NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Ankylosing spondylitis and axial spondyloarthritis (non-radiographic) - adalimumab, etanercept infliximab and golimumab

Enbrel®

Multiple Technology Appraisal (MTA)

Version: Final incorporating addendum
Date originally sent: 3rd September 2014
Date updated: 10th November, 2014

Contents

List of tables ... 3
List of figures ... 9
Abbreviations ... 12
Executive Summary ... 15
1. Background ... 20
1.1 Remit of NICE appraisal ... 22
1.2 Disease background ... 22
1.3 Burden of disease ... 24
1.4 Current management of disease ... 26
1.5 Factors for consideration in anti-TNF therapy ... 32
1.6 Etanercept (Enbrel®) ... 35
2. Systematic review of the clinical literature ... 37
2.1 Identification of studies ... 37
2.2 Study selection ... 38
3. Efficacy and safety of therapy ... 43
3.1 Introduction ... 43
3.2 Summary of efficacy and safety evidence ... 46
3.3 Phase III RCTs : etanercept for the treatment of nr-axSpA ... 53
3.4 Investigator initiated research: etanercept for the treatment of early axSpA ... 62
3.5 Phase III RCTs and related open-label extension studies: etanercept for the treatment of AS ... 72
Etanercept versus sulfasalazine ... 122
3.6 Phase II RCTs and related open-label extension studies: etanercept for the treatment of AS ... 130
3.7 Safety profile of etanercept ... 135
4. Comparative efficacy and safety ... 156
4.1 Summary ... 156
4.2 Efficacy: Meta-analysis and indirect comparisons ... 157
5. Pharmaco-economic Evaluation ... 226
5.1 Summary of cost-effectiveness ... 226
5.2 Design of the economic evaluation ... 227
5.4 Data used in the model ... 261
5.5 Summary of model assumptions ... 291
5.6 Model validation ... 292
5.8 Results ... 299
5.9 Discussion ... 321
6. Wider NHS Implications ... 325

7. References 327
8. Appendix A: 347
8.1 Search strategy for sections 2.1 and 4.2 (clinical literature review and NMA) 347
9. Appendix B 403
9.1 Study 1031- analysis of the primary endpoint based on 2009 EMA guidelines 403
10. Appendix C 404
11. Appendix D 412
12. Appendix E 417
12.1 Studies which met the inclusion criteria of the review but were considered inappropriate to support the NMA 417
12.1 Impact of excluded etanercept studies on AS NMA results 420
13. Appendix F 421
13.1 Additional NMA analyses 421
14. Appendix G 446
14.1 Search strategy for section 5.3.2.2 (Health-related quality of life studies) 446
15. Appendix H: 462
15.1 Search strategy for section 5.3.2.3 (Economic SR) 462
16. Appendix I: 471
16.1 Clinical, utility, resource use and cost inputs for included AS evaluations 471
17. Appendix J: 479
17.1 Quality assessment for section 5.3.2.4 479
18. Appendix K 483
19. Appendix L 485
19.1 Methodological notes 485

List of tables

Table 1: Eligibility criteria used in search strategy 38
Table 2: Summary of main treatment response and remission criteria used in clinical trials of AS and nr-axSpA 49
Table 3: Methodology: Study 1031 53
Table 4: Patient demographics and baseline characteristics (mITT population): Study 1031 56
Table 5: Secondary efficacy outcomes (mITT population): Study 1031 - double-blind period 58
Table 6: Improvement from baseline in QoL measures (mITT population): Study 1031 – double blind period 59
Table 7: Subgroup analysis of the ASAS 40 response in patients with elevated CRP levels or positive MRI at baseline 60
Table 8: Efficacy outcomes at Week 48, open-label period (mITT population): Study 1031 60

Table 9: Methodology: ESTHER trial /Song, 2011 62
Table 10: Patient demographics and baseline disease characteristics: ESTHER trial/ Song, 2011 64
Table 11: Primary efficacy outcome: ESTHER trial/Song, 2011 64
Table 12: Secondary efficacy outcome: ESTHER trial/Song, 2011 65
Table 13: Other clinical outcome parameters: ESTHER trial/Song, 2011 65
Table 14: Methodology and study outcomes: Song, 2012 67
Table 15: Methodology: Song, 2013 69
Table 16: Baseline demographics of patients with nr-axSpA and AS: Song, 2013 70
Table 17: BASDAI, ASAS and ASDAS response rates in patients with nr-axSpA vs patients with AS, Song, 2013 70
Table 18: Summary of Phase III RCTs and related open-label extension studies of etanercept for the treatment of AS 72
Table 19: Methodology: study 314-EU/van der Heijde, 2006 74
Table 20: Patient demographics and baseline characteristics: study 314-EU/van der Heijde, 2006 75
Table 21: Secondary efficacy outcomes: study 314-EU/van der Heijde, 2006 76
Table 22: Methodology study 16.0037/Davis, 2003 78
Table 23: Patient demographics and baseline disease characteristics: study 16.0037/Davis, 2003 79
Table 24: Summary of secondary and other efficacy outcomes (study 16.0037/Davis, 2003) 81
Table 25: Methodology study 16.0040/Davis, 2005 83
Table 26: Additional efficacy outcomes following extended etanercept treatment (study 16.0040/Davis, 2005) 84
Table 27: Methodology: Davis, 2008 86
Table 28: Disease response measures at RCT baseline, OLE baseline and Week 168, Davis, 2008 87
Table 29: Methodology and efficacy outcomes: Boonen, 2008 89
Table 30: Methodology and efficacy outcomes, van der Heijde, 2008 91
Table 31: Patient demographics and baseline characteristics 92
Table 32: Methodology study 311-EU/Calin, 2004 94
Table 33: Patient demographics and baseline characteristics: study 311-EU/Calin, 2004 ... 96
Table 34: Summary of secondary efficacy outcomes, study 311-EU/Calin, 2004 97
Table 35: Comparison of acute phase reactants and spinal mobility values at baseline and at Week 12, study 311-EU/Calin, 2004 98
Table 36: Methodology: study 312-EU/Dijkmans, 2009 99
Table 37: Secondary efficacy outcomes: study 312-EU/Dijkmans, 2009 100
Table 38: mSASSS scores for cervical and lumbar spine: study 312-EU/Dijkmans, 2009 . 101
Table 39: Methodology: study 907-EU/Martin-Mola, 2010 102
Table 40: Primary efficacy outcomes: study 907-EU/Martin-Mola, 2010 103

Table 41: Secondary efficacy outcomes: study 907-EU/Martin-Mola, 2010 104
Table 42: Methodology: SPINE trial/Dougados, 2011 105
Table 43: Patient demographics and baseline characteristics: SPINE trial/ Dougados, 2011 107
Table 44: Absolute changes from baseline to Week 12 in clinical parameters: SPINE trial/Dougados, 2011 108
Table 45: Secondary efficacy outcomes (responder criteria): SPINE trial/Dougados, 2011108
Table 46: Methodology and efficacy outcomes: Dougados, 2012..... 110
Table 47: Absolute changes from baseline to Week 24 in clinical parameters: Dougados, 2012..... 111
Table 48: Secondary efficacy outcomes (responder criteria): Dougados, 2012 112
Table 49: Methodology and efficacy outcomes: Brandt, 2003 113
Table 50: Methodology and efficacy outcomes: Brandt, 2005 116
Table 51: Methodology: Baraliakos, 2013 118
Table 52: Patient demographics and baseline characteristics: Baraliakos, 2013 119
Table 53: Secondary efficacy outcomes (absolute changes from baseline to Year 7 in clinical parameters), Baraliakos, 2013 120
Table 54: Secondary efficacy outcomes, Baraliakos, 2013 121
Table 55: Methodology study 402-WW/Braun, 2011 122
Table 56: Patient demographics and baseline characteristics: study 402-WW/Braun, 2011 124
Table 57: Summary of secondary efficacy outcomes: study 402-WW/Braun, 2011 126
Table 58: Methodology: Giardina, 2010 127
Table 59: Patient demographics and baseline characteristics: Giardina, 2010 128
Table 60: Methodology and efficacy outcomes: study 16.0626/Gorman, 2002 130
Table 61: ASAS 20 results at open-label baseline and Weeks 20, 28 and 40 132
Table 62: ASAS 50 and ASAS 70 results at open-label baseline and Weeks 20, 28 and 40 133
Table 63: Additional efficacy results at Week 40 133
Table 64: Summary of most commonly reported AEs and discontinuations due to AEs in patients receiving etanercept in clinical trials..... 139
Table 65: Summary of SAEs resulting in discontinuations in patients with AS receiving etanercept in clinical trials 144
Table 66: EAMs occurring in AS patients receiving etanercept 145
Table 67: Methodology and safety outcomes: Sieper, 2010 147
Table 68: Methodology and safety outcomes: Gottlieb, 2011 149
Table 69: Methodology and safety outcomes: van der Heijde, 2014..... 152
Table 70: Characteristics of etanercept and comparator treatment RCTs in the AS population eligible for inclusion in the NMA 160
Table 71: Baseline disease status in etanercept and comparator treatment RCTs in the AS population eligible for inclusion in the NMA 167

Table 72: Characteristics of etanercept and comparator treatment RCTs in the nr-axSpA population eligible for inclusion in the NMA 172

Table 73: Baseline disease status in etanercept and comparator treatment RCTs in the nr-axSpA population eligible for inclusion in the NMA..... 176

Table 74: Patients with evidence of nr-axSpA symptoms at baseline in etanercept study 1031 and adalimumab ABILITY-1 trials 179

Table 75: Comparison of MRI sacroiliitis and elevated CRP levels between the three key nr-axSpA trials..... 180

Table 76: AS studies - CRP outcome..... 184

Table 77: AS studies - BASFI outcome 184

Table 78: AS studies - BASDAI outcome 186

Table 79: nr-axSpA studies – ASDAS-CRP outcome 187

Table 80: nr-axSpA studies – BASFI outcome 188

Table 81: nr-axSpA studies – BASDAI outcome..... 189

Table 82: Mean differences in BASFI score among drugs used to treat AS patients (with imputation of SEs)..... 192

Table 83: Mean differences in BASDAI score among drugs used to treat AS patients (with imputation of SEs)..... 193

Table 84: Mean differences in CRP score among drugs used to treat AS patients (with imputation of SEs)..... 194

Table 85: Mean differences in BASFI score among drugs used to treat nr-axSpA patients (with imputation of SEs) 195

Table 86: Mean differences in BASDAI score among drugs used to treat nr-axSpA patients (with imputation of SEs) 195

Table 87: Summary Baseline Characteristics 203

Table 88: Summary of analyses 208

[redacted] 213

Table 90: Final models for BASDAI, restricted to variables adjusted in MAIC 214

Table 91: Final models for BASFI, restricted to variables adjusted in MAIC 215

Table 92: Comparison of BASDAI for placebo arms of etanercept and ABILITY-1 Trials .. 216

Table 93: Comparison of BASDAI for etanercept vs adalimumab based on etanercept and ABILITY-1 trials..... 217

Table 94: Comparison of BASDAI for placebo arms of etanercept and RAPID-axSpA Trials 218

Table 95: Comparison of BASDAI for etanercept vs certolizumab pegol based on etanercept 1031 and RAPID-axSpA trials 219

Table 96: Comparison of BASFI for placebo arms of etanercept and ABILITY-1 Trials 220

Table 97: Comparison of BASFI for etanercept vs adalimumab based on etanercept and ABILITY-1 trials..... 221

Table 98: Comparison of BASFI for placebo arms of etanercept and RAPID-axSpA Trials 222

Table 99: Comparison of BASFI for etanercept vs certolizumab pegol based on etanercept and RAPID-axSpA trials. 223

Table 100: Comparators considered within the primary analysis 228
Table 101: Utility data from studies meeting the NICE reference case 240
Table 102: Mapping algorithms reported in the three included economic evaluations 249
Table 103: Summary of included evaluations 254
Table 104: Key drivers of cost-effectiveness from sensitivity analyses reported in published economic evaluations 259
Table 105: Baseline characteristics of patients sampled in model 261
Table 106: AS patient characteristics in registry data and study 314-EU 262
Table 107: Change in BASDAI and BASFI from baseline at Week 12 in a nr-axSpA population (n=103) 264
Table 108: Change in BASDAI and BASFI from baseline at week 12 in AS population (n=294) 265
Table 109: Relative effects from NMA for nr-axSpA, mean change (with imputation): fixed effects 266
Table 110: Relative effects from NMA for AS, mean change (with imputation): random effects 266
Table 111: MAIC data used in analyses adjusting for differences in study baseline 268
Table 112: Change in BASDAI and BASFI from baseline at Week 48 in nr-axSpA population (n=97) 270
Table 113: Change in BASDAI and BASFI from baseline at Week 48 in AS population (n=43) 271
Table 114: Summary of progression assumptions by study (all converted to 0–100 scale) 274
Table 115: Model fit statistics for treatment discontinuation in nr-axSpA 277
Table 116: Exponential model for time to discontinuation from anti-TNF in the nr-axSpA population 279
Table 117: Exponential model for time to discontinuation from anti-TNF in the AS population 279
Table 118: Assumptions for rebound effects in the previous economic evaluations 280
Table 119: Previously used HRQL algorithms† 282
Table 120: EQ-5D in nr-axSpA-study 1031 (includes all complete EQ-5D observations) .. 285
Table 121: EQ-5D in AS-study 314-EU. (includes all complete EQ-5D observations) 286
Table 122: Drug and administration costs 289
Table 123: Costs of serious infection (using in scenario analysis only) 291
Table 124: Key model assumptions 291
Table 125: Mean baseline and week 12 BASDAI and BASFI in study 1031 for the nr-axSpA population 294
Table 126: Mean baseline and week 12 BASDAI and BASFI in study 314-EU for the AS population 294
Table 127: Results of primary analysis in the AS population 299
Table 128: Incremental results of primary analysis in the AS population 300
Table 129: Results of primary analysis in the nr-axSpA population 300
Table 130: Results of primary analysis in the nr-axSpA population 301

Table 131: Results of etanercept vs certolizumab pegol in AxSpA (using MAIC data)..... 302
Table 132: Incremental results of etanercept vs certolizumab pegol in AxSpA (using MAIC data) 302
Table 133: Results of etanercept vs adalimumab in nr-axSpA (using MAIC data) 303
Table 134: Incremental results of etanercept vs adalimumab in nr-axSpA (using MAIC data) 303
Table 135: Results of analysis including alternative sequential pathways in the AS population 304
Table 136: Results of analysis including Alternative sequential pathways in the nr-axSpA population 304
Table 137: Summary of scenario analysis results in AS population..... 305
Table 138: Scenario analysis results in AS population: Costs 307
Table 139: Scenario analysis results in AS population: QALYs 309
Table 140: Summary of scenario analysis results in nr-axSpA population..... 311
Table 141: Scenario analysis results in nr-axSpA population: Costs 312
Table 142: Scenario analysis results in nr-axSpA population: QALYs. 313
Table 143: Budget Impact calculations for use of etanercept in nr-axSpA patients over 5 years..... 326
Table 144: Proportion of patients achieving ASAS 40 in the open-label period of Study 1031 (mITT population) based on 2009 EMA guidelines..... 403
Table 145: Reclassification in ABILITY-1 410
Table 146: Assessment of certolizumab pegol for nr-axSpA 411
Table 147: Bias potential for relevant RCTs in AS population (assessed using data reported in primary publication)..... 412
Table 148 Bias potential for relevant ax-SpA and nr-axSpA RCTs (assessed using data reported in primary publication)..... 415
Table 149: Mean differences in BASFI score among drugs used to treat AS patients (without imputation of SEs)..... 421
Table 150: Mean differences in BASDAI score among drugs used to treat AS patients (without imputation of SEs) 422
Table 151: Mean differences in CRP level among drugs used to treat AS patients (without imputation of SEs)..... 423
Table 152: Mean differences in BASFI score among drugs used to treat nr-AxSpA patients (without imputation of SEs) 424
Table 153: Mean differences in BASDAI score among drugs used to treat nr-AxSpA patients (without imputation of SEs) 425
Table 154: Direct vs indirect comparisons (edge-splitting) for the PLAC-ETN-INF closed loop for change in BASFI score in the AS population (with imputation of SEs)..... 430
Table 155: Mean differences in BASFI score among drugs used to treat AS patients (with imputation of SEs)..... 431
Table 156: Direct vs indirect comparisons (edge-splitting) for the PLAC-ETN-INF closed loop for change in BASDAI score in the AS population (with imputation of SEs)..... 432
Table 157: Mean differences in BASDAI score among drugs used to treat AS patients (with imputation of SEs)..... 433

Table 158: Mean differences in BASFI score among drugs used to treat AS patients (without imputation of SEs) 435

Table 159: Mean differences in BASDAI score among drugs used to treat AS patients (without imputation of SEs) 437

Table 160: Mean differences in CRP among drugs used to treat AS patients (without imputation of SEs) 439

Table 161: Mean differences in BASFI score among drugs used to treat AS patients (with imputation of SEs) 441

Table 162: Mean differences in BASDAI score among drugs used to treat AS patients (with imputation of SEs) 443

Table 163: Mean differences in CRP among drugs used to treat AS patients (with imputation of SEs) 445

Table 164: Studies that met the inclusion criteria of the SR but were not considered appropriate to the NICE reference case (n=35) 459

Table 165: Eligibility criteria used in search strategy 468

Table 166: Clinical, utility, resource use and cost inputs for included AS evaluations 471

Table 167: Clinical, utility, resource use and cost inputs for included nr-AxSpA evaluations 478

Table 168: Quality assessment of AS full publications (n=8) 479

Table 169: Proposed treatment pathways by population in the full sequential use scenarios 484

List of figures

Figure 1: Recommendations for the management of AS according to ASAS/EULAR guidelines 31

Figure 2: SR flow diagram (original search carried out February 2014) 41

Figure 3: SR flow diagram (updated search carried out May 2014) 42

Figure 4: Summary of etanercept clinical trials in patients with nr-axSpA and AS with their relevant publications 45

Figure 5: Complete network of evidence of randomised clinical trials in the AS population 191

Figure 6: Complete network of evidence of randomised clinical trials in the nr-axSpA population* 194

[redacted] 205

[redacted] 206

Figure 9: Current NICE treatment algorithm 229

Figure 10: Discrete event simulation schematic 235

Figure 11: Schematic of the systematic review of HRQoL evidence 237

Figure 12: Schematic of the systematic review of cost-utility evaluations 252

Figure 13: BASDAI levels in patients treated with infliximab or etanercept 272

Figure 14: Illustrative example of week 48 treatment effect assumption assuming example mean difference of 10 at week 12 272

Figure 15: BASDAI and BASFI over 7 years for etanercept patients 273

Figure 16: Treatment cessation in study 1031 by BASDAI 50 response at Week 12 277

[redacted] 278

[redacted] . 283

[redacted] 283

Figure 20: Predicted vs observed EQ-5D in nr-axSpA population (model 4) 287

Figure 21: Predicted vs observed EQ-5D in AS population (model 3) 287

Figure 22: Initial response for etanercept for 10,000 patients (change in BASDAI and BASFI at Week 12 from baseline) in the nr-axSpA population 293

Figure 23: Initial response for etanercept for 10,000 patients (change in BASDAI and BASFI at wk 12 from baseline) in the AS population 294

Figure 24: Discontinuation after week 12 in the AS population 295

Figure 25: Cumulative mean costs and QALYs for etanercept 296

Figure 26: Cost-effectiveness plane for AS population with 95% confidence ellipses 315

Figure 27: Cost-effectiveness plane for AS population with 95% confidence ellipses and mean costs and QALYs 316

Figure 28: Pairwise cost-effectiveness comparisons for AS with mean and 95% confidence ellipses 317

Figure 29: Cost-effectiveness plane for primary analysis for nr-axSpA population with 95% confidence ellipses 318

Figure 30: Cost-effectiveness plane for primary analysis for nr-axSpA population and mean costs and QALYs† 319

Figure 31: Pairwise cost-effectiveness comparisons for nr-AxSpA with mean and 95% confidence ellipses 320

Figure 32: Impact of excluded etanercept studies on the NMA results for AS 420

Figure 33: Network of randomised clinical trials comparing mean change from baseline in BASFI score among AS population (without imputation of SEs) 421

Figure 34: Network of randomised clinical trials comparing mean change from baseline in BASDAI score among AS population (without imputation of SEs) 422

Figure 35: Network of randomised clinical trials comparing mean change from baseline in CRP (mg/dl) among AS population (without imputation of SEs) 423

Figure 36: Network of randomised clinical trials comparing mean change from baseline in BASFI (cm) among nr-AxSpa population (without imputation of SEs) 424

Figure 37: Network of randomized clinical trials comparing mean change from baseline in BASDAI score (cm) among nr-AxSpa population (without imputation of SEs) 424

Figure 38: Network of randomised clinical trials comparing mean change from baseline in BASFI score among AS population (with imputation of SEs) 426

Figure 39: Network of randomised clinical trials comparing mean change from baseline in BASDAI score among AS population (with imputation of SEs) 427

Figure 40: Network of randomised clinical trials comparing mean change from baseline in CRP (mg/dl) among AS population (with imputation of SEs) 428

Figure 41: Network of randomised clinical trials comparing mean change from baseline in BASFI (cm) among nr-axSpA population (with imputation of SEs) 429

Figure 42: Network of randomised clinical trials comparing mean change from baseline in BASDAI score (cm) among nr-axSpA population (with imputation of SEs) 429

Figure 43: Network of randomised clinical trials comparing mean change from baseline in BASFI score among AS population (with imputation of SEs) 430

Figure 44: Network of randomised clinical trials comparing mean change from baseline in BASDAI score among AS population (with imputation of SEs) 432

Figure 45: Network of randomised clinical trials comparing mean change from baseline in BASFI score among AS population (without imputation of SEs) 434

Figure 46: Network of randomised clinical trials comparing mean change from baseline in BASDAI score among AS population (without imputation of SEs) 436

Figure 47: Network of randomised clinical trials comparing mean change from baseline in CRP among AS population (without imputation of SEs) 438

Figure 48: Network of randomised clinical trials comparing mean change from baseline in BASFI score among AS population (with imputation of SEs) 440

Figure 49: Network of randomised clinical trials comparing mean change from baseline in BASDAI score among AS population (with imputation of SEs) 442

Figure 50: Network of randomised clinical trials comparing mean change from baseline in CRP among AS population (with imputation of SEs) 444

Figure 51: Generating stochastic time-to-events (X) using a random number (U) 486

Abbreviations

ADA	Adalimumab
AIC	Akaike Information Criterion
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
AE	Adverse event
AS	Ankylosing spondylitis
ASAS	Assessment of Spondyloarthritis International Society
ASDAS	Ankylosing Spondylitis Disease Activity Score
AS-QoL	Ankylosing Spondylitis quality of life
ASspiMRI-a	Ankylosing Spondylitis spine MRI score for activity
AS-WIS	Ankylosing Spondylitis Work Instability Index
AUC	Area under the curve
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BAS-G	Bath Ankylosing Spondylitis Patient Global Assessment Score
BASMI	Bath Ankylosing Spondylitis Metrology Index
BIC	Bayesian Information Criterion
BSR	British society of Rheumatology
CFB	Change from baseline
CI	Confidence interval
CHMP	Committee for Medicinal Products for Human Use
CON	Conventional care
COSTART	Coding symbols for a thesaurus of adverse reaction terms
CrIs	Credible intervals
CRP	C reactive protein
CSR	Clinical study report
CZP	Certolizumab pegol
DES	Discrete event simulation
DIC	Deviance information criterion
DMARD	Disease modifying anti-rheumatic drug
DSU	Decision Support Unit
DVU	Discovertebral units
ELISA	Enzyme linked immunosorbent assay
EMA	European Medicines Agency
EQ-5D	EuroQoL-5D
EQ-VAS	EuroQoL visual analogue scale
ERG	Evidence Review Group
ESR	Erythrocyte sedimentation rate

ETN	Etanercept
FDA	Food and Drug Administration
GH-RS	General Health Rating scale
GOL	Golimumab
HADS	Hospital Anxiety and Depression Scale
HLA	Human leucocyte antigen
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility values
HUI-3	Health utilities index mark 3
IBD	Inflammatory bowel disease
ICER	Incremental cost-effectiveness ratio
ICUR	Incremental cost-utility ratio
INF	Infliximab
IPD	Individual patient data
ITT	Intention-to-treat
IV	Intravenous
JIA	Juvenile idiopathic arthritis
LOCF	Last observation carried forward
LRiG	Liverpool Reviews and Implementation Group
MAIC	Matching adjusted indirect comparison
MASES	Maastricht Ankylosing Spondylitis Entheses Score
MCII	Minimum clinically important improvement
MCS	Mental component score
MFI	Multidimensional Fatigue Inventory
mITT	Modified intention-to-treat
mNY	Modified New York
MRI	Magnetic resonance imaging
MOS	Medical Outcomes Study (MOS) Sleep Scale
mSASSS	Modified Stoke Ankylosing Spondylitis Spine Score
MTC	Mixed treatment comparison
NA	Not applicable
NHS	National Health Service
NMA	Network meta-analysis
NR	Not reported
nr-axSpA	Non-radiographic axial spondyloarthritis
NSAID	Non-steroidal anti-inflammatory drug
OASIS	Outcomes in Ankylosing Spondylitis International Study
OECD	Organisation for Economic Co-operation and Development

OL	Open-label
OLS	Ordinary least squares
OMERACT	Outcome Measures in Rheumatology
OW	Once weekly
PAS	Patient Access Scheme
PASS	Patient acceptable symptom state
PCS	Physical component score
PGADA	Patient global assessment of disease activity
PP	Per protocol
PSA	Probabilistic sensitivity analysis
PsA	Psoriatic arthritis
PsO	Plaque psoriasis
PSS	Personal Social Services
QALY(s)	Quality-adjusted life year(s)
QoL	Quality of life
RA	Rheumatoid arthritis
RCT	Randomised, controlled trial
RR	Rate ratios
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SEER	Surveillance, Epidemiology and End Results
SEM	Standard error of the mean
SF-36	Short form 36
SIR	Standardised incidence ratio
SMR	Standardised mortality ratio
SpA	Spondyloarthritis
SPARCC	Spondyloarthritis Research Consortium of Canada
SPC	Summary of product characteristics
SR	Systematic review
SSZ	Sulfasalazine
STC	Simulated treatment comparison
TB	Tuberculosis
TEAE	Treatment emergent adverse event
TNF	Tumour necrosis factor
Tx	Treatment
VAS	Visual analogue scale
WPAI	Work Productivity and Activity Impairment Questionnaire
yr	year

Executive Summary

Patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis have painful, debilitating symptoms that can lead to significant functional impairment and deformity, resulting in disability and loss of quality of life

Axial spondyloarthritis (AxSpA) is a heterogeneous group of inflammatory rheumatic disorders. The Assessment of SpondyloArthritis international Society (ASAS) have produced a classification scheme recognising two classes of patients within AxSpA; those with definite radiographic structural evidence of sacroiliitis (ankylosing spondylitis, AS) and a new class of patients who lack definite radiographic structural evidence of sacroiliitis (non-radiographic axial spondyloarthritis, nr-axSpA). The rationale behind the ASAS classification is that nr-axSpA is an earlier stage of AS with shared underlying pathophysiology.

The onset of symptoms for both groups of patients is typically before the age of 30, and frequently begins during early adulthood. Nr-axSpA and AS patients have a similar disease burden, characterised by inflammatory back pain, stiffness and fatigue, that can cause significant long-term functional impairment and disability.

The prevalence of AxSpA (ASAS definition) in the UK is thought to be around 0.3% of the adult population. AxSpA is thought to comprise of approximately equal populations of nr-axSpA and AS (1). Pfizer estimate this is likely to equate to around 2,500 patients in each classification eligible for anti-TNFs in England and Wales.

AxSpA can result in an appreciable economic burden on patients, employers and the healthcare system. It is characterised by a progressive functional deterioration which compounds the direct cost of healthcare resources with further work-related costs due to early retirement and absenteeism.

