challenging issues and recent breakthroughs - A report of a symposium presented on May 19, 2008 San Diego, California. *Gastroenterology and Hepatology* 2008; 4(11 SUPPL.#25):November.

228. Hanauer, S.B., Rubin, D.T., Sandborn, W.J. Treatment guidelines and clinical practice: Optimizing foundational therapies for ulcerative colitis. *Gastroenterology and Hepatology* 2008; 4(10 SUPPL.#22):October.

229. Heraganahally, S.S., Au, V., Kondru, S., Edwards, S., Bowden, J.J., Sajkov, D. Pulmonary toxicity associated with infliximab therapy for ulcerative colitis. *Internal Medicine Journal* 2009; 39(9):629-630.

230. Herrlinger, K.R., Barthel, D.N., Schmidt, K.J., Buning, J., Barthel, C.S., Wehkamp, J. et al. Infliximab as rescue medication for patients with severe ulcerative/indeterminate colitis refractory to tacrolimus. *Alimentary Pharmacology & Therapeutics* 2010; 31(9):1036-1041.

231. Honeywell, M., Touchstone, K., Caspi, A. Infliximab: A chimeric monoclonal antibody against tumor necrosis factor. *P and T* 2007; 32(8 II):August.

232. Hyams, J.S., Lerer, T., Griffiths, A., Pfefferkorn, M., Stephens, M., Evans, J. et al. Outcome following infliximab therapy in children with ulcerative colitis. *American Journal of Gastroenterology* 2010; 105(6):1430-1436.

233. Hyams, J., Damaraju, L., Blank, M., Johanns, J., Guzzo, C., Winter, H. et al. A randomized, multicenter, open-label phase 3 study to evaluate the safety and efficacy of infliximab in pediatric patients with moderate to severe ulcerative colitis. *Journal of Pediatric Gastroenterology and Nutrition Conference: European Society for Paediatric Gastroenterology, Hepatology, and Nutrition Annual Meeting 2011 Sorrento Italy Conference Start: 20110525 Conference End: 20110528 Conference Publication: (* 2011; 52(pp E7-E8):June.

234. Assasi, N., Blackhouse, G., Xie, F., Gaebel, K., Marshall, J., Irvine, E.J. et al. Anti-TNF-alpha drugs for refractory inflammatory bowel disease: clinical- and cost-effectiveness analyses (Structured abstract). *Database of Abstracts of Reviews of Effects* 2009;1.

235. Jackson, J.M. TNF-alpha inhibitors. *Dermatologic Therapy* 2007; 20(4):July.

236. Jarnerot, G. Infliximab or cyclosporine for severe ulcerative colitis. *Gastroenterology* 2006; 130(1):286.

237. Jimenez, J.M. Infliximab in the treatment of severe ulcerative colitis. *Revista Espanola De Enfermedades Digestivas* 2004; 96(2):89-93.

238. Joob, B., Wiwanikit, V. Infliximab for treatment of pyoderma gangrenosum with ulcerative colitis. *Journal of Crohn's & Colitis* 2013; 7(4):e153.

239. Kaser, A., Tilg, H. Inflammatory bowel diseases: Highlights from the United European Gastroenterology Week 2008 - 18-22 October 2008, Vienna, Austria. *Expert Opinion on Therapeutic Targets* 2009; 13(2):February.

240. Kaur, M., Targan, S.R. Steroid-refractory ulcerative colitis-ciclosporin or infliximab? *Nature Reviews Gastroenterology & Hepatology* 2013; 10(1):8-9.

241. Kerbleski, J.F., Gottlieb, A.B. Dermatological complications and safety of anti-TNF treatments. *Gut* 2009; 58(8):August.

242. Klotz, U., Teml, A., Schwab, M. Clinical pharmacokinetics and use of infliximab. *Clinical Pharmacokinetics* 2007; 46(8):2007.

243. Kohn, A., Prantera, C., Pera, A., Cosintino, R., Sostegni, R., Daperno, M. Infliximab in the treatment of severe ulcerative colitis: a follow-up study. *European Review for Medical & Pharmacological Sciences* 2004; 8(5):235-237.

244. Kohn, A., Daperno, M., Armuzzi, A., Cappello, M., Biancone, L., Orlando, A. et al. Infliximab in severe ulcerative colitis: short-term results of different infusion regimens and long-term follow-up. *Alimentary Pharmacology & Therapeutics* 2007; 26(5):747-756.