Anti-TNFs provide effective treatment for severe AS and nr-axSpA

More than a decade ago, etanercept (ETN) was among the first anti-tumour necrosis factors (anti-TNFs) licensed by the EMA for the treatment of AS. Anti-TNFs were subsequently rapidly incorporated into both national and international clinical guidelines for the treatment of AS (2, 3). NICE published Guidance in 2008 (4) and 2011 (5) recommending ETN, adalimumab (ADA) and subsequently golimumab (GOL) as cost effective treatment options in the AS population.

The available evidence demonstrated that the class of anti-TNFs have clinical efficacy in patients who have failed previous NSAID therapy. Furthermore, evidence is beginning to emerge that anti-TNFs may also be able to reduce the risk of radiographic progression in AS, especially in those with early initiation of the treatment (6).

Since the development of the new ASAS classification, trials of anti-TNFs in nr-axSpA have been conducted and these show a broadly comparable efficacy to AS trials. The most recent recommendations from ASAS consequently endorse the use of anti-TNFs across the whole spectrum of AxSpA patients. This means that a previously untreated population of patients with considerable individual and socioeconomic burden of disease can now be diagnosed and treated effectively.

Based on the above points, it is clear that NICE Guidance is needed across the AxSpA spectrum to ensure both populations of patients have optimal access to effective treatment options.

ETN has a unique mode of action with a broad evidence base demonstrating sustained efficacy and safety

- ETN has a unique mode of action in inhibiting TNF. Whilst ETN is often thought to be in the same class of anti-TNF inhibitors as ADA, GOL, infliximab (INF) and certolizumab pegol (CZP), key differences exist in their molecular structures; ETN is a fully human soluble TNF-a receptor p75 Fc fusion protein whereas ADA, GOL and INF are monoclonal antibodies and CZP is a monoclonal antibody fragment.
- ETN is not associated with clinically relevant immunogenicity. Unlike monoclonal antibodies, ETN has not been found to be associated with the production of neutralising anti-drug antibodies and the potential consequences of immunogenicity such as diminishing drug levels and reduced responses to treatment (7, 8).
- ETN's unique mode of action coupled with the lack of evidence for clinically relevant immunogenicity are both important considerations for patient choice and the provision of alternative treatment options in this therapy area.

  ETN has demonstrated sustained safety and clinical efficacy in providing long-term control of severe AS. This is supported by both the extensive experience in UK clinical practice and a broad evidence base including publication of up to seven years continuous treatment with ETN (9). These data provide clear evidence as to the benefit of treating AS patients with ETN over the course of their lifetime.

Etanercept is licensed for nr-axSpA and demonstrates rapid and significant improvements in disease activity, function and MRI evidence of inflammation

ETN was granted Marketing Authorisation by the European Commission on 28 July 2014 for the treatment of adults with severe non-radiographic axial spondyloarthritis with objective signs of inflammation as indicated by elevated C-reactive protein (CRP) and/or magnetic resonance imaging (MRI) evidence, who have had an inadequate response to non-steroidal anti-inflammatory drugs (NSAIDs).

This licence was granted on the basis of a single phase III RCT (1031) that demonstrated significant and robust benefit of ETN (50 mg weekly) compared with placebo in nr-axSpA patients that had failed at least 2 NSAIDs. The primary endpoint of the study was achieved and was supported by a wide range of secondary endpoints. Clinically relevant improvements were demonstrated at week 12, with onset as early as week 2, and were generally sustained through to 48 weeks in the open-label follow-up period.

- ETN treated patients had significantly better outcomes than patients in the PBO group at 12 weeks for ASAS40 ($P<0.01$), BASDAI50 ($P<0.01$) and a greater mean change from baseline for BASFI ($P = 0.016$), and BASDAI ($P = 0.019$).
- ETN demonstrates significant reductions in objective measures of inflammation compared with PBO at 12 weeks including reduction in hsCRP ($P=0.004$), SPARCC Sacroiliac Joint Score ($P <0.001$) and SPARCC MRI Spinal Score ($P = 0.041$).
- At 48 weeks, these clinical benefits continued to improve on the 12 week data throughout the open label follow up for those patients who remained on ETN and for those patients who crossed over from PBO to ETN upon removal of blinding; ASAS40 (51.9% and

44.0%), BASDAI50 (62.9% and 50%), BASDAI (-2.9 ± 0.2 and -3.3 ± 0.2) and BASFI (-1.9 ± 0.2 and -1.9 ± 0.1).

- ETN reduced measures of systemic and skeletal inflammation on MRI for up to 48 weeks of therapy; mean change in SPARCC Sacroiliac Joint Score for pooled ETN group [redacted]) and SPARCC MRI Spinal Score ([redacted]).

The use of ETN for the treatment of patients with nr-axSpA can improve physical function, reduce pain and disease activity, and lessen the level of inflammation in the spine and SI joints.

ETN is recommended in AS patients and delivers sustained clinical improvement across the full AxSpA spectrum

NICE TA143 previously appraised the efficacy and safety of ETN on the basis of three Phase III, double blind RCTs; trials 16.0037 (10), EU311 (11) and 314 (12). Along with supporting phase II evidence, these studies demonstrated that ETN provides a substantial improvement in disease activity and function in AS patients for up to 3 years.

New data submitted in this dossier, provides evidence of maintained benefit in AS patients over 7 years of continuous ETN treatment. Of the 16 patients completing seven years follow-up, 31.3% were in ASAS partial remission and 43.8% showed ASDAS inactive disease (9). Disease related clinical parameters BASDAI, BASFI, BASMI and ASDAS remained low throughout follow-up.

Evidence from the SPINE study demonstrates that ETN is also effective in a population of older AS patients with advanced and severe spinal ankylosis. Significantly greater efficacy was observed for ETN compared to PBO in improving BASDAI, BASFI, total back pain, BASMI, CRP level and forced pulmonary vital capacity (13).

An additional large multicentre double-blind RCT (study 402) comparing ETN (n=379) with sulfasalazine (SSZ) (n=187) has also confirmed the benefit of ETN in AS patients (14). ETN showed statistically significant benefit in the primary outcome of ASAS20 as early as week 2, a difference that was sustained to week 16. Similar significant differences were also seen in ASAS40, ASAS5/6 and ASAS partial remission, as well as BASMI and BASFI.

Lastly, the ESTHER trial compared ETN with SSZ in a population of early AxSpA patients (AS and nr-axSpA with symptom duration of less than five years) who had active inflammatory SI or joint lesions on MRI (15). Results showed that compared with SSZ, ETN is able to significantly reduce the severity of whole body MRI outcomes for SI joint and spine at 48 weeks. Similar significant improvements were shown for patient reported outcomes at week 48 (BASDAI, BASFI and patients and physician global assessments of disease activity). Importantly, an analysis of this trial on the ETN-treated subgroups of AS and nr-axSpA patients showed that, in populations with similar baseline characteristics, ETN is able to provide an equivalent benefit to both AS and nr-axSpA patients (16).

Overall, the new evidence presented in this submission supports the conclusions that ETN provides significant and sustained efficacy across the broad spectrum of the AxSpA population including nr-axSpA and AS patients, even those who are in the advanced stages of the disease.

ETN has an established safety profile

Data from RCTs (966 patients) and open-label extension studies (687 patients) demonstrate that ETN has an acceptable safety profile in patients across the whole disease spectrum of AxSpA.

Adverse events reported were largely mild to moderate in intensity and were generally resolved without discontinuation of therapy. Injection site reactions, infection (e.g. upper respiratory tract infection, bronchitis) and fever are the most commonly reported adverse reactions associated with ETN therapy. The largest pooled safety evidence dataset presented for ETN to date showed that therapy was associated with low rates of serious or opportunistic infections, malignancies, non-melanoma skin cancers and inflammatory bowel disease in patients with active AS (17). Similarly, data from a pooled analysis showed that incidences of uveitis in patients receiving ETN for active AS tended to be mild to moderate in severity (18).

ETN has a significant clinical benefit vs conventional care and comparable efficacy to other anti-TNFs in nr-axSpA and AS,

Network meta-analyses (NMA) of BASDAI, BASFI and, CRP (AS) or ASDAS-CRP (nr-axSpA), were conducted to estimate the relative efficacy of ETN compared with conventional care (NSAIDs), ADA, INF, GOL and CZP in their respective licensed indications.

AS population

- Anti-TNFs demonstrate statistically significant benefit over conventional care and have comparable clinical efficacy. The comparison with conventional care showed a significant benefit for anti-TNFs in the vast majority of sensitivity analyses. Overall, there were very few differences in BASDAI and BASFI between anti-TNFs in AS, supporting the conclusion of comparable efficacy.

Nr-axSpA population

- Consistent with the available trial evidence, the NMA supports the conclusion that anti-TNFs provide a significant benefit over conventional care.
- Extensive heterogeneity exists between the nr-axSpA trial populations. This confounds the comparisons between anti-TNFs, making the results of the NMA potentially unreliable. In the absence of sufficient information for meta-regression, heterogeneity was addressed using matching adjusted indirect comparison (MAIC) and simulated treatment comparison (STC) analyses of patient level data available to Pfizer. The results of these analyses suggest comparable efficacy of ETN with CZP and ADA.

ETN is a cost effective treatment option versus conventional care and has comparable cost effectiveness with other anti-TNFs

A patient-level simulation model was developed, that considers a lifetime time-horizon to reflect the natural history of nr-axSpA and AS. The perspective used in the model was that of the UK National Health Service and Personal Social Services.

In the base-case, the model reported that the incremental cost-effectiveness ratio (ICER) for ETN vs conventional care was £20,938 and £23,195 in the AS and nr-axSpA populations, respectively.

In both AS and nr-axSpA populations, the ICERs of other anti-TNFs versus conventional care were similar to ETN, with the exception of the INF strategy in AS which had an ICER of £37,741 per QALY gained. The probabilistic and deterministic sensitivity analysis demonstrated that the results were robust to parameter uncertainty with ICERs consistently less than £30,000 per QALY gained.

In the AS population, when ETN was compared incrementally with other anti-TNFs there were no differences observed between ICERs, demonstrating that, in this population, anti-TNFs have similar cost effectiveness estimates. However, in the nr-axSpA population differences in the ICERs between ETN, CZP and ADA were observed based on the traditional NMA which did not account for the heterogeneity between the clinical trial populations. When the adjustment is made for this heterogeneity using the MAIC technique, the differences in cost and QALYs are reduced, thus supporting the clinical conclusions of no difference between anti-TNFs and that they are all cost effective options for the treatment of nr-axSpA. This is further supported by the results of the probabilistic sensitivity analysis which showed that when variation in assumptions and clinical data are considered, the ranges of ICERs are consistently overlapping for ETN, CZP and ADA.

Conclusions

Nr-axSpA and AS patients have a similar disease burden with painful and debilitating symptoms that can lead to significant long-term functional impairment, disability and loss of quality of life. AxSpA can result in an appreciable economic burden on patients, employers and the healthcare system.

Anti-TNF agents are the only reliable treatment option in patients who fail to respond or cannot tolerate NSAIDs and this recommendation has been endorsed by the clinical community within the international clinical guidelines for the management of AxSpA.

The clinical evidence presented here demonstrates that ETN provides a unique mode of action for treating patients across the AxSpA spectrum, furthermore ETN is well tolerated and demonstrates rapid and sustained improvements in measures of disease activity, function, inflammation and quality of life.

The results of the economic analysis demonstrate that ETN is a cost-effective treatment option for both the nr-axSpA and AS populations versus conventional care and has comparable cost effectiveness versus other anti-TNFs (with the exception of infliximab).

Overall, this analysis supports the recommendation for ETN use in AxSpA and represents good value for money for the NHS by delivering considerable quality of life benefits to patients who suffer from a severely debilitating condition.

1. Background

The spondyloarthropathies (SpA) are a heterogeneous group of chronic, inter-related rheumatic diseases including ankylosing spondylitis (AS), psoriatic arthritis, inflammatory bowel disease with associated arthritis, reactive arthritis and "undifferentiated" SpA (19). Clinical characteristics include inflammatory back pain (IBP), enthesitis and peripheral arthritis. Patients are classified according to whether they have predominantly axial involvement (axSpA) or peripheral involvement.

- The two subgroup classifications of AxSpA are:
  - Ankylosing spondylitis (AS) – patients present with IBP, impaired spinal mobility, and radiographic sacroiliitis and/or syndesmophytes
  - Non-radiographic axial spondyloarthritis (nr-axSpA) – patients present with IBP but with no definite structural changes in the spine and sacroiliac joints on x-ray
- The UK prevalence is estimated as 0.3% for AxSpA (ASAS criteria) and 0.15% for AS (modified New York criteria). Due to the newer classification of nr-axSpA, there is currently limited data available on the epidemiology of this disease

Burden of disease

- AxSpA is a chronic disease for which there is currently no cure
- Complications of long-standing disease include osteoporosis and vertebral fractures (20)
- Survival rates in patients with AS range from 92-96% at 5 years to 62% at 25 years (21)
- Patients with nr-axSpA or AS can experience increasing levels of physical and mental disability as their disease progresses, impacting on their quality of life (22, 23)
- Patients with AxSpA often have difficulty maintaining their employment or have to take early retirement due to their disease (24)
  - AS and nr-axSpA patients miss 1.0–3.4 work days per month due to their disease (25)
  - On average, an AxSpA patient experiences 6–7 days reduced productivity per month (25)
  - The most significant factors associated with absenteeism from work are level of disease activity and depression (26)

Current management of disease

- Recent Assessment of Spondyloarthritis international Society (ASAS) criteria classify AxSpA based on both clinical (in patients with no evidence of abnormalities on radiograph or MRI) or imaging assessments (in patients with sacroiliitis on radiograph or MRI) (27)
  - Prior to the new classification system, many patients had to wait for 8–11 years

between the onset of symptoms and the development of radiographic sacroiliitis that allowed them to be diagnosed as AS and receive treatment (28)

- The key aims of treating patients with AxSpA are to improve long term health related quality of life through the control of signs and symptoms, prevention of structural damage, preservation or normalisation of function and minimisation of comorbidities (29)

- Non-steroidal anti-inflammatory drugs (NSAIDs) are conventionally used for the treatment of AxSpA as the first treatment option in combination with exercise and physiotherapy (30)

- Anti-tumour necrosis factor (anti-TNF) agents are currently the only biologic treatment option for severe AxSpA patients with persistently high disease activity or those failing to respond to NSAIDs
  - ASAS recommendations state that patients must have tried ≥2 NSAIDs for a minimum of 4 weeks in total before anti-TNF agents can be started
  - Possible predictors of a positive response to treatment with anti-TNF agents in patients with AxSpA include younger age, functional status, higher CRP level and presence of sacroiliitis on MRI at baseline at initiation of therapy (31)
  - There is emerging evidence that early treatment with anti-TNF treatment can slow radiographic progression in patients with AS (32)
  - According to NICE recommendations, patients receiving anti-TNF agents for the treatment of AxSpA should be assessed for their response to treatment 12 weeks after starting therapy (4)
    - An adequate response to treatment is defined as a reduction in the BASDAI of at least 50% or 2 units, and a reduction in the spinal pain visual analogue score (VAS) ≥2cm
  - Data from randomised controlled trials and from longer term extension studies have demonstrated that anti-TNF therapies have an acceptable and tolerable safety profile although caution with initiation of treatment is still recommended

Etanercept

- Etanercept (Enbrel®) is a recombinant, fully human TNF receptor p75 Fc fusion protein which competitively inhibits the binding of TNF to its cell surface receptors preventing initiation of cell signalling and thereby inhibiting its biological activity (8)
  - etanercept received its first approval in November 1998 from the US Food and Drug Administration, for the treatment of patients with moderately to severely active rheumatoid arthritis (RA) who had an inadequate response to one or more disease-modifying anti-rheumatic drugs (DMARDs) (33)
  - in 2000, etanercept was approved by the European Medicines Agency for the treatment of RA
  - etanercept has since been approved for clinical use in over 90 countries. In the EU, etanercept is approved for the treatment of RA, AS, juvenile idiopathic arthritis,

psoriatic arthritis, adult plaque psoriasis, and paediatric psoriasis (8)

- etanercept was granted a licence for the treatment of nr-axSpA by the European Medicines Agency on 28th July 2014

Unmet need

- Current management of AxSpA focuses on the use of NSAIDs in the first-line
- The nr-axSpA population is largely regarded as an earlier stage of the AxSpA disease spectrum
- The burden of disease in the nr-axSpA population is similar to the AS disease burden
- Etanercept offers an effective treatment option for patients failing to respond to NSAIDs
- The absence of any recommendation for the use of anti-TNF agents in the nr-axSpA population represents a clear unmet need for this newly defined group of patients who previously had to wait for close to a decade to receive a diagnosis and effective treatment
- NICE guidance would address this unmet need by providing patients with access to effective treatments

1.1 Remit of NICE appraisal

NICE are conducting a Multiple Technology Appraisal "to appraise the clinical and cost-effectiveness of TNF inhibitors within their licensed indications for treating ankylosing spondylitis and axial spondyloarthritis without radiographic evidence of ankylosing spondylitis". The purpose of this submission document is to provide a balanced summary of evidence for the clinical and cost-effectiveness of etanercept (Enbrel®, ETN), a soluble tumour necrosis factor (TNF) receptor, compared with the alternative TNF-α inhibitors, addressed in the NICE scope: adalimumab (Humira®), certolizumab pegol (Cimzia®), golimumab (Simponi®) and infliximab (Remicade®).

1.2 Disease background

1.2.1 Ankylosing spondylitis and non-radiographic axial spondyloarthritis

The spondyloarthropathies are a heterogeneous group of inter-related, rheumatic diseases which include ankylosing spondylitis (AS), psoriatic arthritis, arthritis with related inflammatory bowel disease, reactive arthritis, and "undifferentiated" spondyloarthritis (SpA) (19). Main, characteristic clinical features of SpA include inflammatory back pain (IBP) and extra spinal manifestations such as enthesitis (inflammation at the sites where ligaments and tendons attach to bone), dactylitis and peripheral arthritis (34). SpA with predominantly peripheral involvement is classified as peripheral SpA while disease with predominantly spinal involvement is classified as axial spondyloarthritis (axSpA). In addition to the joint related symptoms of AxSpA, patients may also have extra-articular manifestations (EAMs) of the disease. The main EAMs include:

- Acute anterior uveitis, which involves inflammation of the uvea and affects 25–30% of patients (35). Attacks are mainly unilateral and often resolve spontaneously within 3

months, but are frequently recurrent (35). The condition occurs more frequently in patients who are human leucocyte antigen B27 (HLA-B27) positive (36).

- Psoriasis, which is seen in 10–25% of patients (35)
- Inflammatory bowel disease, including ulcerative colitis and Crohn's disease, which afflicts 5–10% of patients (35).
- Cardiovascular dysfunction, in particular aortic insufficiency (also known as aortic regurgitation), affects 1–10% of patients, usually with established or advanced disease (35). Other cardiovascular manifestations include atrioventricular block (37) and cardiac arrhythmias (38).
- Pulmonary complications, which usually manifest as either chest wall restrictions, fibrosis of the upper lobes or interstitial lung disease (39).

There are two sub-classifications of AxSpA; ankylosing spondylitis (AS) and non-radiographic axial spondyloarthritis (nr-axSpA). Patients with AS typically present with IBP, impaired spinal mobility, and radiographic sacroiliitis and/or syndesmophytes (according to the modified New York [mNY] criteria, see section 1.4.1) (19, 40). Patients with nr-axSpA present with IBP and may have sacroiliac joint pain. However, while these patients will have no definite radiographic structural changes in the spine and sacroiliac joints, they may or may not have evidence of inflammatory changes on MRI scans (41). Some patients with nr-axSpA will experience disease progression over time, eventually progressing to AS. Progression of the disease varies and may be rapid or may take several years (34). Other patients may only develop mild sacroiliitis and may not experience further disease progression (42) while some patients may go into remission (43, 44). Both patient sub-classifications may also present with peripheral SpA features.

1.2.2 Epidemiology of disease

AS is approximately three times more common in males than in females (35). Patients usually present with the first symptoms before the age of 30 years (19) and 90–95% of patients are positive for the HLA-B27 gene (20). At present, there is very little information available on the epidemiology of AS and nr-axSpA in the UK. However, a recent estimate based on a UK cohort study of inflammatory back pain calculated the prevalence of AxSpA (ASAS criteria) as 0.3% and AS (mNY criteria) as 0.15% in the general UK population (45). This equates to approximately 26.7 per 100,000 in the population of England and Wales.

A recent systematic review conducted to estimate the global prevalence of AS reported that the mean prevalence of AS in Europe was 23.8 per 10,000 of population. This prevalence was estimated based on the evidence from nine cross-sectional population-based studies conducted in France, Finland, Greece, Hungary, Italy, Lithuania and Turkey (46).

A retrospective study conducted in the US used clinical data from patients in representative rheumatology practices to estimate the prevalence of AxSpA (including AS and nr-axSpA) based on the ASAS criteria. The projections of the study estimates a national prevalence of 0.35% (0.18–0.554%) for AS and 0.35% (0.18–0.554%) for nr-axSpA (47). A similar ratio for nr-axSpA to AS of approximately 1:1 was reported in a recent study among SpA patients conducted in Germany; 52.2% of patients were diagnosed with AS and 47.8% of patients with nr-axSpA (48). The probabilities of AS and nr-axSpA were almost equal if the duration of back pain was between 1 and 6 years (1–3 years: 47.5% and 52.5%, respectively; 3–6 years: 46.3% and 53.7%, respectively) (48).

1.3 Burden of disease

1.3.1 Impact on the patient

1.3.1.1 Morbidity and mortality

AxSpA is a chronic disease for which there is currently no cure. The predominant symptom in AxSpA is back pain which is inflammatory in origin. Continual inflammation of the spine results in bone remodelling, leading to erosions, osteoporosis and the development of new bone formation and bony bridges (syndesmophytes) with consequent ankylosis, although not all patients will develop syndesmophytes. Osteoporosis occurs in patients as a result of decreased bone mineral density (20). Another complication of long-standing disease is vertebral fracture. However, in many cases, the presence of vertebral fractures goes unrecognised resulting in delay of diagnosis and further complications (20). The ongoing inflammation and increase in disease activity may result in loss of physical function and quality of life and long-term disability for the patient (19, 49).

A common symptom among AS patients is fatigue. A recent qualitative study to examine the causes and management of fatigue in patients with AS reported that fatigue negatively impacted on daily living, relationships and social interactions and employment (50). The majority of the patients reported that fatigue had not been present before their diagnosis of AS while some patients reported a higher level of fatigue when their pain levels were greater (50). Patients with nr-axSpA also have high levels of disease activity, pain and fatigue similarly to AS patients. A cross sectional analysis of 226 patients with nr-axSpA and 236 patients with AS in the German Spondyloarthritis Inception Cohort (GESPIC) reported that there were no differences in the levels of disease activity (BASDAI) and inflammatory back pain between patient groups (51). Patients in both groups reported equally high levels of general pain, nocturnal pain and fatigue (51).

Many patients with AS will also experience episodes of higher disease activity followed by asymptomatic periods or periods with less severe symptoms of disease. These fluctuations of symptoms are known as “flares”. An earlier UK based study reported that 96% of patients reported experiencing flares since their disease onset, with 29% of patients reporting flares lasting for several months (52). More than 70% of patients reported experiencing constant disease symptoms (i.e. pain and stiffness) between flares and this patient group had significantly worse BASDAI and BASFI scores compared with patients reporting intermittent symptoms (52).

Compared with the general population and independent of treatment strategies, AS patients have an increased risk of mortality. Survival rates in AS populations range from 92–96% at 5 years, 84–96% at 10 years, to 62% at 25 years (21). Cardiovascular dysfunction affects 1–10% of patients with AS (35) and an earlier systematic review reported that circulatory diseases were the acute cause of death in 18-50% of patients with AS (21). There is currently no mortality data available for patients with nr-axSpA.

1.3.1.2 Quality of life

AxSpA has a considerable negative impact on the health status and quality of life (QoL) of patients. In addition, due to the fact that the onset of disease is usually before the third decade of life, the extent of the overall lifetime burden on the patient is much greater. As their disease progresses, patients with nr-axSpA or AS can experience increasing levels of

physical and mental disability. This can impact on activities of daily living, social interactions and employment (22, 23).

Studies evaluating the impact of AS on patient QoL report that the disease has a greater impact on measures of physical function, such as limitation of role due to physical health and bodily pain, compared with other QoL domains. A cross-sectional study of patients with AS in the UK reported that the disease was associated with poor functional status and impaired QoL (24). Patient physical function as measured by the short-form 36 (SF-36) questionnaire (see section 3.2.1, Table 2) was found to be reduced by at least 35% in the AS population compared with the normal UK population, with limitation of role due to physical health being the worst affected. Patients also reported lower mental health component scores compared with the control population scores matched for age and sex (24). Similarly, an Italian study compared the health related quality of life (HRQoL) scores of 799 patients with inflammatory rheumatic diseases (including 164 patients with AS) with 1,579 healthy age-matched controls. Using the SF-36 questionnaire, it was reported that the most negatively affected domains in patients with AS were physical function (role function) and bodily pain (53).

Patients with nr-axSpA also have impaired physical function. Kiltz et al, 2012 reported that one third of patients had a BASFI score ≥3 (range 0–10), with the higher values indicating greater impairment of physical function (54). The study found no significant differences in physical functioning as measured by the BASFI between nr-axSpA patients and those patients with established AS (54).

1.3.2 Societal and healthcare burden

AxSpA is associated with a considerable economic burden due to direct healthcare costs (i.e. healthcare resource utilisation) and indirect costs (e.g. as a result of lost productivity). Due to the relatively young age of onset of AxSpA, healthcare costs are incurred from an early age. Consequently, the total costs over the lifetime of the patient can be significant. A UK study which investigated the direct healthcare resources utilised by AS patients attending a secondary care rheumatology unit reported that the mean annual cost per patient was £1,852 (55). A significant correlation was reported between the total annual cost per patient and disease activity (BASDAI, $p<0.001$) and functional disability (BASFI, $p<0.001$). The greatest cost drivers were reported as physiotherapy costs (32% of the total costs), hospitalisation costs (21% of the total costs) and drug costs (20% of the total costs). With respect to medication costs in this study, none of the patients were receiving anti-TNF agents; 85% of patients were being prescribed NSAIDs, while 34% of patients were receiving disease modifying anti-rheumatic drugs (DMARDs; see section 1.4.2.3) (55). Similar findings have been reported in Europe. Boonen et al, 2003 reported that the mean annual direct costs per patient with AS in Belgium, France and The Netherlands were €941, €1,402, and €2,122, respectively (56). The greatest overall cost driver was reported as inpatient care (including hospitalisation) which accounted for 27% of the total direct costs. Higher costs were related to greater physical impairment, higher disease activity and longer disease duration (56).

AxSpA has a significant impact on a patient’s ability to perform usual daily activities, including paid employment, and thus is associated with significant indirect costs. As the onset of AxSpA symptoms typically manifest in early adulthood, these indirect costs are incurred early on in the patient’s life. The total indirect costs are further compounded due to the manifestation of the disease throughout most of or over the entire adult life of the patient. A cross-sectional study of 246 patients with AS in the UK reported that while 84% of patients

were of working age, 19% of patients had lost their job due to their disease and 10% of patients had taken early retirement (24). In addition, 49% of patients in employment had severe active disease (BASDAI score >4) and were found to have a significantly greater risk of work instability, defined as risk of impending job loss, ($p<0.001$), as measured by the Work Instability Scale (WIS) (24). Similarly, a multicentre survey conducted among 612 patients in the UK with a diagnosis of AS reported that 40% of patients were unemployed while the most significant factors associated with absenteeism from work were higher levels of disease activity (BASDAI >4) and depression (26). A recent UK study of 512 AS patients estimated the total annual work-related costs of AS to be £11,943 per patient as a result of early retirement, absenteeism and presenteeism. The annual cost of early retirement due to AS was estimated as £8,107 per patient and was significantly higher in AS patients with BASDAI or BASFI scores ≥4 (57).

AS patients in the RAPID-axSpA study reported missing 1.0–2.4 work days per month due to their disease while the mean number of days where work productivity was reduced by ≥50% was reported as 3.8–5.4 days per month (25). As reported for AS patients, nr-axSpA is also associated with negative effects on daily home and work life. Baseline data reported in the RAPID-axSpA study, showed that the burden of disease on productivity in the workplace in patients with nr-axSpA was generally comparable with the burden in AS patients; the mean number of work days lost per month was 1.9–3.4 and the mean number of days with work productivity reduced by ≥50% was 5.6–6.5 (25). In addition, more than one third of nr-axSpA patients required regular assistance from relatives, friends or professional caregivers in their daily activities due to their disease (25)

1.4 Current management of disease

1.4.1 Diagnosis and measurement of disease state

Until recently, the most widely accepted classification criteria for the diagnosis of AS was the mNY criteria (40). On the basis of these criteria, a definite diagnosis of AS was based on evidence of grade 2 bilateral or grade 3–4 unilateral radiological sacroiliitis together with at least one clinical criterion (IBP, limited mobility of the lumbar spine or reduction in chest expansion) (40). However, in many patients with AS, it can be several years after the onset of symptoms before the appearance of radiographic sacroiliitis resulting in a delay in diagnosis of 8–11 years (28). Also, as discussed in section 1.2.1, some patients with nr-axSpA may never progress to AS. In addition, several studies have demonstrated that acute inflammation of the sacroiliac joints and/or the spine as evidenced by magnetic resonance imaging (MRI) does develop prior to the appearance of visible structural changes on conventional radiographs (42, 58). MRI is now regarded as a valuable diagnostic tool in the assessment of patients with AxSpA.

In 2009, the Assessment of Spondyloarthritis international Society (ASAS) developed new classification criteria in order to enable more reliable and accurate diagnosis of AxSpA earlier (27, 59). The advent of the new ASAS criteria allows patients with AxSpA to be identified and treated in the early stages of the disease before significant structural damage has occurred.

Classification of AxSpA according to ASAS criteria is based on both clinical and imaging assessments. According to the ASAS criteria, a patient presenting with chronic back pain of duration ≥3 months with age at onset <45 years can be classified as having AxSpA if there is evidence of sacroiliitis (radiographic or MRI) together with at least one other SpA feature

(e.g. IBP, extra spinal manifestations [heel enthesitis, uveitis, dactylitis, arthritis, psoriasis, Crohn's/ulcerative colitis],a family history of SpA, HLA-B27, good response to NSAIDs and elevated CRP or erythrocyte sedimentation rate [ESR]) (27). In the absence of imaging evidence of sacroiliitis, patients must be HLA-B27 positive and present with at least two other SpA features (27). The new criteria encompasses both patients with evidence of established radiographic sacroiliitis (who will be diagnosed with AS) and also those patients who have not yet developed radiographic sacroiliitis (27). Patients who satisfy the ASAS classification criteria for the diagnosis of AxSpA but present with no definite radiographic structural changes in the spine and sacroiliac joints are classified as having nr-axSpA.

1.4.2 Current therapies

According to recommendations from an international treat-to-target task force from 2011, the primary aims of treating patients with any form of SpA are to improve long term HRQoL which may be achieved through the control of signs and symptoms, prevention of structural damage, preservation or normalisation of function and minimisation of comorbidities (29).

In addition, the most recent guidelines for the management of AS published by ASAS and the European League Against Rheumatism (EULAR) recommend that optimal management of AS should include both pharmacological and non-pharmacological interventions/therapies (30). The overarching principles of the ASAS/EULAR guidelines are broadly in line with those recommendations published by the aforementioned international task force.

1.4.2.1 Non-pharmacological treatments

Non-pharmacological therapies are recommended by the ASAS/EULAR guidelines for the optimal management of patients with AS. This includes patient education and regular exercise (30). Due to the nature of the disease, exercise is recommended in patients with AS to improve or maintain spinal mobility. While home exercises can be effective, physical therapy involving supervised exercises is preferable as this is more effective (30). ASAS/EULAR guidelines also recommend the use of patient associations and self-help groups (30). Non-pharmacological therapies are recommended alongside NSAIDs.

1.4.2.2 NSAIDs

The conventional therapy for the treatment of AxSpA is non-steroidal anti-inflammatory drugs (NSAIDs). ASAS/EULAR guidelines recommend NSAIDs as the first-line therapy for patients with AS (30). In addition, international ASAS recommendations state that patients should have failed at least two NSAIDs for a minimum of four weeks in total prior to starting treatment with anti-TNF agents (see section 1.4.2.5) (60). Similarly, in the UK, NICE technology appraisal guidance 143 and 233 state that these same criteria must be fulfilled before etanercept, adalimumab or golimumab can be recommended as treatment options for patients with severe active AS (4). Although there are currently few published studies demonstrating the efficacy of NSAIDs for the treatment of nr-axSpA, it is expected that the general treatment outcomes are similar to those of AS due to the common symptoms and overlapping disease burden between both disorders (43).

NSAIDs can be effective in reducing pain and stiffness in many patients with axial symptoms. More than 60% of patients receiving NSAIDs for the treatment of AS experience a clinically significant improvement in back pain (61). There is also evidence that continuous use of NSAIDs when compared with NSAID treatment on demand is associated with a reduction in radiographic progression in patients with AS over a 2 year period (62). However,

a recent analysis of 2 year data from the GESPIC Cohort reported no significant differences in the effects of high or low doses of NSAIDs on radiographic progression in patients with nr-axSpA (63).

Continuous treatment with NSAIDs is recommended for patients with persistent active, symptomatic disease (30). The dose of NSAIDs must be adjusted according to the severity of the patient's symptoms. While a moderate dose is sufficient to achieve a therapeutic effect in some patients with AS, many patients require the maximum tolerated therapeutic dose (64). If a patient fails to respond to initial NSAID therapy, it is recommended that they try at least one other different NSAID (43).

NSAID use also carries the risk of serious side effects such as gastrointestinal toxicity (including gastrointestinal bleeding and ulcer), renal and cardiovascular risks (43). Approximately 1% of patients treated with a maximum tolerated dose of NSAIDs will experience serious adverse events (64). Thus the safety risks of NSAID treatment should always be taken into account before commencing therapy (30, 43).

In addition to the above there are several other disadvantages to using NSAIDs in the long-term, including (30):

- High inter-patient variation in treatment response
- Limited data currently available to determine the effect of NSAIDs on structural damage in nr-axSpA
- NSAID use in the elderly has been associated with increased risk of adverse events, particularly gastrointestinal complications such as perforations and gastrointestinal bleeding (65, 66)

1.4.2.3 DMARDs

The ASAS/EULAR recommendations state that there is currently limited evidence that disease modifying anti-rheumatic drugs (DMARDs; such as methotrexate and sulfasalazine [SSZ]) are effective for the treatment of AxSpA (30). However, SSZ may be beneficial for the treatment of early AS with IBP as the primary clinical symptom or in patients with peripheral disease (30, 67).

1.4.2.4 Glucocorticoids

Local glucocorticoid injections directed at the site of inflammation may be considered for treatment of peripheral disease (e.g. arthritis, dactylitis, and enthesitis). They may also be effective for the treatment of active sacroiliitis in AxSpA (30, 43). There is currently no evidence in support of the use of systemic glucocorticoids for the treatment of AxSpA (30).

1.4.2.5 Anti-TNF agents

For those patients with persistently high disease activity who fail to respond to NSAIDs, an anti-TNF agent is currently the only biologic treatment option available at the second-line. Studies have demonstrated that anti-TNF therapies are an effective treatment option for reducing the symptoms of AS and nr-axSpA (14, 68-71). There is now evidence emerging that anti-TNF treatment halts radiographic progression in AS patients. A recent study to assess the effect of anti-TNF agents on progressive spinal damage in AS patients reported that patients receiving anti-TNF treatment had a 50% reduction in the odds of progression compared with patients not receiving anti-TNF agents (odds ratio: 0.52; 95% confidence interval: 0.30-0.88; p=0.02) (32). Also, the odds of progression were significantly lower in

patients taking anti-TNF agents for ≥50% of their disease duration compared with patients who were not (odds ratio: 0.2 95% confidence interval: 0.04-0.92; p=0.04) (32). In addition, the rate of modified Stokes Ankylosing Spondylitis Spine Score (mSASSS) progression was shown to increase with an increasing delay in starting treatment demonstrating that earlier treatment in AS is associated with a reduction in radiographic progression (32). Similarly, Baraliakos et al, 2014 reported a significant difference in the mean mSASSS score (p=0.047) and a significantly lower number of new syndesmophytes per patient in AS patients receiving infliximab treatment over 8 years compared with control patients (p=0.007) (72). The findings from these two studies suggest that anti-TNF treatment of nr-axSpA patients may slow the onset of radiographic progression and provides some circumstantial evidence for the rationale of treating this patient population at an earlier stage of the disease.

According to the most recent update of the ASAS recommendations for the treatment of AxSpA with anti-TNF agents (60), patients with AS or nr-axSpA may be considered for anti-TNF therapy providing they fulfil the following criteria: have a definite diagnosis of AxSpA (i.e. fulfilling either the mNY criteria for AS (40) or the ASAS criteria for AxSpA (27)); have high disease activity (defined as a BASDAI score ≥4 [see section 3.2.1, Table 2]); and have failed treatment with ≥2 NSAIDs over at least 4 weeks in total at the maximum recommended or tolerated dose, unless contraindicated. For AxSpA diagnosis and assessment of response to anti-TNF agents, patients require a positive opinion based on clinical examination, assessment of acute phase reactants, and radiographic progression or positive MRI (43, 60).

There are currently five anti-TNF agents licensed for the treatment of AS: the monoclonal antibodies adalimumab (Humira®), golimumab (Simponi®), and infliximab (Remicade®), the monoclonal antibody fragment certolizumab pegol (Cimzia®), and the human TNF receptor p75 Fc fusion protein, etanercept (Enbrel®). Adalimumab and certolizumab pegol are also licensed for the treatment of adults with severe axial spondyloarthritis without radiographic evidence of ankylosing spondylitis but with objective signs of inflammation (including elevated C-reactive protein (CRP) and/or positive MRI), whose disease has responded inadequately to, or who are intolerant to NSAIDs (73, 74).

1.4.2.6 Analgesics

Analgesics (e.g. paracetamol and opioid-like therapies) may be considered for the relief of residual pain following the failure of previously recommended treatments due to lack of efficacy or poor tolerance or when recommended therapies are contraindicated (30).

1.4.2.7 Surgery

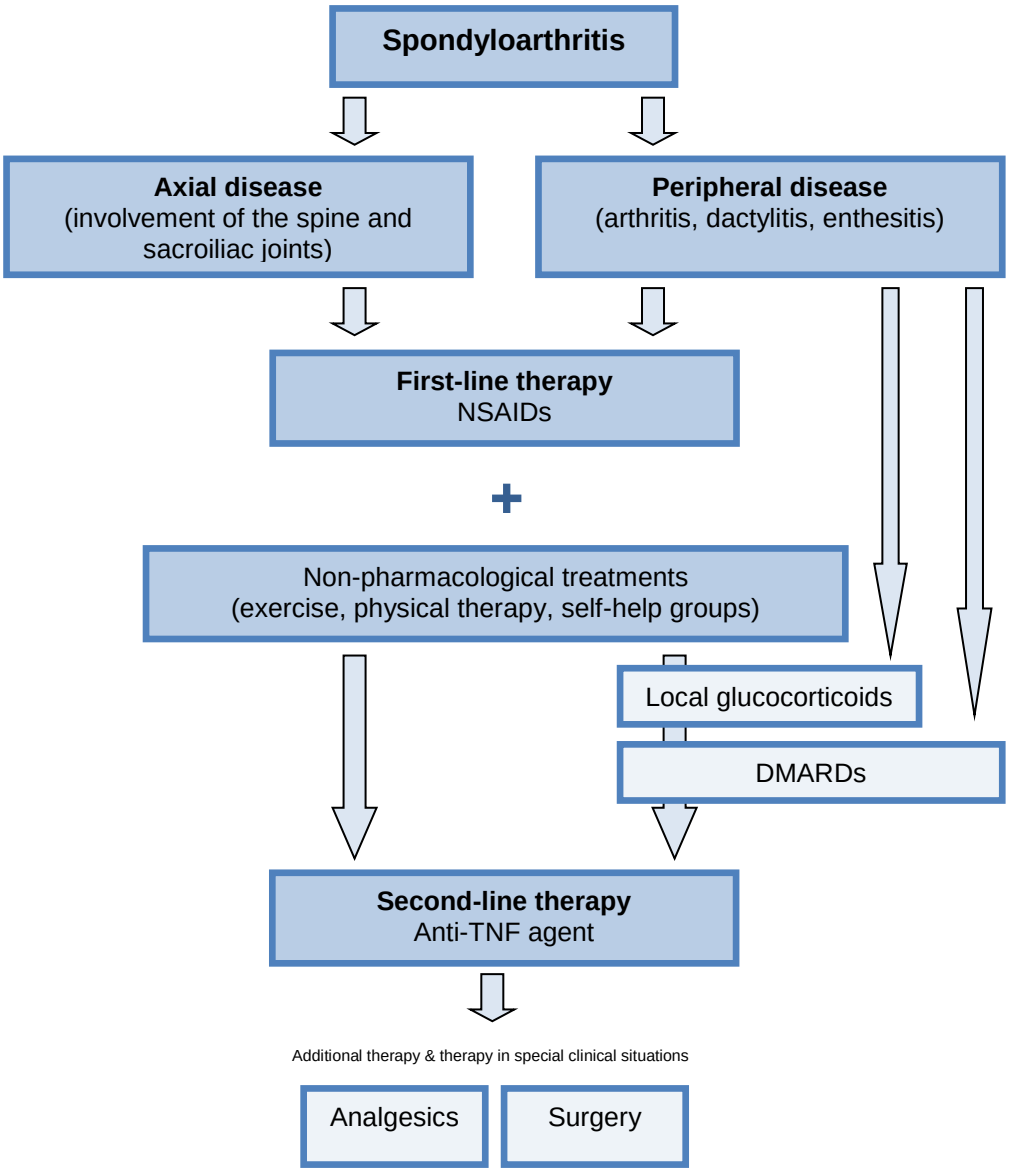
Patients with advanced disease may require surgery. The current ASAS/EULAR guidelines recommend that total hip arthroplasty should be considered in those patients with refractory pain or disability and radiographic evidence of severe structural damage (30). In those patients with advanced AS and severe disabling deformity (e.g. hyperkyphosis), spinal corrective osteotomy may be considered. Consultation with a spinal surgeon is recommended for AS patients with an acute vertebral fracture as this can lead to spinal instability (30).

1.4.3 Treatment pathway

A summary of the ASAS/EULAR recommendations for the management of AS is shown in Figure 1. In the UK, the current treatment pathway for patients following a diagnosis of AS or

nr-axSpA is to firstly consider treatment options (e.g. rheumatology and physiotherapy) and to follow NICE guidance for AS. Patients may then receive conventional treatment which consists of two or more NSAIDs taken sequentially at the maximum tolerated dose. Patients are then assessed after 4 weeks to see if symptoms are controlled. Patients may also receive physiotherapy. Anti-TNF therapy may be considered for patients with severe active AS if conventional treatment fails to control symptoms (see section 1.4.3.1).

Figure 1: Recommendations for the management of AS according to ASAS/EULAR guidelines



Abbreviations: DMARD, disease modifying anti-rheumatic drug; NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor.
Adapted from: Poddubnyy 2013 (43)

1.4.3.1 Treatment guidelines

In the UK, NICE has produced two technology appraisal guidelines on the use of anti-TNF agents for the treatment of AS.

NICE TA 143 (4):

- NICE TA 143 recommends adalimumab or etanercept as treatment options in adults with severe active AS providing that all of the following criteria are met:
  - The patients disease satisfies the mNY criteria for diagnosis of AS
  - There is evidence of sustained active spinal disease as demonstrated by:
    - A BASDAI score ≥4
    - A spinal pain score ≥4 on a 0–10 cm visual analogue scale (VAS)
    - Both of these factors should be demonstrated on two occasions at least 12 weeks apart without any change of treatment
  - Conventional therapy with ≥2 NSAIDs taken sequentially at the maximum tolerated or recommended dose has failed to control symptoms
- Treatment with both adalimumab and etanercept should be carried out by clinicians specialising and experienced in the diagnosis and treatment of AS
- Infliximab is not recommended for the treatment of AS

NICE TA 233 (5):

- NICE TA 233 recommends golimumab as a treatment option for severe, active AS in adults only if the following criteria are met:
  - Golimumab is used as described for adalimumab and etanercept in NICE TA 143, **and**
  - The manufacturer provides the 100 mg dose of golimumab at the same cost as the 50 mg dose in accordance with the patient access scheme
- Those patients receiving golimumab for the treatment of severe active AS who do not meet the criteria described in NICE TA 143 should be given the option to continue therapy until they or their clinician consider it appropriate to stop

1.5 Factors for consideration in anti-TNF therapy

1.5.1 Response to therapy

ASAS clinical practice guidelines on the use of anti-TNF agents for the treatment of AS recommend that response to treatment should be assessed by measuring disease activity levels using the BASDAI and ASAS core set (60). Assessments should be performed after at least 12 weeks following initiation of therapy (60).

In the UK, NICE TA 143 guidance states that patients receiving adalimumab or etanercept for the treatment of severe active AS should be assessed for their response to treatment 12 weeks after starting therapy. The guidelines state that treatment should only be continued in the presence of an adequate response defined as:

- a reduction in the BASDAI of at least 50% of the pre-treatment value or by ≥2 units, and
- a reduction in the spinal pain VAS score ≥2cm.

It is recommended that patients responding to therapy should be monitored at 12 weekly intervals. If the response to treatment is not maintained, the patient should be reassessed after a further 6 weeks, after which treatment should be discontinued if the response has still not been maintained (4).

Previous studies have reported that there are several possible predictors of a positive response to treatment with anti-TNF agents in patients with AxSpA. These include younger age, functional status at initiation of therapy and elevated levels of C reactive protein (CRP) and evidence of MRI sacroiliitis at baseline (31, 44, 75). However, there are limited data available in determining whether these are treatment effect modifiers (i.e. factors that differentially effect treatments compared in a clinical trial), or whether they are simply prognostic factors that identify characteristics of patients who have a better response to any treatment.

A number of patients may discontinue treatment with anti-TNF agents due to inadequate response or adverse events. Also, because there are differences between the mechanisms of action and chemical structures of these therapies, there is a rationale for switching to a different anti-TNF in these circumstances (76). Observational studies investigating the effects of AxSpA and AS patients switching from one anti-TNF drug to another have reported a significant response to a second anti-TNF therapy (76, 77). Pradeep et al, 2008 reported that a better response was observed in patients who switched due to an adverse event compared with those who switched therapies due to inadequate response (76). The current ASAS/EULAR guidelines include a recommendation for switching anti-TNF agents in patients with AS (30). The guidelines state that “switching to a second TNF blocker might be beneficial especially in patients with loss of response” (30).

1.5.2 Adverse events with anti-TNF therapy

Data from randomised controlled trials and from longer term extension studies have demonstrated that anti-TNF therapies have an acceptable and tolerable safety profile (78). However, the use of anti-TNF agents is associated with an increased risk of tuberculosis (TB) or reactivation of latent TB. This risk is also increased by the use of corticosteroids (79). Therefore, prior to initiating therapy with an anti-TNF agent, patients should be screened for the possibility of latent TB infection (78, 79).

Another recognised risk associated with anti-TNF therapy is the risk of serious infection. Thus anti-TNF therapy should not be given in the presence of active serious and/or opportunistic infections (e.g. sepsis (8, 17, 73, 74, 80), listeriosis (8, 73, 80), osteomyelitis (8) and systemic fungal infections (8, 17, 73, 74, 80)). Older adult patients have a 1.5–2-fold higher rate of serious infection compared to younger adult patients (79).

The risk of malignancy is another safety concern when receiving treatment with anti-TNF agents. However, the most recent consensus statement on biological agents for the treatment of rheumatic diseases (2012) reports that the current volume of data available supports the notion that anti-TNF therapy is not associated with cancer overall (79). In addition, a recent study from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis reported that there was no difference in the risk of solid cancers in patients treated with anti-TNF agents compared with patients receiving DMARDs (hazard ratio 0.83, 95% confidence interval 0.64–1.07) (81). The study also reported that there was no difference in the risk of mortality following cancer diagnosis between patients receiving anti-TNF therapy and those receiving DMARDS (hazard ratio 0.90, 95% confidence interval 0.70–1.17) (81). However, vigilance for the occurrence of malignancies is advised in patients

receiving anti-TNF treatment (79). Due to the suggestion of an increased frequency of non-melanoma skin cancers in patients receiving anti-TNF agents, recommendations are that patients should be screened for cutaneous malignancy prior to initiating therapy and subsequently monitored during treatment (79).

A common side effect of anti-TNF therapy in placebo-controlled trials is injection site reaction. However, most of these are mild to moderate in intensity and can be treated with anti-histamines if necessary (79)

1.5.3 Antibodies against anti-TNF agents

Treatment with any biologic such as anti-TNF agents can induce an immune response in the patient resulting in the production of antibodies against the drug which can be neutralising (NAbs) and non-neutralising (82). Both may impact the bioavailability and safety of the drug. NAbs bind to the binding sites of the therapeutic protein, thus neutralising it whereas non-neutralising antibodies bind to the site but do not neutralise it. The formation of NAbs has been increasingly recognised as a mechanism to account for failure to response to therapy or a reduction in response over time in patients receiving anti-TNF agents for the treatment of chronic inflammatory diseases (82-84). The reduction in clinical response or lack of response to therapy is due to the formation of an immune complex between anti-TNF agents and neutralising antibodies (82). This results in suppression of the pharmacokinetic mechanism of action of the anti-TNF agent, thus restricting or inhibiting product efficacy (82, 85). In addition, higher antibody titres are associated with development of infusion-related reactions in patients (84).

The immunogenicity of a biologic can be influenced by several factors including the structural characteristics of the protein itself, the duration of treatment, and the route of administration (85). The immunogenicity of a patient may also be influenced by their genetic characteristics (85). However, while the presence of NAbs is associated with a loss of clinical efficacy, the NAbs status of the patient does not necessarily prevent a clinical response (82). In contrast to several inflammatory auto-immune diseases, AS is not associated with any disease specific neutralising antibodies (7). It has also been suggested that as patients with auto-immune diseases have an already heightened immune system, these patients may therefore be more at risk of developing NAbs to anti-TNF agents (7).

Earlier studies reported the detection of NAbs in 31% and 29% of AS patients treated with adalimumab and infliximab, respectively (86, 87) while NAbs were not detected in AS patients receiving etanercept therapy (88). NAbs formation in patients receiving adalimumab and infliximab was reported to be associated with reduced serum levels of both anti-TNF agents in patients and a reduced response to treatment (86, 87).

A recent study to investigate the effect of NAbs production in response to three separate anti-TNF agents on the clinical response in patients with AS reported that NAbs formation was induced in 20% and 30% of patients receiving infliximab and adalimumab therapy, respectively (7). This was associated with a reduction in clinical efficacy and early discontinuation of treatment in these patients. In contrast, NAbs were not detected in any AS patients receiving treatment with etanercept in this study (7). In addition, in those patients that withdrew from treatment due to a lack of efficacy, NAbs were detected in 4 out of 7 patients treated with adalimumab and in 3 out of 4 patients treated with infliximab (7).

Antibodies to etanercept have been detected in the sera of some patients receiving etanercept. However, they have all been non-neutralising and generally transient (8).

Cumulative rates of antibodies to etanercept were approximately 2% of AS patients (8). There appears to be no relationship between etanercept antibody development and clinical response or adverse events (8). These findings demonstrate firstly that etanercept is considerably less immunogenic compared with other anti-TNF agents. Also, the formation of NAbs may be related to treatment discontinuation due to a lack of efficacy in AS patients receiving anti-TNF therapies. In addition, the formation of Nabs can lead to treatment strategies involving dose increases which subsequently increase costs. Note that the increased likelihood of use of dose escalation as a consequence of Nabs and the impact on costs for other anti-TNF agents is not addressed in this submission.

In summary, many factors impact the efficacy and safety of biologics or how an individual patient responds to these treatments. The immunogenicity of a therapeutic agent is one of these factors and it should not be considered in isolation when evaluating the overall efficacy and safety of a drug. The occurrence of anti-drug antibodies is also influenced by several factors including the co-treatment with immunosuppressive drugs such as methotrexate.

As immunogenicity analyses are product-specific, comparison of antibody rates with those from other products is not appropriate.

1.6 Etanercept (Enbrel®)

TNF is a pro-inflammatory cytokine which plays a dominant role in the inflammatory processes of the spondyloarthropathies. Elevated levels of TNF are found in the serum and synovial tissue of patients with AS (8). Etanercept is a recombinant, fully human TNF receptor p75 Fc fusion protein (34). It is produced in a Chinese hamster ovary mammalian expression line using recombinant DNA technology. Etanercept competitively inhibits the binding of TNF to its cell surface receptors preventing initiation of cell signalling thereby inhibiting its biological activity (8). Etanercept may also modulate biologic responses which are controlled by other downstream molecules which are regulated by TNF (e.g. cytokines, adhesion molecules) (8).

In November 1998, etanercept was approved for reducing signs and symptoms in patients with moderately to severely active RA who had an inadequate response to one or more disease-modifying anti-rheumatic drugs (DMARDS) (33). Etanercept has since been approved for clinical use in over 90 countries and is licensed in the EU for various indications including the treatment of rheumatoid arthritis, AS, juvenile idiopathic arthritis, psoriatic arthritis, adult plaque psoriasis, and paediatric psoriasis (8). Etanercept received its licence extension from the European Medicines Agency for nr-axSpA on the 28th July, 2014.

1.6.1 Unmet need

There is currently no cure for AxSpA. The aims of treatment are to control symptoms and reduce joint damage and disability thus improving patient QoL. Pharmacological management of AxSpA relies primarily on NSAIDs. The development of etanercept and its subsequent use in immunological disease, including AS, represents an innovative approach to treatment and a technological advance over traditional immunosuppressive therapies. Etanercept also provides an effective treatment option for patients whose NSAID treatment has failed on multiple occasions. However, clinicians currently regard the nr-axSpA population as an earlier stage of the AxSpA disease spectrum that includes AS. As discussed in section 1.3, patients with nr-axSpA have a similar burden of disease as patients with AS. In addition, studies on the use of anti-TNF agents demonstrate similar efficacy in both the nr-axSpA and AS populations. The absence of any NICE Guidance for anti-TNF

agents as a treatment option for nr-axSpA patients represents a clear unmet need for this newly defined group of patients who, prior to the development of the new ASAS classification criteria, often had to wait for close to a decade to receive a diagnosis and effective treatment.

The current EMA approved indication for etanercept for the treatment of AS and nr-axSpA is as follows:

- Treatment of adults with severe active ankylosing spondylitis who have had an inadequate response to conventional therapy (8).
- Treatment of adults with severe non-radiographic axial spondyloarthritis with objective signs of inflammation as indicated by elevated C-reactive protein (CRP) and/or magnetic resonance imaging (MRI) evidence, who have had an inadequate response to non-steroidal anti-inflammatory drugs.

The wording for the NICE final scope is as follows:

For the treatment of severe AS:

- People with severe active ankylosing spondylitis whose disease has responded inadequately to, or who are intolerant to, non-steroidal anti-inflammatory drugs (NSAIDs).

For the treatment of severe nr-axSpA:

- People with severe axial spondyloarthritis without radiographic evidence of ankylosing spondylitis but with objective signs of inflammation, whose disease has responded inadequately to, or who are intolerant to NSAIDs.

While the following sections in this submission provide clinical and economic evidence for the use of etanercept in both populations of the final scope, it should also be acknowledged that clinical trials in the AxSpA population (i.e. a combination of both AxSpA and AS patients) also provide relevant information for the decision problem.