245. Kohn, A. Is there a role for infliximab in severe ulcerative colitis?: the European experience.[Erratum appears in Inflamm Bowel Dis. 2009 Sep;15(9):1438-47]. *Inflammatory Bowel Diseases* 2008; 14:Suppl-5.

246. Laharie, D., Bourreille, A., Branche, J., Allez, M., Bouhnik, Y., Filippi, J. et al. Ciclosporin versus infliximab in patients with severe ulcerative colitis refractory to intravenous steroids: a parallel, open-label randomised controlled trial. *Lancet* 2012; 380(9857):1909-1915.

247. Laharie, D., Bourreille, A., Branche, J., Allez, M., Bouhnik, Y., Filippi, J. et al. Cyclosporin versus infliximab in severe acute ulcerative colitis refractory to intravenous steroids: A randomized trial. *Gastroenterology Conference: Digestive Disease Week, DDW 2011 Chicago, IL United States Conference Start: 20110507 Conference End: 20110510 Conference Publication: (Var Pagings)* 2011; 140(5 SUPPL.#1):S112.

248. Leal, R.F., de Lourdes Setsuko, A.M., de Sene Portel, O.P., Fagundes, J.J., Coy, C.S. Impact of a best clinical therapy management on surgery for ulcerative colitis. *Techniques in Coloproctology* 2012; 16(4):321-322.

249. LeBlanc, K., Mosli, M., Baker, K.A., MacDonald, J.K. The impact of biological interventions for ulcerative colitis on health-related quality of life. *Cochrane Database of Systematic Reviews* 2013.

250. Leblanc, S., Allez, M., Seksik, P., Flourie, B., Peeters, H., Dupas, J.L. et al. Successive treatment with cyclosporine and infliximab in steroid-refractory ulcerative colitis. *American Journal of Gastroenterology* 2011; 106(4):771-777.

251. Levesque, B.G., Sandborn, W.J. Infliximab versus ciclosporin in severe ulcerative colitis. *Lancet* 2012; 380(9857):1887-1888.

252. Levy, L.C. Advances in Inflammatory Bowel Diseases. *Expert Review of Gastroenterology and Hepatology* 2009; 3(2):2009.

253. Li, K., Telesco, S., Rutgeerts, P.J., Sandborn, W., Marano, C.W., Ma, K. et al. Characterization of molecular response to golimumab in ulcerative colitis by MRNA expression profiling: Results from pursuit-SC induction study. *Gastroenterology Conference: Digestive Disease Week 2013, DDW 2013 Orlando, FL United States Conference Start: 20130518 Conference End: 20130521 Conference Publication: (Var Pagings)* 2013; 144(5 SUPPL.#1):May.

254. Lichtenstein, G.R. Is infliximab effective for induction of remission in patients with ulcerative colitis? *Inflammatory Bowel Diseases* 2001; 7(2):89-93.

255. Lichtenstein, G.R. Applying clinical trial data to biologic treatment of ulcerative colitis. *Gastroenterology and Hepatology* 2009; 5(5):May.

256. Liu, Y., Reichmann, B., Yan, S., Macaulay, D., Skup, M., Mulani, P. et al. Indirect comparison of adalimumab and infliximab in ulcerative colitis: Cost per remitter analysis. *American Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings)* 2013; 108(pp S520):October.

257. Lofberg, R., Louis, E.V., Reinisch, W., Robinson, A.M., Kron, M., Camez, A. et al. Adalimumab produces clinical remission and reduces extraintestinal manifestations in Crohn's disease: results from CARE. *Inflammatory Bowel Diseases* 2012; 18(1):1-9.

258. Lorenzo-Zuniga, V., Boix, J., Manosa, M., Lezcano, C., Cabre, E., Moreno, d.V., V et al. Local injection of infliximab in symptomatic isolated mucosal lesions: a novel scenario for endoscopic therapy? *Inflammatory Bowel Diseases* 2013; 19(4):E59-E61.

259. Mallow, P., Rizzo, J., Queener, M.K., Gathany, T., Lofland, J. Cost per clinical remission and clinical response of adalimumab, golimumab, and infliximab in the treatment of patients with ulcerative colitis: An induction analysis. *American Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings)* 2013; 108(pp S515-S516):October.