2. Systematic review of the clinical literature

2.1 Identification of studies

A systematic review identified randomised controlled trials (RCTs), in the AS and nr-axSpA populations reporting on the efficacy and safety of etanercept, infliximab, adalimumab, golimumab and certolizumab pegol. In addition, the search strategy allowed for the identification of etanercept RCT open-label extensions (OLEs) and longitudinal observational studies. The objectives of the review were:

- To assess the double-blind and open-label RCTs of etanercept and comparator therapies identified in the search for inclusion in a network meta-analysis (NMA) incorporating an indirect comparison.
- To identify all observational study data for etanercept in the AS and nr-axSpA population areas using a targeted literature search.
- To determine which published studies compared etanercept or another anti-TNF comparator treatment of interest to placebo (includes background NSAIDs) or other treatment falling outside the NICE final scope (which could be used in an indirect comparison of etanercept versus an anti-TNF comparator of interest).
- To determine which published direct and indirect evidence is available for etanercept versus comparators of interest in the approved indications with regard to the following criteria:
  - Trials should report efficacy at 12 weeks for any of the following outcomes:
    - difference between intervention and placebo in change in C-Reactive Protein (CRP) from baseline at 12 weeks (for AS population)
    - difference between intervention and placebo in change in ASDAS-CRP from baseline at 12 weeks (for nr-axSpA population)
    - difference between intervention and placebo in change in BASDAI from baseline at 12 weeks (for AS and nr-axSpA)
    - difference between intervention and placebo in change in BASFI from baseline at 12 weeks.

An initial literature search was performed on 24th February, 2014 with no restrictions on date. An updated search was conducted on 28th May 2014. The following electronic databases were searched; MEDLINE/MEDLINE In-Process, EMBASE, and the Cochrane Library (including the databases Cochrane Reviews, Central Register of Controlled Trials and Cochrane database for health technology assessment [HTA]). The search strategy is provided in Appendix A, section 8.1. The following additional sources were searched:

- NICE TA143 and TA233 (4, 5), including associated manufacturer submissions and Evidence Review Group (ERG)/assessment group reports
- Reference lists of published clinical systematic reviews and meta-analyses, including RCTs, OLEs, and observational studies
- The US Food and Drug Administration (FDA) website for reports relating to the adalimumab ABILITY-1 and the certolizumab pegol RAPID-axSpA trials.

The following conference proceedings were also searched:

- European League against Rheumatism (EULAR), 2011–2014
- The British Society for Rheumatology (BSR), 2011–2014
- American College of Rheumatology (ACR), 2011–2013.

Since the ACR conference is only held in November 2014, it was not possible to update the search for abstracts for the 2014 conference.

2.2 Study selection

2.2.1 Eligibility criteria

Inclusion and exclusion criteria, based on the population, interventions and outcomes reported, are shown in Table 1. The choice of outcomes in the inclusion criteria was based on a recent evidence-based international task force set of recommendations for treating spondyloarthritis (including AS) (29). These recommend the use of validated composite measures of disease activity, including BASDAI, or ASDAS, measures of function such as the BASFI, and acute-phase reactants (such as CRP), to guide treatment decisions in routine clinical practice (29).

The 12 week time point in the outcomes inclusion criteria was selected as several of the key etanercept RCT trials only reported blinded comparative data for up to 12 weeks (11-13, 68, 89). In addition, three of the key adalimumab, golimumab and certolizumab pegol trials (69, 90, 91) had an early escape design after 12 or 16 weeks. This meant that placebo arm non-responders (90, 91) or patients that failed to achieve a satisfactory response (69) were switched from placebo to active treatment after 12 or 16 weeks. This could have potentially compromised the study blinding and randomisation and makes meta-analysis after this time-point problematic.

Table 1: Eligibility criteria used in search strategy

	Description
Inclusion criteria	
Population	Patients with a diagnosis of severe active AS or severe active nr-axSpA
Interventions and comparators	As listed in the final NICE scope: • Intervention – etanercept • Comparators – adalimumab, certolizumab pegol, golimumab, infliximab • Established clinical management without TNF inhibitors (i.e. concomitant/background treatment predominantly consisting of NSAIDS as per the NICE final scope (92))
Outcomes	• BASDAI at 12 weeks (both indications) • BASFI at 12 weeks (both indications) • CRP at 12 weeks (AS) • ASDAS-CRP at 12 weeks (nr-axSpA)
Study design	• Double-blind or open label RCTs (Phase II and above) of etanercept, infliximab, adalimumab, golimumab, and certolizumab pegol were eligible for inclusion in the NMA • Etanercept OLEs and observational studies (longitudinal cohort studies, case-control, cross-sectional and epidemiological studies)
Language restrictions	• English abstracts of non-English publications, if available, were reviewed to assess eligibility

	Description
Exclusion criteria	
Population	• Asian populations =100%
Interventions	• RCTs of conventional therapy (NSAIDs) that did not include a direct comparison to an anti-TNF agent of interest • RCTs of analgesics and DMARDS in line with the final NICE scope
Study design	• Abstracts or conference presentations for which there was also a full paper • Abstracts reporting insufficient information relating to relevant treatment effects • Case studies, letters, reviews or editorials • OLEs and observational studies relating to adalimumab, infliximab, golimumab, and certolizumab pegol as not eligible for inclusion in the NMA • RCTs with a cross-over design • Fewer than N=30 participants • Observational studies of less than 1 year duration
Language restrictions	• Non-English publications

Abbreviations: AS, ankylosing spondylitis; ASDAS, Ankylosing Spondylitis Disease Activity score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; nr-axSpA, non-radiographic axial spondyloarthritis; NMA, network meta-analysis; NSAID, non-steroidal anti-inflammatory drug; OLE, open-label extension; RCT, randomised controlled trial; TNF, tumour necrosis factor.

Titles and abstracts were screened and any studies not meeting the inclusion criteria were excluded. Studies that could not be excluded with certainty on the basis of the title or abstract were retrieved in full for review. Identified studies were independently assessed by a reviewer in order to ascertain whether they met the pre-defined inclusion and exclusion criteria, and any uncertainties were resolved by discussion with a second reviewer. Data were extracted from eligible publications into a pre-defined table by a reviewer. Full details of the databases, search dates, additional searches, and data abstraction strategy are provided in Appendix A, sections 8.1.1 to 8.1.3

Note on established clinical management: While the literature search included search terms for clinical management (NSAIDs, analgesics and DMARDs) based on the draft NICE scope, the latter on their own were excluded from consideration on review of potentially relevant full-texts. This was because the NICE final scope defined conventional therapy for axial spondyloarthritis as including acute anti-inflammatory treatment with NSAIDs, but did not mention analgesics or DMARDs (92). Therefore, studies of analgesics and DMARDs were excluded if they did not report a comparison with an anti-TNF treatment. With regard to NSAIDs, the placebo-controlled anti-TNF RCTs include a proportion of patients who are taking conventional therapies as concomitant/background therapy. In the etanercept AS trials, use of NSAIDs in the placebo arm ranged from 45% in one small trial of N=40 participants (93) to 78-85% in two larger trials (N=356; N=84) where this data was reported (11, 12). In nr-axSpA the single etanercept trial reported 100% concomitant NSAIDs use in both placebo and active treatment arms (68). Therefore, the placebo arms in these trials provide an estimate of conventional therapy (NSAIDs) use without anti-TNF treatment compared to the anti-TNF therapy arms. Consequently, there was no need to consider

studies of NSAIDs that did not directly compare to anti-TNF treatments, and so the former were excluded from consideration on full-text review.

2.2.2 Included studies

Original search

In total, 11350 citations were identified through the original electronic database searches (conducted February 2014). Upon removal of duplicates, 8619 titles and abstracts were screened. Of the identified studies, 412 citations were reviewed. After applying additional inclusion/exclusion criteria as described in Table 1, 33 RCT full-text publications reporting on 24 RCTs of etanercept, adalimumab, infliximab, golimumab and certolizumab pegol in either the AS, nr-axSpA, or AxSpA (where results were stratified by AS and/or nr-axSpA sub-groups) populations were included (11, 12, 14, 41, 69-71, 89-91, 93-114). One additional report of an RCT compared the baseline characteristics and response rates after one year of treatment with etanercept in patients in AS and nr-axSpA sub-groups (115).

Ten open label extension studies of etanercept RCTs (22, 116-124) and three long-term observational studies of etanercept safety (7, 18, 64, 125) were also identified. In addition, seven conference abstracts reporting on four RCTs (126-131) and one OLE in an AS population were identified (132).

A summary of the systematic review procedure for the original search conducted in February 2014 is shown in Figure 2.

Updated search

A total of 597 citations were identified in the updated electronic database search in May 2014 of which 501 were screened on title and abstract (after removal of duplicates). Of the identified studies, 14 citations were reviewed, of which 13 were excluded (Appendix A, section 8.1.4), resulting in one full publication for final inclusion (133). An updated search of the EULAR and BSR 2014 conferences (conducted May 2014) yielded an additional five abstracts reporting on four RCTS for inclusion in the data extraction stage of the systematic review (29, 134-137) across the AS and nr-axSpA indications. Two of these abstracts (136, 138) reported 48-week open-label extension results of an etanercept versus placebo RCT in the nr-axSpA population, in addition to double-blind 12-week data already reported elsewhere.

Targeted searching of the FDA website identified two additional assessment reports of two RCTs focusing respectively on the use of adalimumab (ABILITY-1) and certolizumab pegol (RAPID-axSpA) in the nr-axSpA population(139, 140). Although both trials are reported in publications identified in the literature search, the FDA assessment reports describe additional sub-group analyses relevant to the systematic review and were included for this reason.

Figure 3 shows a summary of the systematic review procedure for the updated searches performed.

Figure 2: SR flow diagram (original search carried out February 2014)

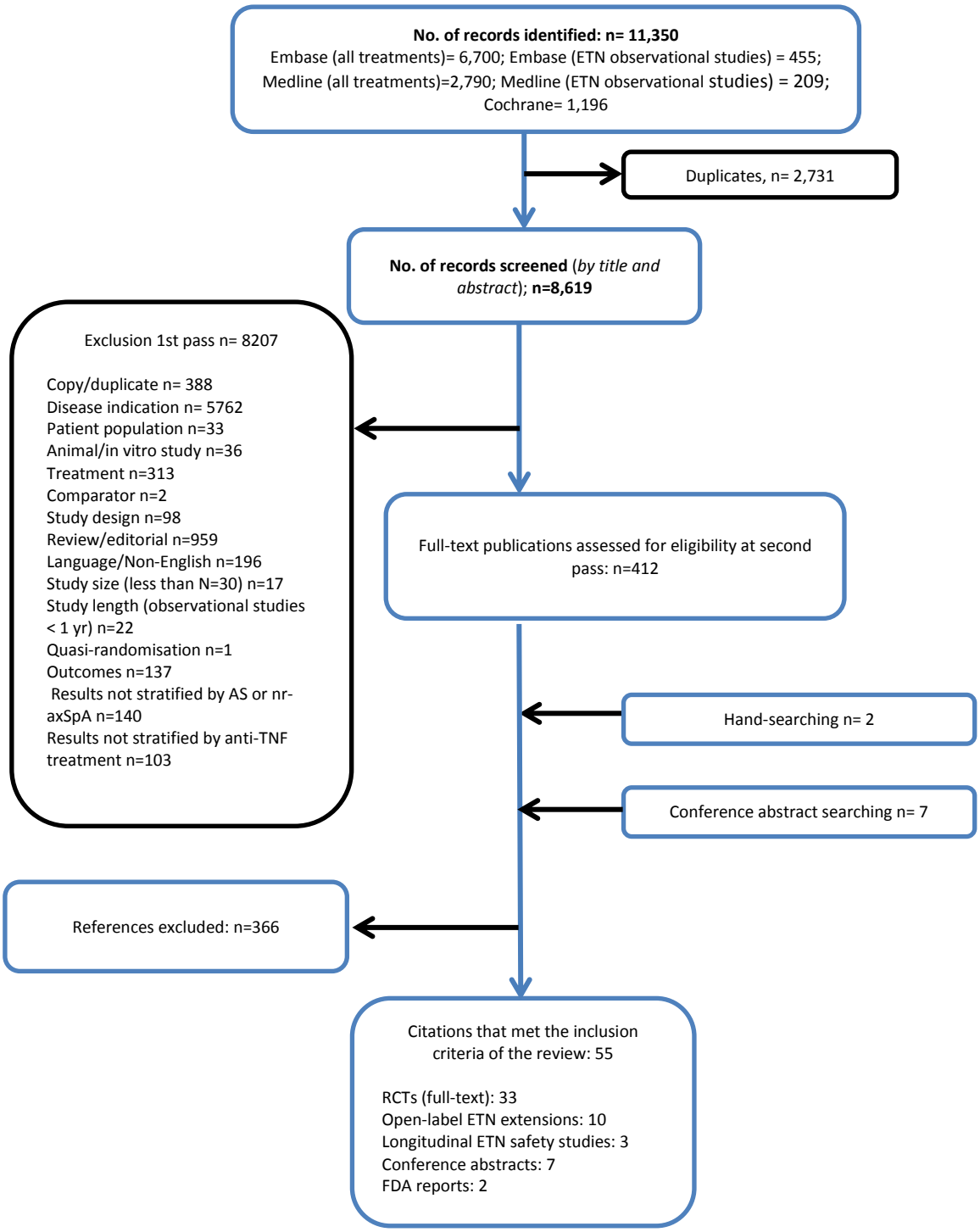
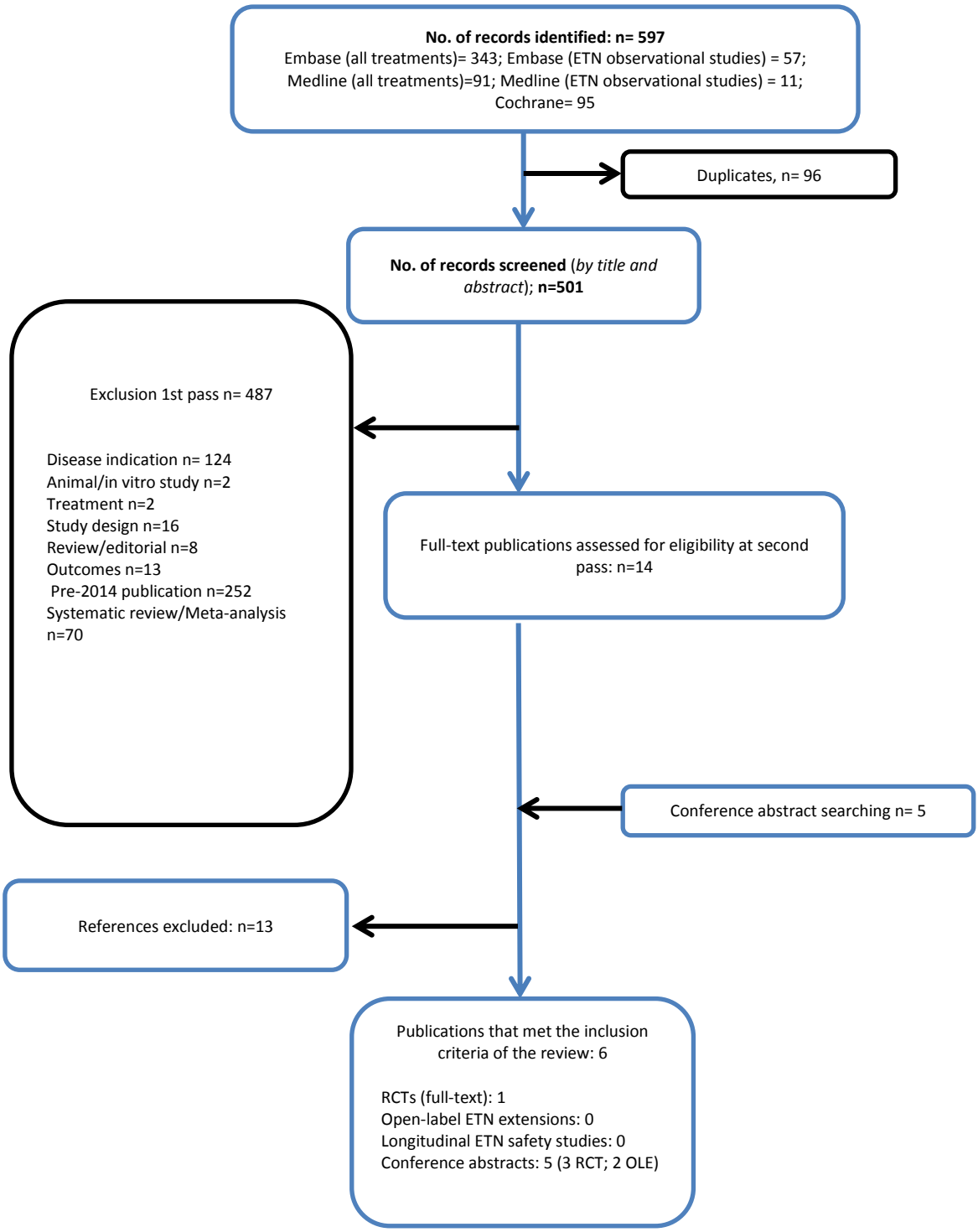


Figure 3: SR flow diagram (updated search carried out May 2014)



3. Efficacy and safety of therapy

3.1 Introduction

In the previous NICE submission to appraise the clinical and cost-effectiveness of etanercept in 2006, data was presented from three Phase III, double-blind, placebo-controlled RCTs (10, 11, 72) and one open-label extension study (119) designed to evaluate the efficacy and safety of etanercept in patients with AS. Supporting efficacy data was also presented from two placebo-controlled Phase II studies (89, 93).

A comprehensive and up-to-date efficacy and safety profile of Etanercept (Enbrel®) for the treatment of both non-radiographic axial spondyloarthritis (nr-axSpA) and ankylosing spondylitis (AS) is now presented in sections 3.3 to 3.6. The sources of evidence presented were identified by means of a systematic literature review, as outlined in section 2.

The current efficacy evidence is presented based on RCTs, open-label extension studies and observational studies. In addition to the data previously presented for etanercept in an AS population, efficacy and safety data from one further Phase III RCT is presented together with supporting data from several open-label extension studies. The current efficacy data for etanercept in an AS population also includes two studies comparing the efficacy of etanercept with sulfasalazine and one study where etanercept is compared with infliximab.

Due to a change in the ASAS definition of AxSpA in 2009 (27), which now includes patients with nr-axSpA, and because of the increased use of MRI to detect inflammation in the sacroiliac joints thus enabling the identification of patients with nr-axSpA, anti-TNF agents may be recommended for the treatment of AxSpA earlier on in the disease. The ASAS treatment guidelines have now been updated to reflect the use of anti-TNFs in the nr-axSpA population (60). In light of these changes, the manufacturer applied for a licence to extend the use of etanercept to the nr-axSpA population. The EMA granted the licence on 28th July 2014.

Previous approval for the use of etanercept in adults with severe AS was granted based on clinical evidence presented from three RCTs in 401 patients (10, 11, 93). The recent application for the extension to the approved indications for etanercept was supported by 12 week double-blind results and the open-label results through to Week 24 from the pivotal 1031 study conducted in an nr-axSpA population. In the current submission, the clinical evidence supporting the use of etanercept for the treatment of patients with nr-axSpA is presented from the pivotal 1031 trial.

A summary diagram of etanercept trials in nr-axSpA and AS is shown in Figure 4. In addition, a 3–4 page summary of each study is provided and is presented in the following order:

Etanercept for the treatment of nr-axSpA

- RCTs and open-label extension studies (n=1, section 3.3)

Etanercept for the treatment of early AxSpA (mixed population studies)

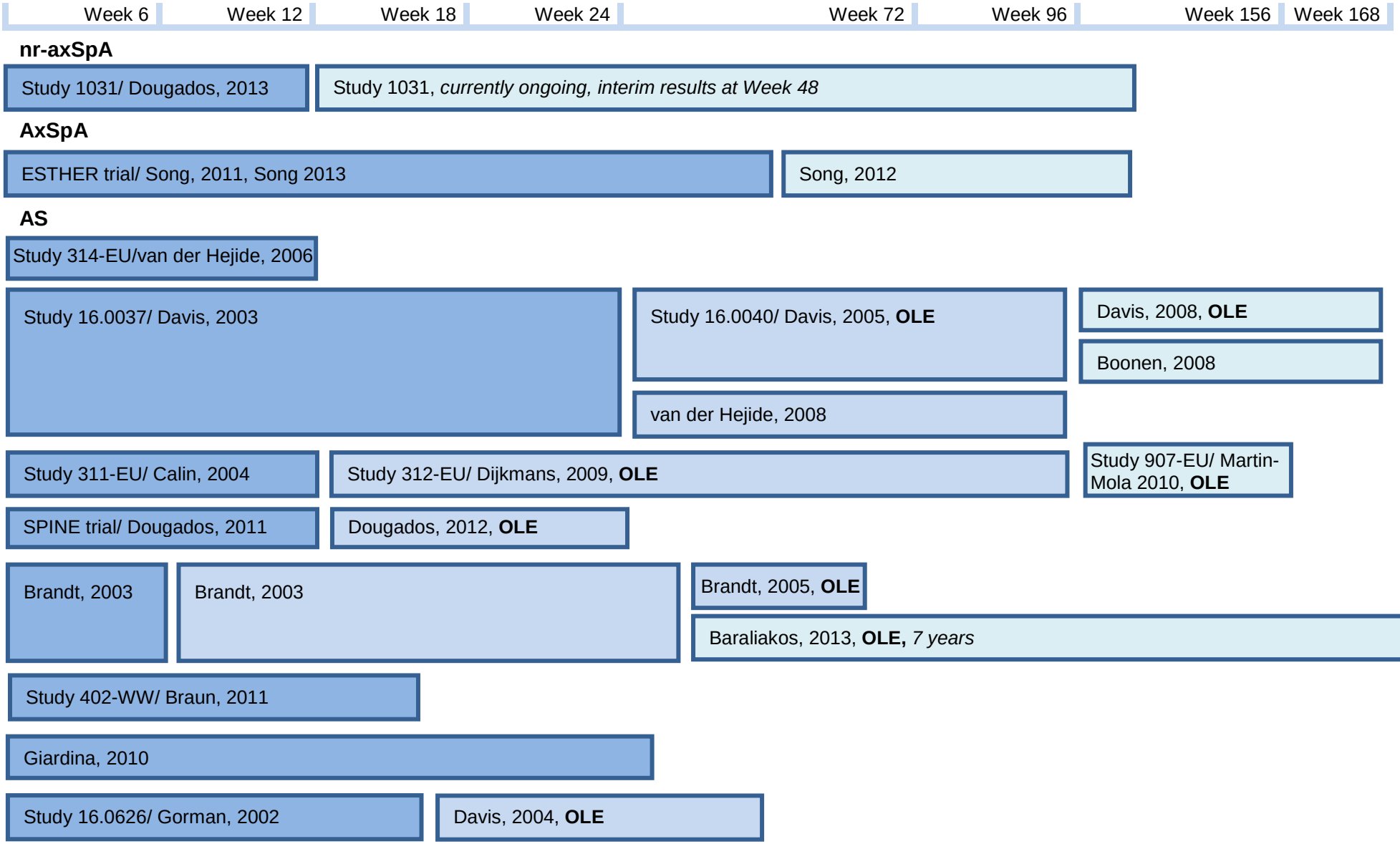
- Investigator-initiated research studies (n = 1, section 3.4.1)
- Follow-up observational studies (n=2, section 3.4.3 and section 3.4.2)

Etanercept for the treatment of AS

- Phase III RCTs, open-label extension studies and observational studies (n=14, section 3.5)
- Etanercept versus sulfasalazine (n = 1, section 3.5.15)
- Etanercept versus infliximab (n=1, section 3.5.16)
- Phase II RCTs and open-label extension studies (n = 2, section 3.6)

The Safety evidence is presented in section 3.7 as an overview from the substantial body of clinical trial evidence and from post-marketing clinical experience, with detailed discussion provided on the occurrence of specific adverse events.

Figure 4: Summary of etanercept clinical trials in patients with nr-axSpA and AS with their relevant publications



3.2 Summary of efficacy and safety evidence

There is a substantial body of clinical evidence which demonstrates that etanercept is an efficacious therapy with a rapid and sustained effect across the broad spectrum of the AxSpA population. Etanercept was previously granted a licence for use in a severe AS population based on clinical evidence from three randomised controlled trials which demonstrated that etanercept is effective in controlling the clinical and functional progression of AS. The current submission adds to the evidence base for etanercept for the treatment of AS and now shows that the clinical benefit in this group of patients is now sustained for up to 7 years. We also present the new clinical data from the pivotal 1031 trial which led to our licence extension and demonstrated that etanercept is a clinically beneficial treatment option for patients with nr-axSpA.

Etanercept for the treatment of nr-axSpA

On the 28th July, 2014, the EMA granted a licence to extend the use of etanercept to the nr-axSpA population. This decision was based on key clinical evidence presented from the pivotal 1031 trial which was designed to focus on the nr-axSpA population. To minimise the risk of including AS patients in the study population, patients enrolled in the 1031 study were aged<50 years with a symptom duration <5 years. To further reduce the risk of including AS patients in the study, diagnosis was based on central assessment of radiographs. The primary objective of study 1031 was to compare the efficacy of etanercept against placebo in improving symptoms of early nr-axSpA at 12 weeks when added to a background NSAID at the optimal anti-inflammatory dose. The efficacy evidence presented in support of etanercept within the nr-axSpA population is based on results obtained from the 12 week double-blind period and the open-label results through to Week 48.

- In the pivotal 1031 trial, 12 weeks of treatment with etanercept in patients with an inadequate response to ≥2 NSAIDs was associated with rapid and significant improvements in disease activity, function and a decrease of inflammation on MRI (68). The level of clinical responses observed during the 12 week period continued to increase up to 48 weeks. The level of MRI inflammation continued to decrease during the open-label period (141). Specifically:
  - After 12 weeks, an ASAS 40 response was achieved in 32% of etanercept patients compared with 16% of placebo patients (p=0.0062)
    - At 48 weeks, an ASAS 40 response was achieved in 53% of all patients continuing with etanercept
  - After Week 12, 44% of etanercept patients achieved BASDAI 50 vs 24% of placebo patients (p<0.01)
    - After 48 weeks, 62% of patients had achieved BASDAI 50
  - Patients receiving etanercept had significantly greater improvement on the BASFI compared with placebo patients after 12 weeks (p=0.016)
    - Improvements in the BASFI continued over the 48 week open label extension period (mean % change from baseline: -53%)

- After 12 weeks:
  - Etanercept patients had significantly greater reductions in inflammation compared with placebo patients as evidenced by the reduction in high sensitivity C-reactive protein (hsCRP) levels [redacted]40 % of etanercept patients had reached ASDAS inactive disease vs 17% of placebo patients (p=0.0004).
  - At Week 48, the mean % change from baseline in hsCRP levels was [redacted]%
- At week 12, etanercept patients had significantly greater reductions in SPARCC MRI sacroiliac joint scores compared with patients receiving placebo (47% vs 11%, respectively; p<0.001)
  - Patients continued to show improvement in SPARCC MRI sacroiliac joint scores over the 48 week extension study (mean % change from baseline: -65%)
- Etanercept patients had significantly greater improvements in SPARCC MRI spine scores at Week 12 compared with placebo patients (p=0.04)
  - Patients continued to show improvement in SPARCC MRI spine scores at Week 48 (mean % change from baseline 60.81%)

Etanercept for the treatment of early AxSpA (disease duration <5 years)

- The ESTHER trial demonstrated that etanercept over 48 weeks of therapy in an AxSpA population was associated with significant reductions in active inflammatory lesions as measured by MRI compared with patients receiving sulfasalazine (SSZ) (15)
  - A subsequent follow-up study showed that etanercept demonstrated equivalent efficacy in nr-axSpA and AS patients with similar baseline characteristics (16)
    - There were no differences in ASAS 40, ASDAS and BASDAI response rates between treatment groups (16)

Etanercept for the treatment of severe AS

- Phase III placebo-controlled trials and related open-label extension studies demonstrated that etanercept is an efficacious therapy for the treatment of active and advanced AS, being associated with rapid and sustained improvements in the clinical outcomes of disease. Data from open label extension studies demonstrated that the clinical benefits of etanercept are sustained for up to 7 years of therapy. Specifically:
  - In placebo-controlled trials, 12 weeks of etanercept, once weekly was associated with achievement of ASAS 40 response in 44–53% of patients vs 22–23% of patients receiving placebo (12, 13)
    - ASAS 40 response was sustained for up to 7 years of etanercept treatment (9)
  - Two RCTs reported that ASAS 50 response was achieved in 42–49% of etanercept patients compared with 10% of placebo patients (10, 11) while 24–28% of patients achieved an ASAS 70 response compared with 5–10% of placebo patients over 12–

24 weeks (10, 11)

- ASAS20 was achieved in 59–71% of etanercept patients compared with 23–37% of placebo patients in four RCTs (10-13)
  - Open-label extension studies demonstrated a sustained durability of response for up to 168 weeks in etanercept patients achieving ASAS 20
- Four placebo-controlled RCTs reported significant improvements in the BASDAI in patients receiving etanercept 25 mg twice weekly versus placebo at 6 weeks (p=0.004) (89), 12 weeks (p<0.001) (11, 12) and 24 weeks (p<0.0001) (10)
  - Etanercept was associated with significantly greater improvements in the BASDAI compared with placebo at 12 weeks (SPINE study, p=0.008; 314-EU, p<0.001) (12, 13)
  - Patients continued to show improvements in the BASDAI for up to 7 years of treatment (9)
- Five placebo-controlled RCTs reported that etanercept was associated with significant improvements on the BASFI at 6 weeks (p=0.008; (89)), 12 weeks (p<0.0001; (10, 12) and p=0.000; (11)) and 24 weeks of therapy (p<0.0001; (10))
  - Open-label extension studies reported that benefits were either sustained or continued to improve for up to 7 years (9)
- Significant improvements in CRP levels in patients receiving etanercept were reported in five placebo-controlled trials
  - Four RCTs reported significant improvements in CRP levels after 6 weeks (p=0.0001; (89)), 12 weeks (p=0.000; (11) and p<0.0001; (12, 13)) and 24 weeks (p<0.0001; (10)) of etanercept treatment
  - In the SPINE trial, ASDAS >2.0 (major improvements) were observed in 38% of etanercept patients with advanced AS compared with 1% of placebo patients (p<0.0001) (13)
  - Open-label extension studies reported that improvements in CRP levels in patients receiving etanercept were sustained up to 168 weeks
- Long term data shows that after 7 years of continuous etanercept therapy, one third of patients achieved ASAS partial remission while 44% of patients achieved ASDAS inactive disease status (9)
- Long term etanercept therapy also had beneficial effects on patients HRQoL which were sustained up to 7 years, as evidenced by significant improvements in the SF-36 (9)

Etanercept versus SSZ

- A significantly greater proportion of etanercept patients achieved ASAS 40 and ASAS 5/6 responses (p<0.0001) compared with SSZ patients over 16 weeks of therapy (14)
- At Week 16 of therapy, a significantly greater proportion of etanercept patients achieved a 50% improvement in the BASDAI compared with SSZ patients

([redacted]) (14)

- Patients receiving etanercept demonstrated significantly greater improvements in mobility (measured on the BASMI, p<0.0001) and disease activity (measured on the BASFI, p<0.0001) compared with patients receiving SSZ (14)

Etanercept versus infliximab

- Note that this is the only head-to-head study currently conducted between anti-TNF agents (104)
- Both etanercept and infliximab produced rapid, significant and sustained improvements in ASAS 20 and ASAS 40 after 12 weeks of treatment (104)
  - Clinical improvements were sustained for up to 2 years of continuous therapy with no significant differences observed between treatment groups at this timepoint(104)

3.2.1 Glossary of outcome measures used in clinical trials of AS and nr-axSpA

Recommendations developed by an international task force of physicians and patients state that the key aims of treating patients with all forms of spondyloarthritis are to improve overall patient quality of life (QoL) (29). This is achieved through the management of signs and symptoms, prevention of structural damage, preservation or improvement of function and by reducing the occurrence of comorbidities. In addition, a patient's treatment should be individualised depending on the "current clinical manifestations of their disease" while disease activity should be measured according to "clinical signs and symptoms" and levels of acute phase reactants (29). Disease specific recommendations for the treatment of AxSpA are that treatment decisions should be based on the use of validated composite measures of disease activity which should be conducted and documented regularly. Other factors which may be considered when setting treatment targets include radiographic progression and axial inflammation as evidenced by magnetic resonance imaging (MRI) (29). A summary of the various measures of treatment response, remission criteria and QoL tools commonly used in the clinical trials presented in this submission are shown in Table 2.

Table 2: Summary of main treatment response and remission criteria used in clinical trials of AS and nr-axSpA

Treatment response criteria	Definition
ASAS 20	An improvement of ≥20% with an absolute improvement of ≥10 units (on a scale of 0–100 measured on a VAS) in at least three of the four ASAS domains with no worsening in the fourth domain (defined as worsening of ≥20% absolute worsening of ≥10 units on a scale of 0–100) (142). The four ASAS domains are: • Inflammation (mean of the two morning stiffness-related BASDAI scores) • Pain (mean of patient assessment of total back pain) • Physical function (represented by the BASFI score) • Patient's global assessment of disease activity (assessment of overall disease activity over the last 48 hours using a pain scale between 0 mm [no pain] and 100 mm [severe pain])

ASAS 40	An improvement of ≥40% in at least three of the four ASAS domains with no worsening in the remaining domain Note, while ASAS 20 was used as a primary endpoint in earlier clinical trials of etanercept in patients with AS, the EMA states that the use of ASAS 20 as a primary efficacy endpoint is acceptable for some products such as NSAIDs. However, in the case of products in different therapeutic classes to NSAIDs, a higher improvement may be required to assess major clinical response (143). Consequently, the more recent trials include ASAS 40 (or ASAS 50/ASAS 70) as a primary endpoint to demonstrate a greater magnitude of the clinical response.
ASAS 50	An improvement of ≥50% in at least three of the four ASAS domains with no worsening in the remaining domain
ASAS 70	An improvement of ≥70% in at least three of the four ASAS domains with no worsening in the remaining domain
ASAS 5/6	An improvement of >20% in at least five of the following six domains: • Inflammation • Pain • Physical function • Patient's global assessment of disease activity • Spinal mobility (measured by modified Schober's test) • CRP levels
ASAS partial remission	Scores <20 units on a 0–100 mm scale for each of the four ASAS domains (142)
ASDAS	A composite index used to assess disease activity in AS. Five disease activity variables are combined in order to discriminate between high and low disease activity state (144): • total back pain • patient global assessment of disease activity • duration of morning stiffness • peripheral pain and swelling • C-reactive protein (CRP) OR erythrocyte sedimentation rate (ESR) The ASDAS-CRP is the version of the ASDAS preferred by ASAS while ASDAS-ESR is used as an alternative version.
BASDAI	A 6-item questionnaire based on the five major symptoms of AS: fatigue, spinal pain, joint pain/swelling, areas of localised tenderness and morning stiffness. Patients answer six questions, one for each symptom (with two questions for morning stiffness). Symptoms are measured using a 10 cm VAS (on a scale of 0–10) with higher scores indicating a greater severity of disease. The mean of the two scores for morning stiffness is added to the remaining five scores and the total score is averaged to give a final 0–10 BASDAI score (145).
BASDAI 20	An improvement in the BASDAI of at least 20%.
BASDAI 50	An improvement in the BASDAI of at least 50%.
BASFI	A 10-item questionnaire to assess a patient's degree of functional ability in normal daily tasks. A 10 cm VAS is used and patients indicate their responses on a horizontal line where responses range from "easy to perform tasks" to "impossible to perform tasks". A BASFI score (a value between 0–10) is produced from the mean of the ten scales (146) with higher scores indicating a greater impact on function.
BAS-G	A questionnaire designed to make a global assessment of the effects of AS on a patient's well-being over the last week/6 months (147).
BASMI	A system designed to assess patients axial status (defined as cervical, dorsal and lumber spine, hips and pelvic soft tissue) to determine clinically significant changes in spinal movement. Higher scores indicate a greater degree of

	limitation of movement (148). The BASMI includes the following clinical measurements: • Cervical rotation • Tragus to wall distance • Lumbar side flexion • Modified Schober's test • Intermalleolar distance
VAS	An instrument used to measure a patients perception of their current health state (e.g. disease state or level of pain). The patient marks their response on a horizontal line (on a scale of 0–10) and the VAS score is determined by measuring from the far left of the line to the point of the patient's response. A higher score indicates a greater severity of disease/pain (149).
Dougados Functional Index	A 20-item self-reported questionnaire to assess a patient's degree of functional ability in various activities of normal daily living. Patients give their responses on a three-point scale of 0–2 where 0= yes, with no difficulty; 1= yes, but with difficulty; 2= impossible to perform. Scores are added to give a total score of dysfunction (147).
Modified Schober's test	A measure of lumbar flexion. The patient stands with feet 30 cm apart and the clinician marks points 15cm apart at the junction of the fourth and fifth lumbar vertebrae. The patient bends forward while keeping the knees straight and the distance between the points is measured. Increases beyond 15 cm represent the amount of movement achieved (148).
Occiput to wall measurement	A measurement of the distance between the occiput (posterior portion of the head) and wall. The patient stands with their back and heels against the wall while attempting to get the occiput against the wall keeping the chin in the horizontal position (150).
Chest expansion	A measure of maximum expiration to maximum inspiration. Normal chest expansion is 1 inch.
m-SASSS	A composite index used to measure radiographic changes in the spine of patients with AS and encompassing both the cervical spine and the lumbar spine. Assessments detect the presence of erosion, sclerosis or squaring, syndesmophytes, and total bony bridging at each site (151).
SPARCC	A scoring system which involves examination of the entire spine for inflammation but only the six most severely affected disco-vertebral units are scored (152). There are two SPARCC scoring systems: the SPARCC sacroiliac joint score (score range 0–72) and the SPARCC spinal score (score range 0-108).
ASspiMRI	AN MRI scoring system for the assessment and quantification of acute spinal lesions in AxSpA patients (153).
C reactive protein (CRP)	A biomarker of disease activity. CRP is an acute phase reactant protein in plasma. Levels are rapidly elevated in response to infection, tissue injury, or inflammatory disease (154).
Erythrocyte sedimentation rate (ESR)	A biomarker of disease activity. A blood test is performed to detect and monitor non-specific inflammation in the body (155).
Health related quality of life instruments	
ASQoL	An 18-item questionnaire to assess the impact of AS on patient quality of life. Questions request a yes/no response and relate to the impact of pain on day to day activities, coping ability, mood, motivation, sleep, independence, relationships and social life (147).
EQ-5D	A standardised, self-administered questionnaire used as a measure of patient health outcome across five domains: mobility; pain/discomfort; anxiety/depression; self-care; and usual activities (156).

SF-36	A self-reported measure of patient health status comprising a 36-item questionnaire which measures QoL across eight domains: physical functioning; pain; general health, role limitations due to physical health; vitality; social functioning; emotional wellbeing; and mental health (157).
HADS	A self-administered 14-item questionnaire used in hospital and community settings to measure anxiety and depression (158).
MCII	A clinical tool defined as the smallest outcome measure change which indicates a significant improvement in a patient's symptom. Results are expressed as a percentage of improved patients (159).
PASS	A single question outcome tool used to determine the threshold beyond which a patient considers themselves well and is satisfied with their treatment (160).

Abbreviations: ASQoL, Ankylosing Spondylitis quality of life; ASAS, Assessment of Spondyloarthritis International Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath ankylosing Spondylitis Functional Index; BAS-G, Bath Ankylosing Spondylitis Global Score; BASMI, Bath Ankylosing Spondylitis Metrology Index; MCII, minimum clinically important improvement; mSASSS, modified Stoke Ankylosing Spondylitis Spine Score; PASS, patient acceptable symptom state; SF-36, short form-36. SPARCC, Spondyloarthritis Research Consortium of Canada; VAS, visual analogue scale.

3.3 Phase III RCTs : etanercept for the treatment of nr-axSpA

One Phase III RCT with an on-going open-label extension study is presented in this section.

3.3.1 Study 1031

3.3.1.1 Methodology: Study 1031

The source of the efficacy data for study 1031 is Dougados et al, 2014 (68) and the Clinical study Report (141) (NCT01258738). Study 1031 is a multicentre, randomised, 12-week, double-blind, placebo-controlled study to evaluate the efficacy of etanercept on a background NSAID in patients with early nr-axSpA with a 92-week open-label extension study. The open-label extension study is currently on-going; therefore the results from the 12-week double-blind period and the latest results from Week 48 of the open-label extension study are presented in this MTA. The source of the 48 week data are EULAR 2014 abstracts Makysymowych et al, 2014 (136) and Sieper et al, 2014 (138) and the clinical study report (161).

The 1031 study of etanercept in patients with nr-axSpA was designed specifically to focus on a pure nr-axSpA population. The early nr-axSpA population consisted of patients aged <50 years, with symptom duration <5 years. This population was selected to reduce the risk of including AS patients in the population, which was further minimised by using central assessment of radiographs for diagnosis; an approach which reduces the variability known to exist in the interpretation of radiographs of the sacroiliac joint.

A summary of the methodology in study 1031 is provided in Table 3.

Table 3: Methodology: Study 1031

Study	Study 1031 (NCT01258738)
Objectives	To compare the efficacy of etanercept against placebo in improving the symptoms of early nr-axSpA when added to a background NSAID at a stable, optimal anti-inflammatory dose.
Study location	The study was conducted at 48 centres in 14 countries in Europe (including the UK), Asia and Latin America.
Design details	<u>Double-blind period</u> A screening period of up to 4 weeks was followed by assessments carried out at baseline and Weeks 2, 4, 8 and 12. Patients were stratified based on positive or negative sacroiliitis on MRI. Patients, investigators, radiographic and MRI assessors and sponsors were blinded to treatment during the 12-week double-blind period. Patients who completed the double-blind period were eligible to enter the 92 week open-label extension study. <u>Open-label period</u> Assessments were carried out at Week 16 and then every 8 weeks up to Week 48. MRI sacroiliac /spine assessments were performed at baseline, Weeks 12 and 48 and were scheduled for Week 104. X-rays were performed at baseline and were scheduled for Year 2.
Interventions	<u>Double-blind period</u> Patients were randomised in a 1:1 ratio to receive either: • Etanercept 50 mg, subcutaneous injection once weekly plus a stable background NSAID at optimal inflammatory dose, or • Placebo subcutaneous injection plus background NSAID for 12 weeks <u>Open-label period</u> All patients continuing into the extension study were treated with:

Study	Study 1031 (NCT01258738)
	• Etanercept 50 mg, subcutaneous injection once weekly plus a stable background NSAID at optimal inflammatory dose
Key inclusion criteria	• Diagnosis of nr-axSpA according to ASAS criteria with duration of symptoms ≥3 months and <5 years at the time of consent • Active symptoms defined by a BASDAI ≥4 at screening • Axial symptoms of back pain with failure of at least 2 NSAIDs (including current NSAID) taken separately at the optimal dose with a total combined duration of ≥4 weeks. • Taking a stable dose of NSAID ≥14 days prior to baseline • Aged 18–50 years
Key exclusion criteria	• Previous anti-TNF, anti-B/T cell therapy or other biologic/immunosuppressive agents for a condition other than IBD • Current treatment or within the last 6 months with any anti-TNF or other immunosuppressant (with the exception of IBD) • Any other medical condition that could cause chronic back pain (other than SpA) • IBD flare within 6 months of baseline • Uveitis within 6 months of baseline • Radiological sacroiliitis grade 3–4 (unilaterally) or grade ≥2 (bilaterally) as determined by a blinded central assessor • Concurrent use of >1 NSAID within 14 days at baseline or altered NSAID dose within 14 days before baseline • Use of DMARDs (not including methotrexate, SSZ and hydroxychloroquine) within 4 weeks of baseline • Use of prednisone (>10 mg/day) or altered dose within 4 weeks before baseline • Corticosteroid treatment (injection or IV) within 4 weeks before baseline • [redacted]
Permitted concomitant medications	• [redacted] Chronic use of no more than one NSAID at a time and at a stable dose • [redacted]
Efficacy outcomes	<u>Primary efficacy endpoint</u> • Proportion of patients achieving ASAS 40 at Week 12 <u>Secondary efficacy endpoints</u> • Patients achieving ASAS 40 at time points other than 12 weeks • Patients achieving ASAS 20 and ASAS 5/6 • Patients achieving ASAS partial remission and time to ASAS partial remission • Changes from baseline in the following: ASDAS; patients assessment of disease activity; physician global assessment; nocturnal and total back pain over time; the BASFI and its individual components; the BASDAI and its individual components; spinal mobility (BASMI, occiput to wall distance and chest expansion); tender and swollen joint counts (44 counts); dactylitis and enthesitis score (MASES); and hsCRP and [redacted] • Proportion of patients achieving BASDAI 50 • Changes in the BAS-G score • Changes in inflammation at Week 12 as measured by MRI of the spine

Study	Study 1031 (NCT01258738)
	(SPARCC MRI SIJ score, SPARCC MRI spinal score and ASspiMRI) • Health outcomes assessments measured using the following: [redacted]; EQ-5D; [redacted]; SF-36; and WPAI
Populations analysed	Double-blind period • Full analysis population: all randomised patients • mITT population (primary analyses): all randomised patients who received at least one dose of the study drug and had at least one on-therapy assessment • Per protocol analysis population: a subset of the mITT population with the exclusion of patients deviating from the protocol. • Safety analysis: based on mITT population at Week 12 and Week 48 Open-label period • mITT population: all patients with 12 week efficacy data who had received at least one dose of study drug in the open-label period
Statistical analysis	The primary endpoint of ASAS 40 response was analysed using the 2005 EMA guidelines on clinical investigation of medicinal products for the treatment of AS. The 2009 EMA guidelines were used for all other ASAS analyses. The more recent EMA guidelines differ from the 2005 guidelines only in terms of the definition of pain assessment, i.e. “nocturnal pain and total back pain” were replaced by “total back pain”. Comparison between two treatment groups was performed using the Cochran-Mantel-Haenszel chi square test, stratified by positive or negative sacroiliitis MRI and geographic region. For the analysis of the primary efficacy endpoint, the LOCF approach was used for patients who discontinued before Week 12 with an additional sensitivity analysis where these patients were counted as non-responders. A generalised estimating equations model, using a logit link, a binomial distribution and an auto-regressive correlation structure, with treatment groups, visits and their interaction as fixed factors was used as a sensitivity analysis for the primary endpoint. The Kaplan-Meier method was used to estimate the time to partial remission and compared using the log-rank test stratified by positive or negative sacroiliitis MRI and geographic region. An ANCOVA model with the terms of treatment, region, baseline sacroiliac status and baseline score was used for analysis of continuous clinical and health outcome endpoints in the double-blind period of the study. For the open-label extension, unadjusted outcomes were reported at the 48-week follow-up period and these were compared with the corresponding unadjusted 12-week double-blind results. All endpoints were tested at two-sided α=0.05 significance levels. No adjustments for multiple testing were made for secondary and supportive analyses. In addition to analysis of the overall population, a subgroup analysis of the ASAS 40 primary endpoint was also conducted. Key demographic and baseline disease characteristics were analysed to determine the impact of these characteristics on response to treatment. These included age, baseline hsCRP status, HLA-B27 status, and MRI sacroiliitis classification at screening. Responses in each subgroup category (i.e. responder count, group size and percentage response) were presented. A logistic regression model with factors for treatment subgroup and treatment by subgroup was also fitted. Post-hoc analyses were also performed to explore the effect of hsCRP and the effect of MRI sacroiliac joint scores on treatment response.

Abbreviations: ANCOVA, analysis of covariance; ASQoL, Ankylosing Spondylitis quality of life; ASAS, Assessment of Spondyloarthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; AS-WIS, Ankylosing Spondylitis Work Instability Index; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BAS-G, Bath Ankylosing Spondylitis Patient Global Assessment Score; BASMI, Bath Ankylosing Spondylitis Metrology Index; DMARD, disease modifying anti-rheumatic drug; EMA, European Medicines Agency; ESR, erythrocyte sedimentation rate; HADS, Hospital anxiety and depression scale; hsCRP, high sensitivity C reactive protein; IBD, inflammatory bowel disease; IV, intravenous; LOCF, last observation carried forward; MASES, Maastricht Ankylosing Spondylitis Enthesitis Score;

MCII, minimum clinically important improvement; MFI, Multidimensional Fatigue Inventory; mITT, modified intention to treat; MOS, Medical Outcomes Study (MOS) Sleep Scale; MRI, magnetic resonance imaging; nr-axSpA, non-radiographic axial spondyloarthritis; NSAID, non-steroidal anti-inflammatory drug; PASS, patient acceptable symptom state;; SF-36, short-form-36; SpA, spondyloarthritis; SSZ, sulfasalazine; TNF, tumour necrosis factor; WPAI, Work Productivity and Activity Impairment Questionnaire.

3.3.1.2 Patient population: Study 1031

Patient disposition – double-blind period (mITT population)

A total of 383 patients were screened for the double-blind period of the study, of which 158 were screen failures. Out of a total of 225 patients randomised for the double-blind period of the study, 215 patients comprised the mITT population which was used for all primary efficacy analyses; 106 patients in the etanercept group and 109 patients in the placebo group. By Week 12, 6 patients in the etanercept group and 3 patients in the placebo group discontinued the study.

A total of 206 patients completed the double-blind period and 205 patients entered the open-label extension study.

Open-label period (mITT population)

[redacted] patients discontinued the study during the open-label period ([redacted] patients in the etanercept group and [redacted] patients previously receiving placebo). Reasons for discontinuation included lack of response ([redacted] patients) and withdrawal of consent ([redacted] patients). A total of 189 patients (90 patients in the etanercept group and 99 in the placebo/etanercept group) reached the 48 week time point.

Patient demographics and baseline characteristics (mITT population)

There were no significant differences in patient demographics and baseline disease characteristics between treatment arms. The majority of patients were male, the average age was 32 years and the mean duration of disease symptoms was 2.4 years. There were no differences in mean NSAID scores between treatment groups. Disease activity levels as measured by the BASDAI and ASDAS and levels of functional impairment as measured by the BASFI were moderate to high in both treatment groups (Table 4).

Table 4: Patient demographics and baseline characteristics (mITT population): Study 1031

Characteristic	Etanercept (n=106)	Placebo (n=109)
Age (years), mean	31.9 (7.8)	32.0 (7.8)
Female (%)	38 (35.9)	47 (43.1)
White (%)	79 (74.5)	79 (72.5)
HLA-B27 positive, n (%)	71 (67.0)	83 (76.2)
Duration of disease symptoms (years), mean (SD)	2.4 (1.9)	2.5 (1.8)
hsCRP, mg/L, mean (SD)	6.8 (10.6)	6.4 (10.5)
Elevated hsCRP, n (%)	48 (45.3)	44 (40.4)
BASDAI total score, mean (SD)	6.0 (1.8)	6.0 (1.9)
BASFI, mean (SD)	4.2 (2.5)	3.9 (2.5)
ASDAS – hsCRP, mean (SD)	3.0 (0.9)	3.0 (1.0)
[redacted]	[redacted]	[redacted]

Patient global assessment, mean (SD)	5.8 (2.3)	5.8 (2.1)
[redacted]	[redacted]	[redacted]
Total back pain (cm), mean (SD)	5.5 (2.4)	5.5 (2.4)
[redacted]	[redacted]	[redacted]
SPARCC MRI sacroiliac score, mean (SD)	8.0 (9.7)	7.7 (10.1)
SPARCC MRI 6DVU spinal score, mean (SD)	4.7 (7.1)	3.5 (5.6)
MRI sacroiliitis positive†, n (%)	87 (82.1)	87 (79.8)
MRI sacroiliitis positive† or elevated hsCRP, n (%)	94 (88.7)	95 (87.2)
MASES total score, mean (SD)	3.1 (3.3)	2.9 (3.1)
Swollen joint count, mean (SD)	0.9 (2.0)	0.9 (2.2)
Tender joint count, mean (SD)	3.3 (4.7)	4.1 (6.2)
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
NSAID score (-4 to 0 weeks), mean (SD)	85.1 (25.7)	77.8 (30.2)
Concomitant DMARD use, n (%)	21 (19.8)	21 (19.3)
[redacted]	[redacted]	[redacted]

Abbreviations: ASDAS, Ankylosing Spondylitis Disease Activity Score; ASQoL, Ankylosing spondylitis quality of life; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; DMARD, disease modifying anti-rheumatic drug; DVU, Discovertebral units; ESR, erythrocyte sedimentation rate; hsCRP, high sensitivity C reactive protein; MASES, Maastricht Ankylosing Spondylitis Enthesitis Score; MRI, magnetic resonance imaging; NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation; SF-36, short form-36; SPARCC, Spondyloarthritis Research Consortium of Canada.
†Baseline MRI scans were read by a central imaging reader.

3.3.1.3 Results: Study 1031

Primary efficacy outcome – double-blind period

The primary endpoint of ASAS 40 response is presented using the 2005 EMA guidelines. However, for completeness, the primary endpoint was also analysed using the 2009 EMA guidelines (Appendix B). At Week 12, a significantly greater proportion of patients achieved ASAS 40 in the etanercept group compared with the placebo group (32% vs 16%, respectively; p=0.0062). Sensitivity analysis using a generalised estimating equations model confirmed the primary analysis results.

Secondary efficacy outcomes- double-bind period

At Week 2 and Week 12, a significantly greater proportion of patients in the etanercept group achieved ASAS 20 compared with the placebo group (30.5% vs 16.0%, respectively; p<0.05 and 52.4% vs 36.1%, respectively; p<0.05). Partial remission was achieved in [redacted]% of etanercept patients compared with [redacted]% of placebo patients at Week 12 ([redacted]).

Significantly greater improvements in clinical and imaging endpoints were observed in the etanercept group compared with the placebo group, with the exception of the BAS-G, dactylitis and mobility endpoints (excluding lateral flexion). Significant differences in favour of the etanercept group versus the placebo group were observed for the following outcomes (Table 5):

- BASDAI 50 (p<0.01)

- ASDAS hsCRP measures of disease activity (p<0.0001)
- Improvements in total BASDAI score and total BASFI score (p=0.0186 and p=0.0164, respectively)
- Improvements in SPARCC MRI sacroiliac joint scores (-3.8 vs -0.8, p<0.001, mean % change from baseline (46.9% etanercept vs 10.9% placebo)
- Improvements in SPARCC-Spine 6 (DVU) total score (-2.12 vs -1.16; p=0.041), mean % change from baseline (45.4% etanercept vs 33.4% placebo)

Table 5: Secondary efficacy outcomes (mITT population): Study 1031 - double-blind period

Outcome	Etanercept† (n=106)	Placebo† (n=109)	p value
Clinical outcomes (Week 12)			
ASAS 20 (%)	52.4	36.1	<0.05
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
Patient assessment of disease activity	-2.1 (0.3)	-1.3 (0.3)	0.010
[redacted]	[redacted]	[redacted]	[redacted]
Total back pain	-2.0 (0.3)	-1.1 (0.3)	0.006
BASDAI morning stiffness (2 items)	-2.3 (0.3)	-1.4 (0.3)	0.0134
BASMI	-0.3 (0.2)	-0.3 (0.1)	0.687
hsCRP	--3.0 (1.1)	0.1 (1.0)	0.004
ASDAS Inactive Disease (%)	40.00	17.4	<0.001
ASDAS hsCRP	-1.1 (0.1)	-0.5 (0.1)	<0.0001
BASDAI total score	-2.0 (0.3)	-1.3 (0.3)	0.019
BASDAI 50 (%)	43.8	23.9	<0.01
BASFI total score	-1.4 (0.2)	-0.8 (0.2)	0.016
[redacted]	[redacted]	[redacted]	[redacted]
BASMI (mean change VAS, cm)	-0.3 (0.2)	-0.3 (0.1)	0.687
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
Imaging endpoints (Week 12)			
SPARCC sacroiliac joint score	-3.8 (0.7)	-0.8 (0.6)	<0.001
SPARCC-Spine 6 (DVU) total score	-2.1 (0.5)	-1.2 (0.5)	0.041
[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ASDAS, Ankylosing Spondylitis Disease Activity Score; ASAS, Assessment of Spondyloarthritis International Society; ASspiMRI-a, Ankylosing Spondylitis spine MRI score for activity; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath

Ankylosing Spondylitis Metrology Index; DVU, Discovertebral units; hsCRP, high sensitivity C reactive protein; SPARCC, Spondyloarthritis Research Consortium of Canada.
†Unless otherwise stated, values are presented as mean change from baseline (SE).