260. Mallow, P., Rizzo, J., Queener, M., Gathany, T., Lofland, J. Comparative efficacy of adalimumab, golimumab, and infliximab in the treatment of patients with ulcerative colitis: An induction analysis. *American Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings)* 2013; 108(pp S514):October.

261. Maser, E.A., Deconda, D., Lichtiger, S., Ullman, T., Present, D.H., Kornbluth, A. Cyclosporine and infliximab as rescue therapy for each other in patients with steroid-

refractory ulcerative colitis. *Clinical Gastroenterology & Hepatology* 2008; 6(10):1112-1116.

262. Matsumoto, T. Treatment of inflammatory immunologic disease 2. Anti-cytokine therapies in inflammatory bowel disease. *Internal Medicine* 2007; 46(16):15.

263. Mazumdar, S., Greenwald, D. Golimumab. *MAbs* 2009; 1(5):September-October.

264. McCann, D.A., Smith, H.L. Infliximab-associated Blastomycosis dermatitidis in treatment of ulcerative colitis. *Colorectal Disease* 2013; 15(2):e102-e103.

265. Molnar, T., Farkas, K., Nagy, F., Szepes, Z., Wittmann, T. Infliximab safety profile and long-term applicability in inflammatory bowel disease: clinical experiences from the eastern side of Europe. *Alimentary Pharmacology & Therapeutics* 2010; 31(10):1152-1153.

266. Molnar, T., Farkas, K., Szepes, Z., Nagy, F., Wittmann, T. Long-term outcome of infliximab therapy is highly comparable in a Danish and in a Hungarian tertiary center. *Scandinavian Journal of Gastroenterology* 2011; 46(2):248-249.

267. Molnar, T., Farkas, K., Szepes, Z., Nagy, F., Wittmann, T. Infliximab rescue therapy in steroid-refractory ulcerative colitis: is more really more? *Alimentary Pharmacology & Therapeutics* 2011; 33(3):412-413.

268. Moss, A.C., Farrell, R.J. Infliximab for induction and maintenance therapy for ulcerative colitis. *Gastroenterology* 2006; 131(5):1649-1651.

269. Nakase, H., Yamamoto, S., Matsuura, M., Honzawa, Y., Chiba, T. Cytomegalovirus affects clinical outcome of infliximab in ulcerative colitis refractory to tacrolimus. *Alimentary Pharmacology & Therapeutics* 2010; 32(3):510-511.

270. Infliximab (Remicade) for paediatric ulcerative colitis (Structured abstract). *Health Technology Assessment Database* 2011.

271. Golimumab (Simponi) for ulcerative colitis - second line (Structured abstract). *Health Technology Assessment Database* 2011.

272. COmparison of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis: a Trial (CONSTRUCT) (Project record). *Health Technology Assessment Database* 2010.

273. Clinical Trials.gov record. 2013; available from: http://clinicaltrials.gov/ [date accessed 15/03/2014]

274. Nguyen, G.C., Prather, C.M. Is keeping the colon the ultimate marker of success in ulcerative colitis? *Gastroenterology* 2009; 137(4):1204-1206.

275. Nielsen, O.H., Jess, T. IBD: Can TNF inhibitors be administered during the third trimester? *Nature Reviews Gastroenterology & Hepatology* 2013; 10(3):130-131.

276. Ochsenkuhn, T., Sackmann, M., Goke, B. Infliximab for acute, not steroid-refractory ulcerative colitis: a randomized pilot study. *European Journal of Gastroenterology & Hepatology* 2004; 16(11):1167-1171.

277. Orlando, A., Renna, S., Mocciaro, F., Cappello, M., Di, M.R., Randazzo, C. et al. Adalimumab in steroid-dependent Crohn's disease patients: prognostic factors for clinical benefit. *Inflammatory Bowel Diseases* 2012; 18(5):826-831.

278. Oussalah, A., Laclotte, C., Chevaux, J.B., Bensenane, M., Babouri, A., Serre, A.A. et al. Long-term outcome of adalimumab therapy for ulcerative colitis with intolerance or lost response to infliximab: a single-centre experience. *Alimentary Pharmacology & Therapeutics* 2008; 28(8):966-972.