At Week 12, QoL improvements in favour of the etanercept group were observed for a number of patient reported outcome measures including the EQ-5D VAS, subscales of the WPAI and SF-36 physical component score [redacted] (Table 6).

Table 6: Improvement from baseline in QoL measures (mITT population): Study 1031 – double blind period

Outcome	Baseline	Etanercept†	Placebo†
EQ-5D improvement >0.05 from baseline, n/N (%)	-	51/85 (60.0)	40/93 (43.0)
EQ-5D VAS	56.4 (20.7)	67.4 (21.2)	60.9 (22.3)
SF-36 physical component score	37.5 (8.5)	43.7 (8.9)	41.0 (7.8)
WPAI – % impairment while working	43.6 (26.4)	25.2 (24.2)	33.2 (26.0)
WPAI –% work time missed	10.4 (26.3)	9.6 (24.6)	5.9 (17.4)
[redacted]	-	[redacted]	[redacted]
[redacted]	-	[redacted]	[redacted]
[redacted]	-	[redacted]	[redacted]

Abbreviations: ASQoL, Ankylosing Spondylitis Quality of Life; AS-WIS, Ankylosing Spondylitis Work Instability Index; HADS, Hospital Anxiety and Depression Scale; SF-36, short form-36; VAS, visual analogue scale; WPAI, Work Productivity and Activity Impairment Questionnaire.
†Data are presented as adjusted mean (SE) change from baseline. Change = baseline score + treatment + region + sacroiliac status

Post-hoc analysis – double-blind period

Etanercept is licensed for “the treatment of adults with severe non-radiographic axial spondyloarthritis with objective signs of inflammation as indicated by elevated C-reactive protein (CRP) and/or magnetic resonance imaging (MRI) evidence, who have had an inadequate response to non-steroidal anti-inflammatory drugs.” However, the 1031 study included a small population of patients who did not meet the licensing indication for etanercept (i.e. patients who had normal CRP and negative MRI).

Patients with elevated CRP levels at baseline had a numerically higher ASAS 40 response to etanercept treatment than those with normal CRP levels, although the results were not statistically significant. Similarly, a numerically higher response to etanercept was observed among patients with Spondyloarthritis Research Consortium of Canada (SPARCC) scores ≥2 (42% etanercept patients vs 18% placebo patients).

Pre-specified post-hoc subgroup analysis of the primary efficacy outcome (ASAS 40 response) at Week 12 showed that the study population with either elevated CRP levels or positive MRI sacroiliitis at baseline have a marginally greater benefit with etanercept treatment compared with the mITT population (Table 7). These data suggested that patients with higher CRP levels at baseline may have higher ASAS 40 response rates to etanercept

treatment, and provides evidence from study 1031 suggestive of baseline CRP levels and evidence of MRI sacroiliitis at baseline being potential treatment effect modifiers.

Based on the results of the post-hoc ASAS 40 subgroup analysis, it is likely that the mITT analyses are underestimating the benefit of etanercept in the licensed population. While subgroup analyses of the BASDAI and BASFI outcomes for the licensed population have not been conducted, it is reasonable to suggest that a subgroup analysis of the results from the mITT analyses on the BASDAI and BASFI would produce a similar pattern of results. In the context of the economic model, the results are therefore likely to provide a conservative estimate of the clinical benefit of etanercept compared with conventional therapy.

Table 7: Subgroup analysis of the ASAS 40 response in patients with elevated CRP levels or positive MRI at baseline

	mITT population			**Elevated baseline hsCRP or positive MRI**		
	Etanercept n=105	**Placebo N=108**	**Difference in proportions**	**Etanercept n=[redacted]**	**Placebo n=[redacted]**	**Difference in proportions**
ASAS 40 (n/N, %)	34/105 (32.4)	17/108 (15.7)	16.6%	[redacted]	[redacted]	[redacted]

Abbreviations: ASAS 40, Assessment of Spondyloarthritis International Society; hsCRP, high sensitivity C reactive protein; mITT, modified intent-to-treat; MRI, magnetic resonance imaging.

The unadjusted efficacy results for the double-blind treatment period of the 1031 study are presented in Table 8 together with the results at the 48-week time-point from the open-label extension phase of the study

Open-label period (mITT population)

At week 48 in a combined population 52.7% of patients met ASAS40. In the pooled analysis group (n=205), an ASAS 40 response was achieved in 52.7% of patients.

[redacted] 8[redacted]

[redacted] Table 8).

During the open-label period, improvement from baseline was shown in the SPARCC sacroiliac joint score with the response at Week 48 (Table 8) corresponding to a mean percent change of 65.2% for pooled open-label patients. Improvements were also seen in the SPARCC-spine 6 DVU total score (Table 8) corresponding to a mean percent change of 60.8% for pooled open-label patients. The mean change in ASspiMRI-a total score was maintained with the response at Week 48 (Table 8) corresponding to a mean percent change of 40.9%.

Table 8: Efficacy outcomes at Week 48, open-label period (mITT population): Study 1031

	Week 12 (Double-blind period)		**Week 48 (Open-label period)**
Outcome†	**Etanercept 50 mg (n=105)**	**Placebo (n=109)**	**Etanercept 50 mg, pooled analysis (n=205)**

[redacted]	[redacted]	[redacted]	[redacted]
ASAS 5/6, n (%)	[redacted](34.3)	[redacted](11.9)	97/205 (47.3)
[redacted]	[redacted]	[redacted]	[redacted]
ASDAS hsCRP inactive disease, n (%)	42/105 (40.0)	19/109 (17.4)	[redacted]
BASDAI 50, %	43.8	23.9	62.4
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ASDAS, Ankylosing Spondylitis Disease Activity Score; ASAS, Assessment of Spondyloarthritis International Society; ASspiMRI-a, Ankylosing Spondylitis spine MRI score for activity, BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; DVU, Discovertebral units; hsCRP, high sensitivity C reactive protein; SPARCC, Spondyloarthritis Research Consortium of Canada.
†Values are presented as the (unadjusted) mean change from baseline.

3.3.1.4 Efficacy conclusion: Study 1031

The results of the 12-week double-blind period of study 1031 demonstrates that treatment with etanercept in patients with nr-axSpA, who had an inadequate response to ≥2 NSAIDs, was associated with rapid and significant improvements from baseline in clinical and imaging outcomes. In addition, patients receiving etanercept demonstrated improvement in QoL measures for disease specific and functional domains compared with placebo. Week 48 data demonstrates that improvements in clinical and health outcomes achieved over the double blind study period were sustained or continued to improve in those patients continuing with etanercept in the open-label study.

3.4 Investigator initiated research: etanercept for the treatment of early AxSpA

One RCT and two follow-up reports of the same study are presented in this section.

3.4.1 Study: ESTHER/Song, 2011

3.4.1.1 Methodology: ESTHER trial/Song, 2011

The source of the efficacy data for the ESTHER trial is Song et al, 2011 (15), (NCT00844142). This was an, investigator initiated research sponsored by Pfizer. The ESTHER trial was a randomised, multicentre, open-label trial to investigate the effects of etanercept versus SSZ on active inflammatory lesions by whole-body MRI in patients with early AxSpA with a symptom duration of less than five years (15).The study population is therefore comprised of both nr-axSpA and AS patients. A summary of the methodology in the ESTHER trial is provided in Table 9. Note that at the time of designing this study, clinical opinion was still that SSZ could provide some benefit to AxSpA patients, and as such the comparator for the trial was relevant to clinical practice.

Table 9: Methodology: ESTHER trial /Song, 2011

Study	ESTHER trial/Song, 2011 (NCT00844142)
Summary	A randomised, multicentre, open-label trial to investigate the effects of etanercept vs SSZ on active inflammatory lesions by whole-body MRI in patients with early AxSpA.
Objectives	To evaluate the effect of treatment with etanercept in comparison with SSZ on active bony inflammation by whole-body MRI in patients with early AxSpA and a symptom duration <5 years.
Study location	Not reported
Design details	Patients were randomised to receive either etanercept or SSZ over 48 weeks. MRI assessments were performed at Weeks 0, 24 and 48 by two radiologists who were blinded to both treatment arm and MRI assessment point.
Interventions	Patients were randomised to receive either: • Etanercept 25 mg subcutaneous injection, twice weekly, or • Oral SSZ 2–3 g/day (patients with intolerance to SSZ could switch to oral methotrexate 15–20 mg, once weekly)
Key inclusion criteria	• Aged 18–50 years • Diagnosis of AxSpA with a symptom duration <5 years (defined as ≥3 months chronic back pain with onset <45 years of age). • Active inflammatory lesions (osteitis/bone marrow oedema) in either the spine or sacroiliac joints as evidenced by whole-body MRI. • Three of the following: inflammatory back pain; good or very good response to NSAIDs; one or more extra-spinal manifestations such as uveitis, peripheral arthritis, enthesitis; HLA-B27 positive; family history of spondyloarthritis • Fulfilment of the ASAS criteria for AxSpA • BASDAI score ≥4 and a back pain score ≥4, despite NSAID therapy
Key exclusion criteria	• DMARD or prednisone (≥7.5 mg/day) therapy had to be discontinued at least 4 weeks prior to the study
Efficacy outcomes	<u>Primary efficacy endpoint</u> • The change in active inflammatory lesions in the spine and sacroiliac joints as detected by MRI at Week 48

Study	ESTHER trial/Song, 2011 (NCT00844142)
	Secondary efficacy endpoints • MRI evidence of reduced active inflammatory lesions on the posterior elements of the spine • Evidence of reduced peripheral enthesitis on MRI • Achievement of 50% improvement on the BASDAI • Achievement of ASAS 20, ASAS40 and ASAS partial remission Other outcomes • Improvements in BASDAI, BASFI and BASMI scores, swollen joints count (64 joints), a modified enthesitis Maastricht AS enthesitis score, patient's and physician's global assessments of disease activity and CRP levels.
Populations analysed	Efficacy analysis was performed as an ITT analysis.
Statistical analysis	Comparisons between treatment groups at baseline were performed using a Mann-Whitney test. A non-parametric ANCOVA was used to compare changes in spine and sacroiliac joint MRI scores. The possible influence of the diagnosis of nr-axSpA and disease characteristics at baseline on the primary outcome was assessed by including these parameters as covariables in the ANCOVA model as a sensitivity analysis. The LOCF approach was used to handle missing MRI data. Secondary outcomes were analysed using the non-parametric or parametric ANCOVA as appropriate, with the baseline value as a covariable. The correlation between changes in clinical endpoints and change in MRI score after adjustment for treatment were assessed using the parametric ANCOVA. The threshold value for statistical significance was 0.05. Repeated significance testing of the primary outcome was adjusted by application of the Holm procedure.

Abbreviations: ANCOVA, analysis of covariance; AxSpA, axial spondyloarthritis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; HLA, human leucocyte antigen; ITT, intention to treat; LOCF, last observation carried forward; MRI, magnetic resonance imaging; NSAID, non-steroidal anti-inflammatory drug; SSZ, sulfasalazine.

3.4.1.2 Patient population: ESTHER trial/Song, 2011

Patient disposition

A total of 76 patients fulfilled the inclusion criteria for the study; 40 patients were randomised to etanercept treatment and 36 were randomised to SSZ. The mean dose of SSZ was 1688 mg (n=32) at Week 24 and 1650 mg (n=30) at Week 48. Due to intolerance to SSZ, 4 patients switched to methotrexate therapy (15.0 mg /week). Seven patients discontinued before Week 24 and a further 3 patients discontinued before Week 48 (5 patients in each treatment group). Reasons for discontinuation included lost to follow up (both treatment arms) and side effects and lack of efficacy in the SSZ group only.

Patient demographics and baseline characteristics

There were no significant differences in patient demographics and baseline disease characteristics between treatment arms (Table 10). At baseline, active inflammatory lesions were present in the sacroiliac joints of 94.6% of patients and in the spine in 47.3% of patients. Only 5.4% of patients had active inflammatory lesions in the spine but not in the sacroiliac joints.

Table 10: Patient demographics and baseline disease characteristics: ESTHER trial/ Song, 2011

Characteristic	Etanercept (n=40)	SSZ (n=36)	Total (n=76)
Age, years (mean, ± SD)	34.5 (8.6)	32.8 (8.4)	33.7 (8.5)
Disease duration, years (mean, ± SD)	2.6 (1.7)	3.0 (1.8)	2.9 (1.7)
Male, n (%)	23 (57.5)	21 (58.3)	44 (57.9)
HLA-B27 positive, n (%)	34 (85)	28 (77.8)	62 (81.6)
Elevated CRP (>5 mg/L), n (%)	24 (60.0)	16 (44.4)	40 (52.6)
Clinical arthritis, n (%)	15 (37.5)	18 (50.0)	33 (43.4)
Clinical enthesitis, n (%)	25 (62.5)	24 (66.7)	49 (64.5)
Fulfilled New York criteria, n (%)	20 (50)	19 (52.8)	39 (51.3)

Abbreviations: CRP, C reactive protein; HLA, human leucocyte antigen; SD, standard deviation; SSZ, sulfasalazine.

3.4.1.3 Results: ESTHER trial/Song, 2011

Primary efficacy outcome

At Week 48, patients treated with etanercept showed a significant reduction in inflammation of both the spine and the sacroiliac joint compared with SSZ treated patients (Table 11).

Table 11: Primary efficacy outcome: ESTHER trial/Song, 2011

	Etanercept (n=40)		SSZ (n=36)		
Parameter	**Baseline†**	**Week 48**	**Baseline†**	**Week 48**	**p value‡ (ETN vs SSZ)**
MRI sacroiliac joint score (0–24)	7.8 (6.3)	2.4 (3.2)	5.4 (5.1)	3.5 (3.8)	0.02
MRI spine score (0–69)	2.3 (3.5)	1.0 (2.1)	1.4 (3.1)	1.3 (2.9)	0.01

Abbreviations: ETN, etanercept; MRI, magnetic resonance imaging; SSZ, sulfasalazine.
†No significant differences were observed between etanercept and sulfasalazine groups at baseline.
‡Differences in MRI scores analysed by analysis of covariance.

Secondary efficacy outcome

At Week 48, patients treated with etanercept showed a significant reduction in peripheral enthesitis on MRI compared with SSZ patients. Etanercept patients also showed a numerical improvement in active inflammatory lesions on the posterior elements of the spine compared with SSZ patients, but this was not statistically significant (Table 12).

Six patients in the etanercept group and two patients in the SSZ group became free of any active inflammatory lesions on MRI in the sacroiliac joints, the spine, entheses and posterior elements. These six etanercept patients who were completely free of inflammation also achieved clinical remission according to ASAS criteria. However, neither of the two patients in the SSZ group who reached clinical remission were free from inflammation as evidenced by MRI.

Table 12: Secondary efficacy outcome: ESTHER trial/Song, 2011

	Etanercept (n=40)		SSZ (n=36)		
Parameter	**Baseline†**	**Week 48**	**Baseline†**	**Week 48**	**p value‡ (ETN vs SSZ)**
MRI enthesitis	26 sites (in 15 patients)	11 sites (in 11 patients)	24 sites (in 16 patients)	26 sites (in 14 patients)	0.04
MRI posterior segments of the spine	37 sites (in 9 patients)	26 sites (in 2 patients)	18 sites (in 4 patients)	13 sites (in 1 patient)	0.92

Abbreviations: ETN, etanercept; MRI, magnetic resonance imaging; SSZ, sulfasalazine.
†No significant differences were observed between etanercept and sulfasalazine groups at baseline.
‡Differences in MRI scores analysed by analysis of covariance.

There was a significant reduction in all clinical outcome parameters in the etanercept group compared with the SSZ group with the exception of BASMI, CRP levels, swollen joints, and clinical enthesitis (

Table 13). There was a statistically significant relationship between changes in MRI sacroiliac joint scores and changes in the BASDAI and BASFI (p=0.04 and p=0.0069, respectively). At Week 48, 50% of etanercept patients reached ASAS partial remission compared with 19% of patients in the placebo group.

Table 13: Other clinical outcome parameters: ESTHER trial/Song, 2011

	Etanercept† (n=40)		SSZ† (n=36)		
Parameter	**Baseline‡**	**Week 48**	**Baseline‡**	**Week 48**	**p value§ (ETN vs SSZ)**
BASDAI	5.5 (1.3)	2.5 (2.0)	6.0 (1.2)	4.4 (2.4)	0.001
BASFI	4.3 (2.3)	2.0 (2.1)	4.3 (1.8)	3.3 (2.2)	0.001
Patient global assessment	6.7 (2.1)	2.6 (2.2)	7.1 (1.6)	4.9 (3.0)	<0.001
Physician global assessment	6.4 (1.2)	1.8 (1.8)	6.1 (1.5)	4.1 (2.9)	<0.001
Joint count	2.3 (5.8)	0.2 (0.7)	1.3 (1.7)	0.3 (1.5)	0.48
Enthesitis score	4.4 (4.6)	1.8 (4.2)	3.4 (3.4)	1.4 (2.9)	0.70
BASMI	1.9 (1.7)	1.6 (1.8)	1.7 (1.4)	1.9 (1.7)	0.32
EQ-5D	0.6 (0.3)	0.8 (0.2)	0.6 (0.3)	0.7 (0.3)	0.047
AS-QoL	9.7 (4.5)	4.4 (4.8)	9.3 (3.6)	7.5 (5.4)	<0.001
CRP level (ref 5mg/L)	11.9 (13.2)	4.3 (3.7)	10.6 (14.9)	8.7 (12.5)	0.13

Abbreviations: AS-QoL; ankylosing spondylitis quality of life questionnaire; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; ETN, etanercept; SSZ, sulfasalazine.
†All values presented as mean ± SD
‡ No significant differences were observed between etanercept and sulfasalazine groups at baseline.
§ Differences in MRI scores analysed by analysis of covariance.

3.4.1.4 Efficacy conclusion: ESTHER trial/Song, 2011

The results from the ESTHER trial demonstrate that treatment with etanercept 25 mg twice weekly in patients with early AxSpA is associated with significantly greater reductions in active inflammatory lesions as evidenced by whole-body MRI when compared with SSZ treated patients at 48 weeks. These results correlated with a good clinical response in patients receiving etanercept.

3.4.2 Study: Song, 2012

3.4.2.1 Methodology and efficacy outcomes: Song, 2012

The source of the efficacy data for Song, 2012 is Song et al (162). This was a follow up open-label extension study to the ESTHER trial. Patients with early active AxSpA and a disease duration <5 years who had received etanercept or SSZ treatment for 1 year were assessed at Year 2 to determine the frequency and duration of drug-free remission in patients treated with etanercept after experiencing a flare of disease. The methodology and study outcomes for Song, 2012, are presented in Table 14.

Table 14: Methodology and study outcomes: Song, 2012

Study	Song, 2012
Objectives	The two aims of this follow-up study were: 1. To evaluate the frequency and duration of drug free remission in patients with early active AxSpA who had previously received etanercept or SSZ for 1 year in the ESTHER trial. 2. To assess the efficacy of etanercept therapy at Year 2 in those patients who did not reach remission at Week 48 of the ESTHER trial.
Design details	Patients who had received etanercept or SSZ for 48 weeks in the ESTHER trial and had achieved ASAS plus MRI remission by Week 48 were followed up every 6 weeks without active treatment up to 1 year. Patients who experienced a flare of disease (defined as an increase of 2 points from Week 48 on the BASDAI) resumed etanercept treatment for a further year. Patients who did not have a flare were removed from the study at the end of Year 2 and were classified as being in permanent drug-free remission. Patients who had not achieved ASAS plus MRI remission at Week 48 of the ESTHER trial (both etanercept and SSZ patients) were treated with etanercept in Year 2. Patients were scored for active inflammation using whole-body MRI at Week 104.
Interventions	• Patients had previously received etanercept 25 mg subcutaneous injection, twice weekly over 48 weeks
Key inclusion criteria	• Full inclusion criteria are provided in section 3.4.1.1, Table 9
Key exclusion criteria	• Full exclusion criteria are provided in section 3.4.1.1, Table 9
Efficacy outcomes	• Proportion of patients remaining in drug-free remission • In those patients experiencing a flare and receiving etanercept therapy – achievement of ASAS remission and improvement in the BASDAI, ASDAS, CRP levels and MRI imaging variables
Populations analysed	Investigations were performed on the ITT population from the ESTHER trial.
Statistical analysis	Analysis was performed as an ITT, LOCF analysis. Changes within groups were analysed using the one-sample t-test and Wilcoxon rank test. The 95% CIs were calculated using the Blyth-Still-Casella method.
Patient disposition	Out of a total of 76 patients, 60 patients completed the study by the end of Year 2. There were 12 patients in the remission-flare group (9 patients had previously received etanercept and 3 patients had previously received SSZ. There were 48 patients in the non-remission group (22 patients had previously received etanercept and 26 patients had previously received SSZ).
Patient demographics and baseline characteristics	Full details are provided in section 3.4.1.2, Table 10 Table 10

Study	Song, 2012
Efficacy outcomes	• During Year 2, 4/17 patients (23.5%) remained in drug-free remission (3/13 etanercept patients and 1/4 SSZ patients). • There was no significant difference in the mean time to flare between the etanercept group (24.4 weeks) and the SSZ group (39.6 weeks) and no difference in the proportion of patients with flares between these two groups (69% of etanercept patients vs 75% of SSZ patients). • At Week 108, a significant and sustained improvement in the BASDAI and ASDAS ($p<0.05$) and a sustained MRI response was observed in patients previously treated with etanercept not reaching ASAS plus MRI remission. ASAS remission was achieved in 56% of these patients by Week 108 while 13.6% of non-remission ETN patients reached ASAS plus MRI remission. • Significant improvements in the BASDAI ($p=0.0013$) and ASDAS ($p=0.0019$) were observed in patients switching from SSZ to etanercept in Year 2. Improvements seen in MRI and sacroiliac joint scores in this patient group were not statistically significant.
Study conclusion	There were no statistically significant differences in drug-free remission rates of AxSpA patients treated with etanercept over 1 year compared with patients treated with SSZ. However, the results of this study demonstrated that resultant flares responded well to resumed etanercept therapy. BASDAI and ASDAS responses were sustained for up to 2 years in patients on SSZ in Year 1 who switched to etanercept in Year 2.

Abbreviations: AxSpA, axial spondyloarthritis; ASAS, Assessment of Spondyloarthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CI, confidence interval; ITT, intention-to-treat; LOCF, last observation carried forward; MRI, magnetic resonance imaging; SSZ, sulfasalazine.

3.4.3 Study: Song 2013

3.4.3.1 Methodology: Song, 2013

The source of the efficacy data for Song, 2013 is Song et al (16). This was a follow up subgroup analysis of the ESTHER trial. As the study cohort in the ESTHER trial was a mixed AxSpA population, the aim of Song, 2013 was to assess whether there were any treatment effect differences in two subgroups of patients who had received etanercept for one year; nr-axSpA patients and AS patients. The methodology for Song, 2013 is presented in Table 15.

Table 15: Methodology: Song, 2013

Study	Song, 2013
Objectives	To investigate any differences in treatment effect with etanercept in patients with nr-axSpA compared with AS patients with a disease duration <5 years.
Design details	Patients who received etanercept in the ESTHER trial for 1 year were compared for differences in baseline characteristics and treatment effect. Patients had previously been scored for active inflammation using whole-body MRI of the sacroiliac joints and spine by radiologists blinded to the treatment arm.
Interventions	• Patients (n=40) had previously received etanercept 25 mg subcutaneous injection, twice weekly over 48 weeks
Key inclusion criteria	• Full details are provided in section 3.4.1.1, Table 9
Key exclusion criteria	• Full details are provided in section 3.4.1.1, Table 9
Efficacy outcomes	• Achievement of ASAS 40 • Achievement of ASAS partial remission • 50% improvement on the BASDAI • Scores for inactive disease, clinically important change, and major improvement on the ASDAS • Improvements in the BASDAI, BASFI, patient's and physician's global assessments of disease activity, and CRP levels
Populations analysed	Investigations were performed on the etanercept group from the ITT population in the ESTHER trial.
Statistical analysis	A mixed-model approach was used to compare outcome parameters between nr-axSpA and AS patients after adjusting for baseline. This approach also allowed the inclusion of dropouts (n=3) in the analysis. The 95% CIs of response rates were calculated using the Blyth-Still-Casella method.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; CRP, C reactive protein; ITT, intention-to-treat; MRI, magnetic resonance imaging; nr-axSpA, non-radiographic axial spondyloarthritis.

3.4.3.2 Patient population: Song, 2013

A total of 40 patients who had received 1 year of etanercept treatment in the ESTHER trial were included in the observational study: patients with nr-axSpA, n=20; patients with AS, n=20. With respect to baseline characteristics, there were no significant differences between patient groups (Table 16).

Table 16: Baseline demographics of patients with nr-axSpA and AS: Song, 2013

Parameter	Etanercept nr-axSpA n=20	Etanercept AS n=20
Age in years, mean (±SD)	33.3 (±9.1)	33.8 (±8.2)
Disease duration in years, mean (±SD)	2.6 (±1.8)	2.7 (±1.6)
Male (%)	55	60
HLA-B27 positive, (%)	80	90
BASDAI (0–10), mean (±SD)	5.6 (±1.3)	5.5 (±1.4)
CRP (<5 mg/l), mean (±SD)	14.4 (±15.7)	9.4 (±9.9)
ASDAS (0–∞), mean (±SD)	3.6 (±0.9)	3.2 (±0.7)
BASFI (0–10), mean (±SD)	4.5 (±2.4)	4.2 (±2.2)
MRI spine score (0–69), mean (±SD)	1.3 (±1.9)	3.4 (±4.3)
MRI SI-joint score (0–24), mean (±SD)	7.1 (±5.9)	8.6 (±6.8)

Abbreviations: AS, ankylosing spondylitis; ASDAS; Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; HLA, human leucocyte antigen; MRI, magnetic resonance imaging; nr-axSpA, non-radiographic axial spondyloarthritis; SD, standard deviation; SI-joint, sacroiliac joint.

3.4.3.3 Results: Song, 2013

Efficacy outcome parameters

There were no significant differences in BASDAI, ASAS and ASDAS response rates between the nr-axSpA and AS patient groups after 1 year of treatment with etanercept. Similarly, no significant differences were observed between CRP levels, MRI scores or in BASFI scores. However, non-significant, small differences in favour of the nr-axSpA group were observed in the BASDAI 50 (75% vs 55%), ASDAS clinically important improvement (75% vs 55%), and ASAS partial remission (60% vs 40%) outcomes (Table 17).

Table 17: BASDAI, ASAS and ASDAS response rates in patients with nr-axSpA vs patients with AS, Song, 2013

Parameter	Etanercept nr-axSpA n=20	Etanercept AS n=20	p value
BASDAI 20, % (95% CI)	85 (65–96)	80 (59–93)	>0.05
BASDAI 50, % (95% CI)	75 (54–90)	55 (34–75)	>0.05
ASAS 40, % (95% CI)	65 (43–83)	75 (54–90)	>0.05
ASAS partial remission, % (95% CI)	60 (38–79)	40 (21–62)	>0.05
ASDAS clinically important improvement (%) †	75 (54–90)	55 (34–75)	>0.05
ASDAS major improvement (%) ‡	25 (10–46)	30 (14–52)	>0.05
ASDAS inactive disease (<1.3), % (95% CI)	40 (21–62)	40 (21–62)	>0.05

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; ASDAS; Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CI, confidence interval; nr-axSpA, non-radiographic axial spondyloarthritis.
†ASDAS clinically important improvement: ≥1.1 change on the ASDAS between baseline and follow-up.
‡ASDAS major improvement: ≥2.0 change on the ASDAS between baseline and follow-up.

3.4.3.4 Conclusion: Song, 2013

The results of this analysis demonstrate that in the context of similar baseline characteristics such as duration of symptoms and disease activity, etanercept is similarly efficacious in the treatment of patients with nr-axSpA and AS. Response rates in the BASDAI, ASDAS and ASAS 40 were not significantly different between treatment groups. This study therefore demonstrates that etanercept offers a broad range of effectiveness over different disease stages of AxSpA.

3.5 Phase III RCTs and related open-label extension studies: etanercept for the treatment of AS

Fourteen studies assessing the efficacy of etanercept are presented in this section; five placebo-controlled RCTs, seven open-label extension studies, and two observational extension studies. A summary of all included studies is provided in Table 18 (see also Figure 4).