279. Oussalah, A., Evesque, L., Laharie, D., Roblin, X., Boschetti, G., Nancey, S. et al. A multicenter experience with infliximab for ulcerative colitis: outcomes and predictors of response, optimization, colectomy, and hospitalization. *American Journal of Gastroenterology* 2010; 105(12):2617-2625.

280. Panccione, R., Ghosh, S., Middleton, S., Marquez, J.R., Khalif, I., Flint, L. et al. Infliximab, azathioprine, or infliximab + azathioprine for treatment of moderate to severe ulcerative colitis: The UC success trial. *Gastroenterology Conference: Digestive Disease Week, DDW 2011 Chicago, IL United States Conference Start: 20110507 Conference End: 20110510 Conference Publication: (Var Pagings)* 2011; 140(5 SUPPL.#1):May.

281. Panaccione, R., Ghosh, S., Middleton, S., Marquez, J., Flint, L., Hoogstraten, J.v. et al. Improvement in patient quality of life during treatment with infliximab, azathioprine, or combination infliximab+azathioprine for moderate-to-severe ulcerative colitis. *American Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings)* 2013; 108(pp S544-S545):October.

282. Pardi, D.S., Sandborn, W.J. Infliximab in ulcerative colitis is associated with an increased risk of postoperative complications after restorative proctocolectomy: Commentary. *Diseases of the Colon and Rectum* 2008; 51(8):August.

283. Pastore, S., Londero, M., Gortani, G., Abate, M.V., Marchetti, F., Di, L.G. et al. Infliximab-related vasculitis in patients affected by ulcerative colitis. *Journal of Pediatric Gastroenterology & Nutrition* 2010; 51(2):226-228.

284. Pastorelli, L., Pizarro, T.T., Cominelli, F., Vecchi, M. Emerging drugs for the treatment of ulcerative colitis. *Expert Opinion on Emerging Drugs* 2009; 14(3):September.

285. Pearce, C.B., Lawrance, I.C. Careful patient selection may improve response rates to infliximab in inflammatory bowel disease. *Journal of Gastroenterology & Hepatology* 2007; 22(10):1671-1677.

286. Peyrin-Biroulet, L., Laclotte, C., Roblin, X., Bigard, M.A. Adalimumab induction therapy for ulcerative colitis with intolerance or lost response to infliximab: an open-label study. *World Journal of Gastroenterology* 2007; 13(16):2328-2332.

287. Pola, S., Fahmy, M., Evans, E., Tipps, A., Sandborn, W.J. Successful use of infliximab in the treatment of corticosteroid dependent collagenous colitis. *American Journal of Gastroenterology* 2013; 108(5):857-858.

288. Reinisch, W.F., Sandborn, W.J. Infliximab concentration and clinical outcome in patients with ulcerative colitis. *Gastroenterology* 2012; 142 (5 Suppl 1):S114.

289. Rizzello, F., Pratico, C., Calabrese, C., Gionchetti, P. Rescue therapy: Ciclosporin or infliximab? *Expert Review of Clinical Immunology* 2013; 9(6):June.

290. Rostholder, E., Ahmed, A., Cheifetz, A.S., Moss, A.C. Outcomes after escalation of infliximab therapy in ambulatory patients with moderately active ulcerative colitis. *Alimentary Pharmacology & Therapeutics* 2012; 35(5):562-567.

291. Rubin, D.T., Mulani, P., Chao, J., Pollack, P.F., Bensimon, A.G., Yu, A.P. et al. Effect of adalimumab on clinical laboratory parameters in patients with Crohn's disease: results from the CHARM trial. *Inflammatory Bowel Diseases* 2012; 18(5):818-825.

292. Russell, G.H., Katz, A.J. Infliximab is effective in acute but not chronic childhood ulcerative colitis. *Journal of Pediatric Gastroenterology & Nutrition* 2004; 39(2):166-170.

293. Rutgeerts, P. Infliximab for ulcerative colitis: the need for adequately powered placebo-controlled trials. *American Journal of Gastroenterology* 2002; 97(10):2488-2489.

294. Rutgeerts, P., Vermeire, S., Van, A.G. Predicting the response to infliximab from trough serum levels. *Gut* 2010; 59(1):7-8.