Table 18: Summary of Phase III RCTs and related open-label extension studies of etanercept for the treatment of AS

Study (Study no/relevant publications)	Design	Study duration	Intervention and Population (N)
314-EU/van der Heijde, 2006 (12)	Double-blind, placebo-controlled, non-inferiority RCT	12 weeks	• ETN 50 mg once weekly: n=155 • ETN 25 mg twice weekly: n=150
16.0037/Davis, 2003 (10)	Double-blind, placebo-controlled RCT	24 weeks	• ETN 25 mg twice weekly: n=138 • Placebo: n=139
16.0040/Davis, 2005 (119)	Open-label extension study of 16.0037	72 weeks	ETN 25 mg twice weekly, n=257: • ETN/ETN†: n=128 • Placebo/ETN‡: n=129
Davis, 2008 (118)	Second open-label extension study of 16.0037	168 weeks	ETN 25 mg twice weekly, n=257: • ETN/ETN†: n=128 • Placebo/ETN‡: n=129
Van der Heijde, 2008 (124)	Observational follow-up study to 16.0037	72 weeks	• ETN 25 mg twice weekly: n=257
Boonen, 2008 (22)	Observational follow-up study of 16.0337 and 16.0040	72 weeks	• ETN 25 mg twice weekly: n=257
311-EU/Calin, 2004 (11)	Double-blind, placebo-controlled RCT	12 weeks	• ETN 25 mg twice weekly: n=45 • Placebo: n=39
312-EU/Dijkmans, 2009 (120)	Open-label extension study of 311-EU	96 weeks	ETN 25 mg twice weekly, n=81: • ETN/ETN†: n=42 • Placebo/ETN‡: n=39
907-EU/Martin-Mola (122, 163)	Second open-label extension study of 312-EU	156 weeks	• ETN 25 mg twice weekly, n=59
SPINE/Dougados, 2011 (13)	Double-blind, placebo-controlled RCT	12 weeks	• ETN 50 mg once weekly: n=39 • Placebo: n=43
Dougados, 2012 (121)	Open-label extension study of the SPINE trial	12 weeks	ETN 50 mg once weekly: n=77
Brandt, 2003 (89)	Double-blind, placebo-controlled RCT followed by an observational phase	RCT: 6 weeks Observational phase: 24 weeks	Double-blind period: • ETN 25 mg twice weekly: n=14 • Placebo: n=16 Observational phase: • ETN/ETN: n=14 • Placebo/ETN: n=15
Brandt, 2005 (116)	Open-label, observational extension study of Brandt, 2003	54 weeks	ETN 25 mg twice weekly: n=26

Baraliakos,2013 (9)	Open-label extension study of Brandt, 2003	7 years	ETN 25 mg twice weekly: n=21

Abbreviations: ETN, etanercept; RCT, randomised controlled trial;
†Patients had received etanercept in the previous study and continued to receive etanercept in the open-label extension study
‡Patients had switched from placebo to etanercept in the open-label extension study

3.5.1 Study 314-EU/van der Heijde, 2006

3.5.1.1 Methodology: Study 314-EU/van der Heijde, 2006

The source of the efficacy data for study 314-EU is van der Heijde et al, 2006 (12) (NCT00418548). This was a Phase III, multicentre, randomised, double-blind, placebo-controlled non-inferiority trial to compare the efficacy, pharmacokinetics and safety of etanercept 50 mg once weekly with etanercept 25 mg twice weekly and placebo in patients with AS. A summary of the methodology for study 314-EU/van der Heijde, 2006 is provided in Table 19**.**

Table 19: Methodology: study 314-EU/van der Heijde, 2006

Study	**314-EU/van der Heijde, 2006 (NCT00418548)**
Objectives	A twelve week study to determine the non-inferiority of etanercept 50 mg once weekly vs etanercept 25 mg twice weekly on the basis of the ASAS working group criteria (ASAS 20). A placebo group was included to quantify the magnitude of the treatment effect.
Study location	The study was conducted at 38 centres in 11 European countries: UK, Belgium, France, Germany, Greece, Hungary, Italy, The Netherlands, Poland, Portugal, and Spain.
Design details	Assessments were carried out at screening, baseline and at Weeks 2, 4, 8 and 12.
Interventions	Patients were randomised in a 3:3:1 ratio to receive either: • Etanercept 50 mg once weekly • Etanercept 25 mg twice weekly • Placebo
Key inclusion criteria	• Aged 18–70 years • Active AS based on the modified New York criteria for AS. Active disease was defined as an average VAS score ≥30 for duration and intensity of morning stiffness and two or more of the following: patient global assessment of disease activity, BASFI scores ≥30, or mean of nocturnal and total pain VAS scores >30.
Key exclusion criteria	• Previously treated with TNF inhibitors • Use of DMARDs (excluding hydroxychloroquine, sulfasalazine and methotrexate if stable for ≥4 weeks before randomisation) <4 weeks prior to baseline • Complete ankylosis of the spine based on radiological assessment • Concurrent medical events, including congestive heart failure, unstable angina pectoris, uncontrolled hypertension, severe pulmonary disease, cancer, demyelinating diseases and serious infections
Permitted concomitant medications	• Concomitant NSAIDs and oral corticosteroids (≤10 mg/day) were permitted if stable for ≥2 weeks prior to randomisation
Efficacy outcomes	<u>Primary efficacy endpoint</u> • The percentage of patients achieving an ASAS 20 response at Week 12 <u>Secondary efficacy endpoints</u> • The proportion of patients achieving an ASAS 40 response • The proportion of responders based on ASAS 5/6 • Patient and physician global assessments of disease activity • BASFI score • BASDAI score

Study	314-EU/van der Heijde, 2006 (NCT00418548)
	• Nocturnal and total back pain assessments • Patients achieving partial remission and time to partial remission (as defined in Anderson 2001) (142) • Spinal mobility score (modified Schober's test, chest expansion and occiput to wall measurements) • Joint assessment (70 joints) • CRP levels
Other Assessments	• The pharmacokinetics of etanercept were analysed from patient blood samples collected at baseline and at Weeks 2, 4 and 12, or at early discontinuation. • An ELISA was used to determine serum etanercept concentrations
Populations analysed	Assessment of efficacy and safety was based on the modified ITT population which included all patients who received at least one dose of the study drugs.
Statistical analysis	One way ANOVA for continuous variables, Fisher's exact test or the chi-square test for categorical variables were used to analyse demographic and baseline characteristics. Non inferiority of 50 mg etanercept to 25 mg etanercept was defined as the lower limit of the one-sided 97.5% CI of their differences in the primary endpoint above a non-inferiority margin of -18.5% at Week 12. Secondary efficacy analyses were assessed using two-sided Fisher's exact tests. Missing data was handled using a last observation carried forward approach.

Abbreviations: ANOVA, analysis of variance; AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; ELISA, enzyme linked immunosorbent assay; ITT, intention-to-treat; NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor; VAS, visual analogue scale.

3.5.1.2 Patient population: study 314-EU/van der Heijde, 2006

Patient disposition

Out of a total of 361 patients randomised, 356 comprised the mITT population. A total of 35 (9.8%) patients discontinued treatment before Week 12: 3.9% of patients receiving etanercept 50 mg and 5.3% of patients receiving etanercept 25 mg discontinued due to adverse events. There were no discontinuations due to adverse events in the placebo group. In the placebo group, 5.9% of patients discontinued due to lack of efficacy compared with 1.3% and 2.0% of patients in the 50 mg and 25 mg etanercept groups, respectively

Patient demographics and baseline characteristics

With the exception of a history of inflammatory bowel disease, there were no significant differences in demographic and baseline disease characteristics between patient groups (Table 20).

Table 20: Patient demographics and baseline characteristics: study 314-EU/van der Heijde, 2006

Characteristic	Etanercept 50 mg once weekly (n=155)	Etanercept 25 mg twice weekly (n=150)	Placebo (n=51)
Age in years, mean (SD)	41.5 (11.0)	39.8 (10.7)	40.1 (10.9
Male, n (%)	108 (69.7)	114 (76.0)	40 (78.4)
Duration of disease in years, mean (SD)	9.0 (8.7)	10.0 (9.1)	8.5 (6.8)
Concomitant NSAIDs, n (%)	124 (80.0)	127 (84.7)	40 (78.4)

Concomitant oral corticosteroids, n (%)	19 (12.3)	16 (10.7)	9 (17.6)
Concomitant DMARDs, n (%)	65 (41.9)	55 (36.7)	17 (33.3)
History of uveitis, n (%)	16 (10.3)	17 (11.3)	4 (7.8)
History of inflammatory bowel disease, n (%)	5 (3.2)	0 (0)	4 (7.8)*
History of psoriasis, n (%)	10 (6.5)	8 (5.3)	2 (3.9)
BASFI score, mean (SD)	60.6 (20.3)	57.7 (20.1)	59.7 (19.3)
BASDAI score, mean (SD)	62.4 (17.0)	59.4 (16.7)	61.1 (13.7)
VAS physician global assessment, mean (SD)	62.0 (14.8)	60.7 (15.4)	62.1 (16.2)
VAS patient global assessment, mean (SD)	68.1 (18.9)	65.41 (18.1)	66.2 (16.3)
Nocturnal back pain, mean (SD)	64.9 (21.0)	64.3 (21.7)	62.5 (22.4)
Total back pain, mean (SD)	63.9 (19.2)	63.5 (21.1)	63.1 (18.4)
CRP level (mg/l), mean (SD)	21.7 (24.6)	19.8 (20.8)	22.0 (22.9)
Modified Schober's test, mean (SD)	2.7 (1.9)	2.3 (2.1)	2.1 (2.6)
Chest expansion, mean (cm) (SD)	3.6 (2.5)	3.4 (1.7)	3.5 (1.7)
Occiput-to-wall measurement, mean (cm) (SD)	4.4 (5.9)	4.8 (5.5)	5.8 (6.2)

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; C reactive protein; DMARD, disease modifying anti-rheumatic drug, NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation; VAS, visual analogue scale.
*p<0.01, Fisher's exact test (two-tailed).

3.5.1.3 Results: study 314-EU/van der Heijde, 2006

Primary efficacy outcome

At Week 12, significantly more patients receiving etanercept 50 mg once weekly (74.2%) or 25 mg twice weekly (71.3%) achieved an ASAS 20 response compared with patients receiving placebo (37.3%; p<0.001). The lower limit of the CI for the actual response proportion difference between treatment groups was -7.1%, demonstrating non-inferiority of etanercept 50 mg once weekly to 25 mg twice weekly.

Secondary efficacy outcome

At Week 12, significantly greater ASAS 5/6 and ASAS 40 responses were achieved in patients receiving etanercept 50 mg once weekly and 25 mg twice weekly compared with patients in the placebo group (Table 21). There were no significant differences in the proportion of patients achieving ASAS response between the etanercept groups.

At Week 12, a significantly greater proportion of patients receiving etanercept 50 mg once weekly and 25 mg twice weekly, achieved partial remission compared with patients in the placebo group (Table 21). In the etanercept groups, onset of partial remission occurred significantly earlier than in the placebo group (p=0.003 and p=0.025, log rank test; for etanercept 50 mg once weekly and 25 mg twice weekly, respectively).

Table 21: Secondary efficacy outcomes: study 314-EU/van der Heijde, 2006

	Etanercept		

Outcome	50 mg (once weekly)	25 mg (twice weekly)	Placebo	p value
ASAS 5/6 (%)	70.3	72.0	27.5	p<0.001†
ASAS 40 (%)	58.1	53.3	21.6	p<0.0001†
Partial remission (as defined in Anderson 2001) (%)	31.6	21.3	5.9	p<0.05†

Abbreviations: ASAS, Assessment of Spondyloarthritis International Society.
†Etanercept 50 mg and etanercept 25 mg vs placebo

Significant improvements in measures of back pain, disease activity, and CRP levels were evident in both etanercept dosing regimen groups from as early as Week 2 compared with the placebo group (p<0.05). At all assessment points after Week 2 , a significantly greater proportion of patients receiving etanercept (both dosing regimens) achieved at least a 50% improvement in the BASDAI compared with the placebo group (p<0.05). At Week 12, 60.0% and 58.0% of patients in the etanercept 50 mg once weekly and 25 mg twice weekly groups, respectively, achieved a 50% improvement in the BASDAI compared with 19.6% patients receiving placebo (p<0.001).

[redacted] At Week 12, significant improvements in spinal mobility were observed in both etanercept groups compared with the placebo group when assessed using the modified Schober's test, chest expansion and occiput to wall measurements (p<0.001, p=0.024 and p<0.05, respectively).

Pharmacokinetics

There were no significant differences in serum etanercept exposure between the two etanercept dosing regimens.

3.5.1.4 Efficacy conclusion: Study 314-EU/van der Heijde, 2006

This randomised, placebo-controlled trial demonstrated that the efficacy of etanercept 50 mg once weekly was non-inferior to the dosing regimen of 25 mg twice weekly in patients with AS. In the pharmacokinetic analysis of both etanercept dosing regimens, there were no significant differences in serum etanercept exposure. Therefore, clinical outcomes for AS patients are expected to be comparable when treated with either the 50 mg once weekly or 25 mg twice weekly dose of etanercept.

3.5.2 Study 16.0037/Davis, 2003

3.5.2.1 Methodology: Study 16.0037/Davis, 2003

The source of the efficacy data for study 16.0037 is Davis et al, 2003 (10). Study 16.0037 was a Phase III, multicentre, randomised, double-blind, placebo-controlled 24 week trial to determine the safety and efficacy of etanercept in adults with moderate to severe AS. A summary of the methodology for study 16.0037/Davis, 2003 is provided in Table 22.

Table 22: Methodology study 16.0037/Davis, 2003

Study	16.0037/Davis, 2003
Objectives	The primary objective was to evaluate the clinical response to etanercept on the basis of the percentage of patients achieving ASAS 20 at Weeks 12 and 24.
Study location	The study was conducted at 28 sites in five countries: Canada, France, Germany, The Netherlands and the US.
Design details	Patients were stratified on the basis of concomitant use of DMARDs at the time of assignment to treatment. Efficacy and safety evaluations were performed at Weeks 2, 4, 8, 12 and 24.
Interventions	Patients were randomised in a 1:1 ratio, in blocks of two within each of the two strata, to receive either: • Etanercept 25 mg subcutaneous injection, self-administered twice weekly, or • Placebo 25 mg subcutaneous injection, self-administered twice weekly

Key inclusion criteria	• Aged 18–70 years
	• Diagnosis of AS according to modified New York criteria. Active AS was defined as:
	  ○ A score ≥30 mm for morning stiffness (average of two scores on a 100 mm VAS)
	  ○ Scores of ≥30 mm for two of the following: patient's global assessment of disease activity; average of two VAS scores assessing nocturnal or total back pain; functional ability as assessed by the BASFI

Key exclusion criteria	• Complete ankylosis (fusion) of the spine based on radiographic assessment • Previous use of TNF inhibitors • Serious infection (associated with hospitalisation or IV antibiotics) within 4 weeks prior to screening • Had received any DMARDs (excluding hydroxychloroquine, sulfasalazine or methotrexate) within 4 weeks of baseline
Permitted concomitant medications	• Doses of NSAIDS or prednisone (up to 10 mg/day) must have been stable for 2 weeks prior to baseline and must remain stable for the duration of the study
Efficacy outcomes	<u>Primary efficacy endpoint</u> • To compare the proportion of patients in both treatment arms achieving ASAS 20 at Weeks 12 and 24. <u>Secondary efficacy endpoints</u> • Achievement of ASAS 50 and ASAS 70 • Percentage of patients experiencing a significant decrease in disease activity and achievement of near or partial remission <u>Other outcomes</u> • Effects of treatment in individual ASAS domains, improvements in the BASDAI, spinal mobility scores (modified Schober's test, chest expansion score and occiput to wall distance), peripheral joint counts, CRP and ESR levels and assessors global assessments (measured on a 100 mm VAS) over time

Study	16.0037/Davis, 2003
Populations analysed	Assessment of efficacy and safety was based on the ITT population which included all patients who received at least one dose of the study drugs.
Statistical analysis	All statistical tests were two-sided. Population size was determined based on an earlier RCT in patients with AS (93) and power calculations were performed using a two-sided Fisher's exact test. The Cochran-Mantel-Haenszel test, stratified by the presence or absence of concomitant DMARDs, was used to compare the proportion of patients in each treatment arm achieving ASAS 20, ASAS 50, ASAS 70 and partial remission at each time point. The van Elteren stratified rank test was used for assessment of continuous measures calculated as the percentage improvement from baseline. Patients were considered non-responders at all assessment points for binary endpoints after study drug discontinuation. Missing data was handled using a last observation carried forward approach. For safety analyses, the proportion of patients with AEs and infections was compared between treatment arms using Fisher's exact test.

Abbreviations: AE, adverse events; AS, ankylosing spondylitis; BASDAI., Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; ESR, erythrocyte sedimentation rate; ITT, intention-to-treat; NSAID, non-steroidal anti-inflammatory drug; VAS, visual analogue scale.

3.5.2.2 Patient population: study 16.0037/Davis, 2003

Patient disposition

A total of 277 patients were enrolled in the study and were randomly assigned to receive 25 mg injections of etanercept (n=138) or placebo (n=139). In the etanercept group, 132 patients (96%) completed the first 12 weeks of the study while 126 patients (91%) completed 24 weeks of treatment. In the placebo group, 134 patients (96%) completed the first 12 weeks while 120 patients (86%) completed 24 weeks. Seven patients in the etanercept group discontinued treatment due to adverse events (AEs), three patients withdrew due to lack of efficacy and three patients were lost to follow up. In the placebo group, there was one discontinuation due to an AE, 13 withdrawals due to lack of efficacy, two withdrawals of consent, two discontinuations due to lack of adherence to the protocol and one patient was lost to follow up.

Patient demographics and baseline characteristics

For the majority of patients in the etanercept and placebo groups, demographics and baseline disease characteristics were well matched (Table 23).

Table 23: Patient demographics and baseline disease characteristics: study 16.0037/Davis, 2003

Characteristic	Etanercept (n=138)	Placebo (n= 139)
Age, mean (range) years	42.1 (24–70)	41.9 (18–65)
Male sex	105 (76)	105 (76)
White	130 (94)	127 (91)
Concomitant therapy at baseline		
Any DMARD†	44 (32)	43 (31)
Sulfasalazine	29 (21)	30 (22)
Methotrexate	15 (11)	17 (12)

Characteristic	Etanercept (n=138)	Placebo (n= 139)
Hydroxychloroquine	3 (2)	1 (1)
NSAIDs‡	126 (91)	128 (92)
Corticosteroids§	18 (13)	20 (14)
Duration of AS, mean (range) years	10.1 (0–30.7)	10.5 (0–35.3)
Extra-spinal disease manifestation, no. (%)		
History of Crohn's disease or ulcerative colitis	7 (5)	6 (4)
History of uveitis or iritis	39 (28)	43 (31)
History of psoriasis	11 (8)	15 (11)
HLA-B27 positive, no. (%)	108 (84)	109 (84)
ASAS component, mean (range) score		
Patient's global assessment (100-mm VAS)	62.9 (16–100)	62.9 (9–100)
Total back pain (100-mm VAS)	61.1 (7–100)	63.5 (0–99)
BASFI§	51.7 (4.3–97.7)	56.3 (11.5–97.0)
Inflammation (0–100 scale) ¶	61.4 (17–100)	64.3 (7–100)
Spinal mobility measure, mean ± SEM		
Modified Schober's test, cm	3.06 ± 0.15	2.97 ± 0.13
Chest expansion, cm	3.26 ± 0.19	3.21 ± 0.15
Occiput-to-wall, cm	5.59 ± 0.50	5.33 ± 0.56
BASDAI score, mean ± SEM††	58.1 ± 1.5	59.6 ± 1.4
CRP level, mg/dl, mean ± SEM (normal 0–1.0 mg/dl)	1.9 ± 0.2	2.0 ± 0.2
ESR, mm/hour‡‡, mean ± SEM	25.9 ± 1.8	25.4 ± 1.9

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; ESR, erythrocyte sedimentation rate; HLA, human leucocyte antigen; NSAID, non-steroidal anti-inflammatory drug; SEM, standard error of the mean; VAS, visual analogue scale.
†Some patients were receiving >1 DMARD. ‡Received within 6 months of the screening assessment. §Average of scores for 10 questions (possible range 0–100). ¶Mean score for the last 2 questions on the 6-quesiton BASDAI. ††Mean score for 6 questions related to disease activity. ‡‡Normal: 1–17 (males), 1–25 mm/hour (females).

3.5.2.3 Results: study 16.0037/Davis, 2003

Primary efficacy outcome

At Week 12, 82 (59%) of patients in the etanercept group and 39 (28%) of patients in the placebo group achieved an ASAS 20 response (p<0.0001). At Week 24, 78 patients (57%) and 31 patients (22%) achieved ASAS 20 in the etanercept and placebo groups, respectively (p<0.0001). A statistically significant difference in the ASAS responses between the two groups was evident from Week 2 and was sustained over the 24 weeks of treatment.

Secondary efficacy outcomes

Significant improvements in ASAS 50 and ASAS 70 responses were evident at Week 2 and were sustained throughout the study.

[redacted]

[redacted]

Other efficacy outcomes

[redacted] Etanercept patients had significantly greater improvements in all individual components of the ASAS response criteria, CRP and ESR levels, BASDAI, BASFI and spinal mobility measures, compared with placebo patients at 24 weeks (Table 24).

Table 24: Summary of secondary and other efficacy outcomes (study 16.0037/Davis, 2003)

	Etanercept (n=138)		Placebo (n=139)		
Outcome	**Baseline†**	**24 weeks†**	**Baseline†**	**24 weeks†**	**p value**
ASAS component					
Patient's global assessment (100-mm VAS)	62.9 ± 1.5	36.0 ± 2.4	62.9 ± 1.6	56.3 ± 2.3	<0.0001
Back pain (100-mm VAS)	61.1 ± 1.8	37.1 ± 2.4	63.5 ± 1.7	57.8 ± 2.3	<0.0001
BASFI‡	51.7 ± 1.8	36.0 ± 2.2	56.3 ±1.7	54.7 ± 2.2	<0.0001
Inflammation§	61.4 ± 1.9	33.4 ± 2.3	64.3 ± 1.8	56.6 ± 2.1	<0.0001
CRP level, mg/dl (normal 0–1.0 mg/dl)	1.9 ± 0.2	0.6 ± 0.1	2.0 ± 0.2	1.9 ± 0.2	<0.0001
ESR, mm/hour¶	25.9 ± 1.8	11.2 ± 1.0	25.4 ± 1.9	25.9 ± 1.9	<0.0001
Spinal mobility measure††					
Modified Schober's test, cm	3.06 ± 0.15	3.34 ± 0.15	2.97 ± 0.13	2.88 ± 0.13	<0.0014
Chest expansion, cm	3.26 ± 0.19	3.85 ± 0.20	3.21 ± 0.15	3.01 ± 0.16	<0.0001
Occiput-to-wall, cm	5.59 ± 0.50	4.53 ± 0.50	5.33 ± 0.56	6.01 ± 0.68	<0.0001
BASDAI‡‡	58.1 ± 1.5	34.5 ± 2.1	59.6 ± 1.4	55.1 ± 2.0	<0.0001

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; VAS, visual analogue scale.
†Values are reported as the mean ±SEM.
‡Mean scores for 10 questions (possible range 0–100).
§Mean score for the last 2 questions on the 6-question BASDAI.
¶Normal: 1–17 mm/hour for males, 1–25 mm/hour for females.
††Represented by increases in the modified Schober's test and chest expansion measurements and by a decrease in the occiput-to-wall measurement.
‡‡Mean scores for 6 questions related to disease activity.

Subgroup analyses

Despite previous treatment with NSAIDs or corticosteroids and regardless of concomitant DMARD therapy, etanercept demonstrated efficacy in subgroup analyses.

[redacted].

3.5.2.4 Efficacy conclusion: study 16.0037/Davis, 2003

This randomised, placebo-controlled trial demonstrates that etanercept 25 mg, self-administered twice weekly is associated with rapid and significant improvements in the ASAS 20 and in higher levels of the ASAS response (ASAS 50 and ASAS 70) in patients diagnosed with moderate to severe active AS according to the modified New York criteria. From as early as Week 2, patients receiving etanercept also demonstrated rapid and significant improvements on the BASDAI and BASFI compared with patients receiving placebo. Etanercept also demonstrated efficacy in patients with persistent disease, despite previous treatments with NSAIDs and corticosteroids and regardless of concomitant DMARD therapy.

3.5.3 Study 16.0040/Davis, 2005

3.5.3.1 Methodology: study 16.0040/Davis, 2005

The source of the efficacy data for study 16.0040 is Davis et al, 2005 (119). This was an open-label extension study to:

- Evaluate the extended safety of etanercept in the treatment of patients with AS who had previously enrolled in study 016.0037 (Davis, 2003).
- To establish the efficacy of etanercept in patients with AS who had previously received placebo in study 016.0037.

A summary of the methodology for study 16.0040/Davis, 2005 is provided in Table 25.

Table 25: Methodology study 16.0040/Davis, 2005

Study	16.0040/Davis, 2005
Objectives	To evaluate the longer term safety and durability of the clinical response in patients receiving etanercept for treatment of AS.
Study location	The study was conducted at 28 sites in Canada, France, Germany, The Netherlands and the US.
Design details	Patients who had previously received etanercept in study 16.0037 and had enrolled in the extension study continued to receive etanercept for a further 72 weeks (up to 96 weeks of treatment). Patients who had previously received placebo in study 16.0037 switched to etanercept therapy in the open-label extension study for up to 72 weeks.
Interventions	Patients received 25 mg etanercept subcutaneous injection, self-administered twice weekly for up to 72 weeks.
Key inclusion criteria	• Detailed inclusion criteria are provided in Table 22 • All patients who had completed study 16.0037 were eligible • Patients who had discontinued study 16.0037 due to lack of efficacy were eligible provided that they had returned for all the required follow up visits
Key exclusion criteria	• Detailed exclusion criteria are provided in Table 22
Efficacy outcomes	<u>Primary efficacy endpoint</u> • The durability of response, as assessed by the proportion of ASAS 20 responders at Week 72, in those patients continuing etanercept from study 16.0037 <u>Secondary efficacy endpoints</u> • Achievement of ASAS 50 or ASAS 70 improvement • The mean of the following four components of the ASAS outcome: patient global assessment, pain score, BASFI, BASDAI (two morning stiffness-related scores) <u>Other outcomes</u> • Nocturnal and total back pain, spinal mobility score (modified Schober's test, chest expansion and occiput to wall measurements), peripheral joint counts, CRP and ESR levels, assessors global assessment and the proportion of patients achieving ASAS 40 or ASAS 5/6
Populations analysed	All patients who received at least one dose of study drug were included in the safety analysis.
Statistical analysis	The proportion of ASAS responders was determined over time by original treatment group. Efficacy analyses were reported without statistical inference between treatment groups. Measurements from the baseline visit of study 16.0037 were used as baseline values for all patients. The durability of the ASAS response rates

Study	16.0040/Davis, 2005
	was assessed by combining data with the results of study 16.0037.

Abbreviations: AS, ankylosing spondylitis; ASAS Assessment of Spondyloarthritis International Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; ESR, erythrocyte sedimentation rate.

3.5.3.2 Patient population: study 16.0040/Davis, 2005

Patient disposition

Of the 277 patients who received treatment in study 16.0037, a total of 257 patients were enrolled in the extension study (128 patients from the etanercept arm of the previous RCT and 129 patients beginning etanercept at Week 0). Of the 12 etanercept patients who discontinued study 16.0037, three patients enrolled in the extension study. Of the 19 placebo patients who discontinued study 16.0037 due to lack of efficacy, eight patients enrolled in the open-label extension study. Two hundred and nine patients received etanercept for at least 72 weeks (104 patients continuing from study 16.0037) and 95 patients received 96 weeks of treatment. Sixty seven patients discontinued the extension study.