295. Rutgeerts, P., Feagan, B., Marano, C., Strauss, R., Johanns, J., Zhang, H. et al. Phase 2/3 randomized, placebo-controlled, double-blind study of SC golimumab induction in moderate to Severe UC. *Journal of Gastroenterology and Hepatology Conference: Asian Pacific Digestive Week 2013, APDW 2013 - World Congress of Gastroenterology, WCOG 2013 Shanghai China Conference Start: 20130921 Conference End: 20130924 Conference Publication: (Var Pagings)* 2013; 28(pp 591):October.

296. Rutgeerts, P., Feagan, B.G., Marano, C., Strauss, R., Johanns, J., Zhang, H. et al. Phase 3 study to evaluate the SC golimumab maintenance therapy in moderate to severe UC: Pursuit-maintenance. *Journal of Gastroenterology and Hepatology Conference: Asian Pacific Digestive Week 2013, APDW 2013 - World Congress of Gastroenterology, WCOG 2013 Shanghai China Conference Start: 20130921 Conference End: 20130924 Conference Publication: (Var Pagings)* 2013; 28(pp 593):October.

297. Salvana, E.M.T., Salata, R.A. Infectious complications associated with monoclonal antibodies and related small molecules. *Clinical Microbiology Reviews* 2009; 22(2):April.

298. Sandborn, W.J., Wagner, C.L., Fasanmade, A.A., Diamond, R.H., Olson, A., Marano, C.W. et al. Effects of immunomodulators on pharmacokinetics and immunogenicity of inffiximab adminstered as 3-dose induction followed by systematic maintenance therapy in IBD. *Gastroenterology* 2007; 132(4, Suppl. 2):A504-A505.

299. Sandborn, W.J. Mucosal healing with infliximab: Results from the active ulcerative colitis trials. *Gastroenterology and Hepatology* 2012; 8(2):February.

300. Sandborn, W.J., Van Assche, G.A., Reinisch, W., Colombel, J.F., D'Haens, G.R., Wolf, D.C. et al. Induction and Maintenance of Clinical Remission by Adalimumab in Patients With Moderate-to-Severe Ulcerative Colitis. *Gastroenterology* 2011; 140(5, Suppl. 1):S123-S124.

301. Sandborn, W.J., Wolf, D.C., Van, A.G., Yang, M., Mulani, P.M., Reilly, N. et al. Rapid onset of adalimumab and long-term efficacy among week-8responders in adults with moderate to severe active Ulcerative Colitis. *Inflammatory Bowel Diseases Conference: 2011 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20111201 Conference End: 20111203 Conference P* 2011; 17(pp S4):December.

302. Sandborn, W., Feagan, B., Yang, M., Skup, M., Thakkar, R., Lazar, A. et al. Adalimumab therapy reduces hospitalization rates in patients with ulcerative colitis among initial responders. *Inflammatory Bowel Diseases Conference: 2012 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20121213 Conference End: 20121215 Conference P* 2012; 18(pp S40):December.

303. Sandborn, W.J., Colombel, J.F., D'Haens, G., Van, A.G., Wolf, D., Kron, M. et al. One-year maintenance outcomes among patients with moderately-to-severely active ulcerative colitis who responded to induction therapy with adalimumab: subgroup analyses from ULTRA 2. *Alimentary Pharmacology & Therapeutics* 2013; 37(2):204-213.

304. Sandborn, W.J., Feagan, B.G., Marano, C.W., Strauss, R., Johanns, J., Zhang, H. et al. A phase 2/3 randomized, placebo-controlled, double-blind study to evaluate the safety and efficacy of subcutaneous golimumab induction therapy in patients with moderately to severely active ulcerative colitis: Pursuit sc. *Gastroenterology Conference: Digestive Diease Week 2012, DDW 2012 San Diego, CA United States Conference Start: 20120519 Conference End: 20120522 Conference Publication: (Var Pagings)* 2012; 142(5 SUPPL.#1):May.

305. Sandborn, W., Feagan, B., Lazar, A., Thakkar, R., Skup, M., Yang, M. et al. Adalimumab induction dose reduces hospitalization risk in patients with ulcerative colitis during the first 8 weeks of therapy. *Inflammatory Bowel Diseases Conference: 2012 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20121213 Conference End: 20121215 Conference P* 2012; 18(pp S40-S41):December.

306. Sandborn, W., Feagan, B., Marano, C., Strauss, R., Johanns, J., Zhang, H. et al. Clinical response is a meaningful endpoint in ulcerative colitis clinical studies. *Inflammatory Bowel Diseases Conference: 2012 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20121213 Conference End: 20121215 Conference P* 2012; 18(pp S26-S27):December.

307. Sandborn, W.J., Loftus, E.V. Balancing the risks and benefits of infliximab in the treatment of inflammatory bowel disease. *Gut* 2004; 53(6):780-782.

308. Sands, B.E., Tremaine, W.J., Sandborn, W.J., Rutgeerts, P.J., Hanauer, S.B., Mayer, L. et al. Infliximab in the treatment of severe, steroid-refractory ulcerative colitis: a pilot study. *Inflammatory Bowel Diseases* 2001; 7(2):83-88.

309. Scholmerich, J. Balancing the risks and benefits of prolonged use of infliximab. *Gut* 2009; 58(4):April.

310. Sciaudone, G., Pellino, G., Guadagni, I., Selvaggi, F. Infliximab avoiding colectomy and maintaining remission in pediatric immunosuppressive-naive ulcerative colitis achieving remarkable mucosal healing. *Panminerva Medica* 2010; 52(1):91-92.

311. Sciaudone, G., Pellino, G., Guadagni, I., Selvaggi, F. Education and imaging: gastrointestinal: herpes simplex virus-associated erythema multiforme (HAEM) during infliximab treatment for ulcerative colitis. *Journal of Gastroenterology & Hepatology* 2011; 26(3):610.

312. Seirafi, M., Treton, X., De Vroey, B., Cosnes, J., Roblin, X., Allez, M. et al. Anti-TNF Therapy and Pregnancy in Inflammatory Bowel Disease: A Prospective Cohort Study From the GETAID. *Gastroenterology* 2011; 140(5, Suppl. 1):S175.

313. Siemanowski, B., Regueiro, M. Efficacy of infliximab for extraintestinal manifestations of inflammatory bowel disease. *Current Treatment Options in Gastroenterology* 2007; 10(3):June.

314. Simmons, J., Jewell, D.P. Infliximab for ulcerative colitis. *Digestive & Liver Disease* 2002; 34(9):616-618.

315. Singh, S., Loftus, J. Management of severe steroid-refractory ulcerative colitis: Cyclosporine or infliximab? *Gastroenterology* 2013; 144(5):May.

316. Sjoberg, M., Walch, A., Meshkat, M., Gustavsson, A., Jarnerot, G., Vogelsang, H. et al. Infliximab or cyclosporine as rescue therapy in hospitalized patients with steroid-refractory ulcerative colitis: a retrospective observational study. *Inflammatory Bowel Diseases* 2012; 18(2):212-218.

317. Smith, K. Golimumab shows promise in treatment of active ulcerative colitis. *Nature Reviews Gastroenterology and Hepatology* 2013; 10(7):July.

318. Sokol, H., Seksik, P., Carrat, F., Nion-Larmurier, I., Vienne, A., Beaugerie, L. et al. Usefulness of co-treatment with immunomodulators in patients with inflammatory bowel disease treated with scheduled infliximab maintenance therapy. *Gut* 2010; 59(10):1363-1368.

319. Stein, O.M., Stein, P. Prediction of clinical responders to treatment with DIMS0150 a Toll-Like Receptor 9 agonist in therapy refractory ulcerative colitis patients. *Journal of Crohn's & Colitis* 2013; 7:S173.

320. Su, C., Salzberg, B.A., Lewis, J.D., Deren, J.J., Kornbluth, A., Katzka, D.A. et al. Efficacy of anti-tumor necrosis factor therapy in patients with ulcerative colitis. *American Journal of Gastroenterology* 2002; 97(10):2577-2584.

321. Taxonera, C., Estelles, J., Fernandez-Blanco, I., Merino, O., Marin-Jimenez, I., Barreiro-de, A.M. et al. Adalimumab induction and maintenance therapy for patients with ulcerative colitis previously treated with infliximab. *Alimentary Pharmacology & Therapeutics* 2011; 33(3):340-348.

322. Thorlund, K., Druyts, E., Mills, E., Fedorak, R., Marshal, J. Adalimumab versus infliximab for the treatment of moderate-to-severe ulcerative colitis in adult patients with no prior anti-TNF experience: An indirect comparison meta-analysis. *American Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference*

Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings) 2013; 108(pp S543):October.

323. Toedter, G., Ma, K., Li, K., Marano, C.W., Macoritto, M., Park, J. et al. Genes associated with reduced epithelial permeability and epithelial- mesenchymal transition: Changes in expression following infliximab therapy in ulcerative colitis. *Gastroenterology Conference: Digestive Disease Week, DDW 2010 New Orleans, LA United States Conference Start: 20100501 Conference End: 20100506 Conference Publication: (Var Pagings)* 2010; 138(5 SUPPL.#1):May.

324. Toedter, G., Li, K., Marano, C., Ma, K., Sague, S., Huang, C.C. et al. Gene expression profiling and response signatures associated with differential responses to infliximab treatment in ulcerative colitis. *American Journal of Gastroenterology* 2011; 106(7):1272-1280.

325. Travis, S. Does it all ADA up? Adalimumab for ulcerative colitis. *Gut* 2011; 60(6):741-742.

326. Tursi, A., Elisei, W., Brandimarte, G., Giorgetti, G., Penna, A., Castrignano, V. Safety and effectiveness of infliximab for inflammatory bowel diseases in clinical practice. *European Review for Medical & Pharmacological Sciences* 2010; 14(1):47-55.

327. Van Assche G. The optimization of immunosuppressive and biologic cotherapies in inflammatory bowel disease. *Gastroenterology and Hepatology* 2008; 4(3):March.

328. Van Assche G., Vermeire, S., Rutgeerts, P. Management of loss of response to anti-TNF drugs: Change the dose or change the drug? *Journal of Crohn's and Colitis* 2008; 2(4):December.

329. van Casteren-Messidoro, C., Prins, G., van, T.A., Zelinkova, Z., Schouten, J., de, M.R. Autoimmune hepatitis following treatment with infliximab for inflammatory bowel disease. *Journal of Crohn's & Colitis* 2012; 6(5):630-631.

330. Velayos, F., Mahadevan, U. Management of steroid-dependent ulcerative colitis: immunomodulatory agents, biologics, or surgery? *Clinical Gastroenterology & Hepatology* 2007; 5(6):668-671.

331. Vermeire, S., Ghosh, S., Panes, J., Dahlerup, J.F., Luegering, A., Sirotiakova, J. et al. The mucosal addressin cell adhesion molecule antibody PF-00547,659 in ulcerative colitis: A randomised study. *Gut* 2011; 60(8):August.

332. Warner, B., Harris, A.W. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2012; 143(1):e42.

333. Waters, H., McKenzie, R.S., Lunacsek, O., Lennert, B.J. Gastrointestinal-related service utilization and surgical procedures following infliximab initiation in patients with inflammatory bowel disease. *Journal of Managed Care Pharmacy Conference: AMCP's 2009 Educational Conference San Antonio, TX United States Conference Start: 20091007 Conference End: 20091009 Conference Publication: (Var Pagings)* 2008; 15(7):September.

334. Waters, H., McKenzie, R., Lunacsek, O., Franklin, M., Lennert, B., Piech, C. Stability of infliximab dosing in ulcerative colitis: Results from a chart review.

Inflammatory Bowel Diseases Conference: 2009 Advances in Inflammatory Bowel Diseases, Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20091203 Conference End: 20091206 Conference 2007; 15(pp S57):December.

335. Willert, R.P., Lawrance, I.C. Use of infliximab in the prevention and delay of colectomy in severe steroid dependant and refractory ulcerative colitis. *World Journal of Gastroenterology* 2008; 14(16):2544-2549.

336. Wolf, R., Matz, H., Orion, E., Ruocco, V. Anti-TNF therapies - The hope of tomorrow. *Clinics in Dermatology* 2002; 20(5):September/October.

337. Wolf, D., D'Haens, G., Sandborn, W., Colombel, J.-F., Van, A.G., Lazar, A. et al. Rate of and response to dose escalation in patients treated with adalimumab for moderately-to-severely active ulcerative colitis: Ultra 2 subanalysis. *Inflammatory Bowel Diseases Conference: 2012 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20121213 Conference End: 20121215 Conference P* 2012; 18(pp S22):December.

338. Yamamoto, T., Shiraki, M. The impact of preoperative infliximab therapy on postoperative complications in patients with ulcerative colitis. *Journal of Crohn's & Colitis* 2012; 6(10):1042.

339. Yamamoto-Furusho, J.K., Uzcanga, L.F. Infliximab as a rescue therapy for hospitalized patients with severe ulcerative colitis refractory to systemic corticosteroids. *Digestive Surgery* 2008; 25(5):383-386.

340. Yapali, S., Hamzaoglu, H.O. Anti-TNF treatment in inflammatory bowel disease. *Annals of Gastroenterology* 2007; 20(1):2007.

341. Sandborn, W., Colombel, J.-F., Feagan, B., Marano, C., Strauss, R., Han, C. et al. Early and sustained remission after treatment with subcutaneously administered golimumab is associated with normalized health-related quality of life in patients with moderate to severe ulcerative colitis: Post-HOC analysis from pursuit induction and maintenance trials. *American Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings)* 2013; 108(pp S519):October.

342. Panaccione, R., Ghosh, S., Middleton, S., Velazquez, J.R.M., Khalif, I., Flint, L. et al. Infliximab, azathioprine, or infliximab + azathioprine for treatment of moderate to severe ulcerative colitis: The UC SUCCESS trial. *Journal of Crohn's and Colitis Conference: 6th Congress of the European Crohn's and Colitis Organisation, ECCO Dublin Ireland Conference Start: 20110224 Conference End: 20110226 Conference Publication: (Var Pagings)* 2011; 5(1):February.

343. Lichtenstein, G.R., Diamond, R.H., Wagner, C.L., Fasanmade, A.A., Olson, A.D., Marano, C.W. et al. Clinical trial: benefits and risks of immunomodulators and maintenance infliximab for IBD-subgroup analyses across four randomized trials. *Alimentary Pharmacology & Therapeutics* 2009; 30(3):210-226.

344. Reinisch, W., Sandborn, W., Feagan, B., Ghosh, S., Robinson, A., Skup, M. et al. Association between week eight mayo subscores and hospitalization rates in adalimumab-treated patients with ulcerative colitis from ultra 1 and ultra 2. *American*

Journal of Gastroenterology Conference: 78th Annual Scientific Meeting of the American College of Gastroenterology San Diego, CA United States Conference Start: 20131011 Conference End: 20131016 Conference Publication: (Var Pagings) 2013; 108(pp S206):October.

345. Reinisch, W., Sandborn, W., Thakkar, R., Lazar, A., Huang, B., Mulani, P. et al. Adalimumab induction therapy improves health-related quality of life in patients with moderately to severely active ulcerative colitis. *American Journal of Gastroenterology Conference: 75th Annual Scientific Meeting of the American College of Gastroenterology San Antonio, TX United States Conference Start: 20101015 Conference End: 20101020 Conference Publication: (Var Pagings)* 2010; 105(pp S441-S442):October.

346. Feagan, B., Gibson, P., Marano, C., Strauss, R., Han, C., Johanns, J. et al. Impact of golimumab sc on disease specific and generic hrqol in patients with moderately to severely active UC: Pursuit-SC induction. *Inflammatory Bowel Diseases Conference: 2012 Advances in Inflammatory Bowel Diseases Crohn's and Colitis Foundation's National Clinical and Research Conference Hollywood, FL United States Conference Start: 20121213 Conference End: 20121215 Conference P* 2012; 18(pp S43):December.

347. Feagan, B.G., Reinisch, W., Rutgeerts, P., Sandborn, W.J., Yan, S., Eisenberg, D. et al. The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. *The American Journal of Gastroenterology* 2007; 102:794-802.

348. Metz, D.C., Vakil, N., Keeffe, E.B., Lichtenstein, G.R. Advances in gastrointestinal pharmacotherapy. [Review] [102 refs]. *Clinical Gastroenterology & Hepatology* 2005; 3(12):1167-1179.

349. Reinisch, W., Sandborn, W.J., Bala, M., Yan, S., Feagan, B.G., Rutgeerts, P. et al. Response and remission are associated with improved quality of life, employment and disability status, hours worked, and productivity of patients with ulcerative colitis. *Inflammatory Bowel Diseases* 2007; 13(9):1135-1140.