Patient demographics and baseline characteristics

The baseline demographics of the study patients were similar to those in study 16.0037 (Table 23). The patient population was predominantly male (75%) and white (93%). The mean disease duration of the study patients was 10.7 years (range 0.5–35.8). Approximately one third of patients were taking DMARDs, including SSZ (21.8%), methotrexate (11.3%) and hydroxychloroquine (1.6%). The mean (SD) patient global assessment and BASFI at baseline were 63.3 (17.6) and 54.0 (20.6), respectively.

3.5.3.3 Results: study 16.0040/Davis, 2005

Primary efficacy outcome

Those patients who had previously received etanercept in study 16.0037 continued to show improvement in ASAS response after an additional 72 weeks of etanercept treatment (74% of patients achieved ASAS 20 while 46% of these patients achieved an ASAS 70 response). Of those patients who had previously received placebo in study 16.0037, at Week 24, 70% achieved an ASAS 20 response.

Secondary efficacy outcomes

Improvements in spinal mobility were seen in both those patients who had previously received etanercept and in those who had received placebo in study 16.0037 (Table 26). Those patients who had received etanercept in study 16.0037 showed sustained benefit in all additional outcome criteria (Table 26). The patients who had received placebo in study 16.0037 showed improvement in all measures of additional outcome criteria, including CRP levels, duration and intensity of morning stiffness (mean of the last two questions on the BASDAI) and the BASFI (Table 26). Disease improvement as measured by the ASAS 40 and ASAS 5/6 response was observed in both patient groups.

Table 26: Additional efficacy outcomes following extended etanercept treatment (study 16.0040/Davis, 2005)

	Etanercept		Placebo	
Other outcomes	**Baseline (study**	**Extension study**	**Baseline (study**	**Etanercept Week 72‡**

	16.0037) (n=128)	Week 96† (n=95)	16.0037) (n=129)	(n=105)
Spinal mobility measure, mean (SE)				
Modified Schober's test, cm	3.1 (0.2)	3.5 (0.2)	3.0 (0.1)	4.6 (0.7)`
Chest expansion, cm	3.4 (0.2)	4.0 (0.2)	3.2 (0.2)	3.5 (0.2)`
Occiput-to-wall, cm	5.4 (0.5)	4.0 (0.6)	5.0 (0.6)	4.6 (0.7)`
ASAS component, mean, (SE)				
Patient's global assessment (0–100)	63.2 (1.5)	26.6 (2.6)	63.3 (1.6)	28.4 (2.3)
Back pain (0–100)	61.7 (1.8)	24.3 (2.6)	63.7 (1.8)	27.5 (2.6)
BASFI (0–100)	51.1 (1.8)	25.4 (2.4)	56.8 (1.8)	32.3 (2.5)
Inflammation§ (0–100)	60.9 (1.9)	22.8 (2.6)	65.3 (1.8)	28.0 (2.4)
CRP levels, mean (SE)	1.8 (0.2)	0.5 (0.1)	2.0 (0.2)	0.7 (0.1)

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; SE, standard error.
†Consisting of 72 weeks of open-label etanercept and 24 weeks of etanercept in study 16.0037.
‡Consisting of 72 weeks of open-label etanercept.
§Mean score for the last 2 questions of the BASDAI (duration and intensity of morning stiffness).

3.5.3.4 Efficacy conclusion: study 16.0040/Davis, 2005

As demonstrated by the results of the ASAS 20, ASAS 50 and ASAS 70, responses to treatment were maintained or improved in patients who received etanercept during the open-label period, regardless of prior treatment group in study 016.0037. A sustained response was also demonstrated in the BASFI, measures of spinal mobility and in measures of inflammation (mean of the last two questions on the BASDAI). The results of this open-label extension study demonstrated that etanercept 25 mg twice weekly provided sustained durability of response in the improvement of the signs and symptoms of AS for 96 weeks.

3.5.4 Study: Davis, 2008

3.5.4.1 Methodology and efficacy outcomes: Davis, 2008

The source of the efficacy data for Davis, 2008 is Davis et al (118). This was the second open-label extension study of study 16.0037/Davis, 2003 to assess the long-term efficacy and safety of etanercept in patients with AS. A summary of the methodology for Davis, 2008 is presented in Table 27.

Table 27: Methodology: Davis, 2008

Study	Davis, 2008
Objectives	To evaluate the long-term safety and efficacy of etanercept in patients with AS receiving therapy for up to 192 weeks.
Study location	The study was conducted at 28 sites in Canada, France, Germany, The Netherlands and the US.
Design details	Study 16.0040, the 96 week, open-label extension trial of study 16.0037 was further extended for a total of 168 weeks. Patients who had previously received etanercept in the initial RCT could receive etanercept for up to 192 weeks. Patients who received placebo in the RCT could potentially receive etanercept for up to 168 weeks.
Interventions	Following Week 72 of study 16.0040, patients received etanercept 50 mg per week, administered as two 25 mg subcutaneous injections twice weekly.
Key inclusion criteria	• Full details are provided in section 3.5.2.1, Table 22 • Discontinuation of DMARDs prior to enrolment
Key exclusion criteria	• Full details are provided in section 3.5.2.1, Table 22 • Previous therapy with any other anti-TNF agent
Permitted concomitant medications	• Doses of methotrexate, hydroxychloroquine or SSZ must be stable if concomitant use was deemed necessary
Efficacy outcomes	• Achievement of ASAS 20, ASAS 40, ASAS 5/6 and partial remission • Effects of etanercept in individual ASAS domains (pain, function, inflammation and patient global assessment) • BASFI and BASDAI scores • Spinal mobility assessed by the modified Schober's test, chest expansion and occiput-to wall measurement
Populations analysed	All patients who had at least one post-dose assessment were included in the analysis.
Statistical analysis	The LOCF approach for missing data was used to calculate ASAS 20 response, ASAS 40 response and partial remission in patients who had at least one post-dose assessment. Efficacy analyses were reported without statistical inference between treatment groups. Baseline values were determined for the initial RCT. Data relating to all individual components of the efficacy variables were collected at 12- or 24-week intervals for 168 weeks with the exception of chest expansion, which was not recorded for all patients at the final time points, thereby limiting ASAS 5/6 analysis to 144 weeks.

Abbreviations: AE, adverse event; AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis International Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; DMARD, disease modifying anti-rheumatic drug; LOCF, last observation carried forward; RCT, randomised controlled trial; SSZ, sulfasalazine.

3.5.4.2 Patient population: Davis, 2008

Patient disposition

A total of 257 patients were enrolled in the open-label extension; 128 from the etanercept group and 129 from the placebo group. A total of 126 patients completed the additional extension study while 131 patients discontinued. Reasons for discontinuation included: patients declined to re-enrol following the amendments to lengthen the study duration (13.2%), patient refusal (10.1%) and AEs (8.2%).

Patient demographics and baseline characteristics

The patient population was predominantly male (75%) and Caucasian (93%). The mean disease duration of the study patients was 10.7 years. Approximately one third of patients were taking DMARDs at baseline, including SSZ (21.8%), methotrexate (11.3%) and hydroxychloroquine (1.6%).

3.5.4.3 Results: Davis, 2008

After 192 weeks of continuous etanercept treatment, 81% of patients achieved a sustained ASAS 20 response while 69% achieved and sustained ASAS 40. By Week 168, a sustained ASAS 20 response was achieved in 82% of patients switching from placebo while ASAS 40 was achieved and sustained in 68% of patients. Patients continuing with etanercept showed durability of response in the ASAS 5/6; 49% of patients responded by Week 8 and 60% responded by Week 144. After 24 weeks of etanercept, 48% of patients switching from placebo achieved ASAS 5/6 while 40% reached ASAS 5/6 at Week 144. At Week 192, partial remission was achieved by 44% of patients continuing with etanercept and 28% of patients switching from placebo.

Improvements in disease activity and function as measured by the BASDAI and the BASFI were seen from baseline in the open-label extension study and were sustained throughout the study. Patients in both arms demonstrated sustained improvements in patient global assessment, total back pain and CRP levels. Results of the analyses of the individual composite efficacy outcomes are shown in Table 28.

Table 28: Disease response measures at RCT baseline, OLE baseline and Week 168, Davis, 2008

Parameter (mean, SD)	ETN/ETN			Placebo/ETN		
	RCT baseline (n=128)	OLE baseline (n=124)	Week 168 (n=54)	RCT baseline (n=129)	OLE baseline (n=127)	Week 168 (n=50)
Patient global assessment†	63.2 (17.2)	29.6 (25.6)	19.7 (21.5)	63.3 (18.1)	33.0 (26.2)	25.9 (23.2)
Total back pain†	61.7 (20.7)	29.5 (27.4)	18.8 (21.6)	63.7 (20.2)	33.6 (28.1)	24.1 (23.5)
BASDAI morning stiffness‡	60.9 (21.6)	26.6 (26.0)	19.0 (22.1)	65.3 (20.2)	30.4 (24.1)	26.0 (23.3)
CRP level (mg/dl)	1.8 (1.9)	0.7 (1.8)	0.3 (0.5)	2.0 (2.4)	0.6 (0.8)	0.5 (0.7)
Chest expansion (cm)§	3.4 (2.2)	3.9 (1.8)	4.1 (1.9)¶	3.2 (1.8)	3.3 (2.0)	3.7 (2.4)¶
Modified Schober's test (cm)§	3.1 (1.8)	3.3 (1.7)	3.5 (1.7)	3.0 (1.6)	3.0 (1.5)	3.0 (1.5)
Occiput-to-wall (cm)§	5.4 (5.7)	4.2 (6.2)	3.6 (5.9)	5.0 (6.4)	4.8 (7.0)	5.4 (7.3)

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C reactive protein; ETN, etanercept; OLE, open-label extension; RCT, randomised controlled trial; SD, standard deviation.
†Scored 0–100 mm on a VAS

‡Average of the last 2 questions on the BASDAI relating to morning stiffness
§Data collected at 24 week rather than 12 week intervals. At the OLE baseline the number of patients in the placebo-/ETN group was n=122 and for the ETN/ETN group n=113
¶LOCF from Week 144

3.5.4.4 Efficacy conclusion: Davis, 2008

As demonstrated by the results of the ASAS 20 and the ASAS 40, treatment responses to etanercept were sustained throughout the open-label period in both treatment arms. A rapid response rate was also observed in the ASAS 5/6 and was sustained throughout the study period in both treatment groups. Substantial improvements in measures of disease activity were also maintained throughout the course of the open-label period. The results from this 168 week open-label study therefore demonstrate that long-term therapy with etanercept in patients with AS results in major clinical improvements in the signs and symptoms of AS which are sustained for up to 192 weeks.

3.5.5 Study: Boonen, 2008

3.5.5.1 Methodology and efficacy outcomes: Boonen, 2008

The source of the data is Boonen et al, (22). This was an observational follow-up study of studies 16.0037 and 16.0040 and was designed to investigate the longer term effects of etanercept on the health related quality of life (HRQoL) in patients with AS. A summary of the methodology and efficacy outcomes is provided in Table 29.

Table 29: Methodology and efficacy outcomes: Boonen, 2008

Study	Boonen, 2008
Objectives	To examine the long-term effect of etanercept therapy on HRQoL and utility in patients with AS.
Study location	The study was conducted at 28 sites in Canada, France, Germany, The Netherlands and the US.
Design details	Patients who had previously received etanercept in study 16.0037 and had enrolled in the 16.0040 extension study continued to receive etanercept for a further 72 weeks (up to 96 weeks of treatment). Patients who had previously received placebo in study 16.0037 switched to ETN therapy in the open-label extension study for up to 72 weeks. Patient-reported outcomes were used to assess HRQoL and utility at open-label baseline and every 12 weeks thereafter.
Interventions	Patients received 25 mg etanercept subcutaneous injection, self-administered twice weekly for up to 72 weeks.
Key inclusion criteria	• Detailed inclusion criteria are provided in Table 22 • All patients who had completed study 16.0037 were eligible provided that data was available at baseline and Week 12 for at least one patient-reported outcome • Patients who had discontinued study 16.0037 due to lack of efficacy were eligible provided that they had returned for all the required follow up visits
Key exclusion criteria	• Detailed exclusion criteria are provided in Table 22
Permitted concomitant medications	• Permitted and concomitant medications are detailed in Table 22
Efficacy outcomes	• Mean change from baseline to Week 12 in SF-36, EQ-5D, SF-6D, and EQ-VAS scores and to Week 72 • Differences between mean SF-6D and EQ-5D scores in each treatment arm at baseline and between mean changes in SF-6D and EQ-5D scores from baseline to Week 12 • Mean 72-week QALY gain per person in patients switching from placebo
Populations analysed	The population of patients included in the analyses were those who initiated the open-label phase of the trial and for whom data was available at baseline and at Week 12 for at least one patient-reported outcome. In total this included 232 patients; 119 patients continuing with etanercept and 113 patients switching from placebo to etanercept in the open label study.

Study	Boonen, 2008
Statistical analysis	Descriptive statistics were used to summarise SF-36, EQ-5D and SF-6D scores at baseline and each subsequent time point. Comparison of these scores was conducted using the t test. Baseline values and mean change from baseline were calculated for EQ-5D and SF-6D utilities and compared using the t test. QALYs over the 72 weeks were calculated based on EQ-5D and SF-6D utilities. It was assumed that the utility score would not have changed if patients had not received active treatment. Patients who withdrew from the trial were assigned their mean baseline utility score before starting active treatment.
Patient disposition	Of the 257 patients who initiated the open-label extension phase of the trial (128 patients continuing with etanercept and 129 patients switching from placebo), a total of 220 (85%) completed the 72-week open-label extension study, including 105 patients in the etanercept continuation arm and 115 patients switching from placebo. Data at baseline and Week 12 were available for at least one patient-reported outcome for 232 (90%) patients (113 etanercept patients and 119 patients switching from placebo, respectively). Of these patients, all had at least two assessments of the EQ-5D, while 227 patients (88%) provided two assessments of the SF-36 and SF-6D.
Patient demographics and baseline characteristics	• The two study groups were similar in terms of patient baseline demographics, including mean age, height, weight, and duration of disease. The majority of subjects were Caucasian men, with similar proportions in each treatment arm. • For patient-reported outcomes, open-label mean baseline values were significantly higher in patients continuing with etanercept compared with patients switching from placebo for the SF-36 PCS (p<0.0001), the EQ-5D (p<0.0001), and the SF-6D (p=0.0007) scores. Mean baseline values were also higher in patients continuing etanercept for SF-36 MCS and EQ VAS scores, although these differences were not significant.
Efficacy outcomes	• Despite significantly lower SF-36 PCS, EQ-5D, SF-6D and EQ-VAS scores in patients switching from placebo to etanercept at baseline, by Week 12, scores were similar between treatment groups and were sustained in both treatment arms over the 72-week treatment period. • At open-label baseline, mean SF-6D scores were significantly higher than EQ-5D scores in the patients switching from placebo (0.65 vs 0.49, respectively; p<0.0001) but not in patients continuing with etanercept (0.71 vs 0.69, respectively; p=0.4) • The mean change in utility scores from baseline to Week 12 was significantly lower for the SF-6D than for the EQ-5D in patients switching from placebo (0.06 vs 0.18, respectively; p<0.0001) • The average QALY gain in patients switching from placebo was 0.24 calculated using the EQ-5D, and 0.10 using the SF-6D
Efficacy conclusion	Etanercept treatment in patients with AS is associated with rapid and sustained improvements in measures of HRQoL. Etanercept treatment produces sustained improvement in SF-36 as well as EQ-5D and SF-6D utility scores over a 72-week period. Despite lower scores at baseline, patients switching from placebo to etanercept showed a rapid improvement in HRQoL and utility and, by Week 12, showed similar scores to patients continuing with etanercept. These improvements were sustained over the 72-week treatment period.

Abbreviations: AS, ankylosing spondylitis; EQ-5D, EuroQoL-5D; EQ-VAS, EuroQoL visual analogue scale; HRQoL, health-related quality of life; MCS, mental component score, PCS, physical component score; QALY, quality-adjusted life year; SF-36, short form-36.

3.5.6 Study: van der Heijde, 2008

3.5.6.1 Methodology and efficacy outcomes: van der Heijde, 2008

The source of the efficacy data is van der Heijde et al (124). This was an observational study to assess the effect of etanercept treatment on radiographic progression in patients with AS versus a cohort of patients who had never received treatment with anti-TNF agents. The methodology and efficacy outcomes of van der Heijde, 2008 are presented in Table 30.

Table 30: Methodology and efficacy outcomes, van der Heijde, 2008

Study	Van der Heijde, 2008
Objectives	To investigate the effect of etanercept therapy on radiographic progression in patients with AS.
Study location	The study was conducted at 28 sites in Canada, France, Germany, The Netherlands and the US. Patients in the control arm of the study were enrolled in the OASIS study which was conducted in Belgium, France and The Netherlands.
Design details	Patients in the etanercept arm of this controlled study had previously completed the 24 week double-blind study 16.0037 and were enrolled in a 72 week open label extension trial (study 16.0040). Patients received etanercept for up to 96 weeks. Patients in the control arm of the present study were enrolled in the OASIS study† and had not been treated with anti-TNF agents. Radiographs of the cervical and lumbar spine from etanercept patients were compared with radiographs of patients in the control arm. Radiographs from both study arms were assessed by two independent, blinded readers.
Interventions	Patients in the etanercept arm received etanercept 25 mg subcutaneously, twice weekly for up to 96 weeks.
Key inclusion criteria	• All patients from study 16.0037 and the 16.0040 extension study who had received at least 1 dose of etanercept, and had baseline radiographs of the cervical and lumbar spine were eligible • All patients from the OASIS study with radiographs at baseline and at 2 years were eligible
Key exclusion criteria	• Patients from the OASIS study with complete spinal fusion at baseline were not eligible
Permitted concomitant medications	• NSAIDs • Analgesics • DMARDs (including corticosteroids).
Efficacy outcomes	<u>Primary efficacy endpoint:</u> • Direct comparison of the 96-week change from baseline in the mSASSS between patients receiving etanercept and those in the control group, adjusted for baseline mSASSS. <u>Secondary endpoints and sensitivity analysis:</u> • Change in mean baseline cervical and lumbar radiography scores • Comparison between all patients with baseline radiographs in the etanercept group and patients in the OASIS group who would have fulfilled the entry criteria for study 16.0037 (section 3.5.2) • Comparison between all patients in the etanercept group with baseline and 96 week radiographs and patients from the OASIS study who would have fulfilled the entry criteria for study 16.0037 (section 3.5.2)
Populations	All patients from the RCT and the open-label extension who received at least 1 dose of etanercept and had baseline radiographs of the cervical and lumbar spine

Study	Van der Heijde, 2008
analysed	were analysed. All patients from the OASIS study with radiographs at baseline and at 2 years were evaluated as controls.
Statistical analysis	Missing data on the 96-week change from baseline in the mSASSS due to missing post-baseline radiographs was calculated using the median change from baseline score observed in patients in the control arm with the same baseline score. Baseline characteristics of the patients in the etanercept trial and patients in the OASIS study were compared by chi-square test for categorical variables and t-test for continuous variables. Radiographic progression was compared between the patients in the etanercept trial and those in the OASIS study using Quade rank analysis of covariance, adjusted for baseline mSASSS on the change in mSASSS from baseline to 96 weeks.

Abbreviations: AS, ankylosing spondylitis; DMARDs, disease-modifying anti-rheumatic drugs; NSAIDs, non-steroidal anti-inflammatory drugs; mSASSS, modified Stoke AS Spine Score; OASIS, Outcome Assessment in Ankylosing Spondylitis International Study.
†The OASIS study was an international observational study on outcome in AS patients which included consecutive patients who were followed up for 10 years. The protocol included assessment of radiographs of the cervical and lumbar spine at baseline at 1-year and 2-year follow-up, and every 2 years thereafter.

3.5.6.2 Patient population: van der Heijde, 2008

Patient disposition

Of the 257 patients who enrolled in the open label extension study, 76% received etanercept for at least 48 weeks while 50% received etanercept for at least 72 weeks and had data available for radiographic analysis. A total of 175 of the 219 original OASIS patients had available radiographs for analysis. Forty-three percent of OASIS patients (n=76) met the inclusion criteria for the original RCT at baseline (section 3.5.2).

Patient demographics and baseline characteristics

Patient age and duration of disease were similar between treatment groups. Patients in the OASIS arm has significantly less active disease as indicated by lower BASDAI, BASFI, and patients global assessment scores (all p<0.0001). Baseline spinal mobility and mSASSS scores were similar between treatment groups (Table 31).

Table 31: Patient demographics and baseline characteristics

Baseline characteristic	Etanercept (n=257)	OASIS (n=175)	OASIS meeting study 16.0037 entry criteria (n=76)
Age, mean ±SD years	41 ± 10.2	44 ± 12.5	48 ± 12.3
Male, no. (%)	194 (75.5)	121 (69.1)	54 (71.1)
Disease duration, mean ±SD years	10 ± 8.5	11 ± 8.5	12 ± 9.8
CRP, mean ± SD mg/dL	2.0 ± 2.20	1.5 ± 1.98	1.50 ± 1.81
HLA-B27 positive, no. (%)	201 (78.2)	142 (81.1)	64 (84.2)
DMARD use at baseline, no. (%)	83 (32.3)	18 (10.3)*	4 (5.3)
Patients global assessment, mean ±SD	63 ± 18.0	38 ± 27.8*	27 ± 27.2
BASFI, mean ±SD	54 ± 20.7	34 ± 25.5*	55 ± 16.6
BASDAI, mean ± SD	63 ± 20.9	35 ± 20.9*	47 ± 19.8

Modified Schober's score, mean ± SD	3.0 ± 1.7	2.9 ± 1.4	2.3 ± 1.4
mSASSS, mean ± SD	16 ± 18.3	14 ± 17.6	19 ± 20.8

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; HLA, human leucocyte antigen; mSASSS, Modified Stoke Ankylosing Spondylitis Spine Score; SD, standard deviation.
*p<0.0001 versus etanercept treated patients.

3.5.6.3 Results: van der Heijde, 2008

Primary efficacy outcome

There was no significant difference in the change in the mean mSASSS score from baseline among patients who received etanercept (0.91 ± 2.45) versus those from the OASIS group (0.95 ± 3.18), or versus OASIS patients who met RCT entry criteria at baseline (1.27 ± 3.64).

Secondary efficacy outcomes

At 96 weeks, there were no significant differences in the change in mean cervical radiography scores from baseline between etanercept patients (0.49 ± 1.40) versus OASIS patients (0.42 ± 2.11), or versus OASIS patients who met RCT entry criteria at baseline (0.53 ± 2.29). At 96 weeks, there were no significant differences in change in mean lumbar radiography scores from baseline between etanercept patients (0.42 ± 1.84) versus OASIS patients (0.53 ± 1.88), or versus OASIS patients who met RCT entry criteria at baseline (0.73 ± 2.00).

In multiple subgroup analyses, there were no significant differences in the mean change in the mSASSS between treatment groups when patients receiving etanercept were subgrouped according to duration of treatment. There were no significant differences in the mean change in mSASSS when patients who had received etanercept during study 16.0037 and patients who had previously received placebo were compared with OASIS patients.

3.5.6.4 Study conclusion: van der Heijde, 2008

While etanercept has been shown to be highly effective in treating the clinical signs and symptoms of AS, this study did not find any evidence that structural changes in spine disease were slowed or reversed in AS patients treated with etanercept for nearly 2 years.

3.5.7 Study 311-EU/Calin, 2004

3.5.7.1 Methodology: study 311-EU/Calin, 2004

The source of the efficacy data for study 311-EU is Calin et al, 2004 (11) (NCT00421915). This was a Phase III, multicentre, randomised, double-blind, placebo-controlled trial to evaluate the efficacy and safety of etanercept in the treatment of adult patients with active AS. A summary of the methodology for study 311-EU/Calin, 2004 is provided in Table 32.

Table 32: Methodology study 311-EU/Calin, 2004

Study	311-EU/Calin, 2004 (NCT00421915)
Objectives	The primary objective was to evaluate the clinical response to etanercept on the basis of response criteria recommended by the ASAS Working Group.
Study location	The study was conducted in 14 centres in eight countries: Belgium, Finland, France, Germany, Italy, The Netherlands, Spain and the UK.
Design details	A screening period of up to 4 weeks was followed by a 12 week treatment period during which patients received etanercept or placebo. Efficacy and safety evaluations were performed at Weeks 2, 4, 8 and 12. Patients were stratified on the basis of concomitant use of DMARDs at baseline. Patients were randomised to receive either etanercept or placebo. Etanercept and placebo study drugs were similar in appearance to maintain blinding.
Interventions	Patients were randomised to receive either: • Etanercept 25 mg subcutaneous injection, self-administered twice weekly, or • Placebo 25 mg subcutaneous injection, self-administered twice weekly
Key inclusion criteria	• Aged 18–70 years • Active AS diagnosed using the modified New York criteria. Active disease was diagnosed if the patient had an average score ≥30 for spinal inflammation and a score ≥30 on at least two out of the remaining three domains.
Key exclusion criteria	• Complete ankylosis of the spine • Previous use of TNF inhibitors (including etanercept) • Use of DMARDs (except for hydroxychloroquine, sulfasalazine or methotrexate) within 4 weeks of baseline • Use of multiple NSAIDs • Receiving >10 mg prednisone daily • Had changed doses of NSAIDs or prednisone within 2 weeks of baseline
Efficacy outcomes	<u>Primary efficacy endpoint</u> • The proportion of patients achieving an ASAS 20 response <u>Secondary efficacy endpoints</u> • Percentage of ASAS 20 responders at Weeks 2, 4 and 8 • Achievement of ASAS 50 or ASAS 70 at any visit • Symptom improvements in individual ASAS domains • Improvement on the BASDAI • Improvements in levels of CRP and ESR • Spinal mobility improvement (chest expansion, Schober's test and occiput to wall distance)
Populations analysed	Assessment of efficacy and safety was based on the ITT population which included all patients who received at least one dose of the blinded study drugs.
Statistical analysis	Differences between treatment arms were assessed using the Cochrane-Mantel-Haenszel test, stratified by baseline DMARD use. Direct and interactive effects of DMARDs on ASAS responses were assessed using Chi-square and Breslow-Day

Study	311-EU/Calin, 2004 (NCT00421915)
	tests. For the safety analysis, the percentage of adverse events occurring in both treatment arms was assessed using Fisher's exact test. All statistical tests were two-sided. Response rates of 75% and 35% in the etanercept and placebo groups, respectively, were expected based on a previous trial. Assuming similar response rates, the study design provided 90% power with 40 patients in each group. Missing data for continuous and ordinal end points was handled using the LOCF approach. Non-responders were patients who withdrew from the study prematurely at each assessment interval for ASAS and other patient reported responses.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drugs; ESR, erythrocyte sedimentation rate; ITT, intention to treat; LOCF, last observation carried forward; NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor.

3.5.7.2 Patient population: study 311-EU/Calin, 2004

Patient disposition

A total of 84 patients were enrolled in the study and were randomly assigned to receive 25 mg injections of etanercept (n=45) or placebo (n=39). Eighty two patients completed the study and two patients discontinued prematurely; both in the etanercept arm. The first patient discontinued due to not meeting the inclusion criterion for active disease while the second withdrew consent 8 days after treatment was started.

Patient demographics and baseline characteristics

Patient demographics were similar between treatment groups except that patients in the etanercept group were, on average, 5 years older and had a diagnosis of AS for 5 years longer than patients in the placebo group. The majority of patients were white (94%) and there was a higher proportion of males (79%). Baseline disease activity scores and concomitant use of DMARDs, NSAIDS and corticosteroids was similar in both treatment arms. A summary of patient demographics and baseline characteristics is provided in Table 33.

Table 33: Patient demographics and baseline characteristics: study 311-EU/Calin, 2004

Characteristic	Etanercept (n=45)	Placebo (n=39)	Total
Age (years), mean (SD)	45.3 (9.5)*	40.7 (11.4)	43.2 (10.6)
Sex (male/female), %	80/20	77/23	79/21
Race (white/other), %	93/7	95/5	94/6
Disease duration (years), mean (SD)	15.0 (8.8)**	9.7 (8.2)	12.5 (8.9)
Concomitant use of DMARD, † No (%)	16 (36)	16 (41)	32 (38)
Hydroxychloroquine	0	1 (3)	1 (1)
Methotrexate	6 (13)	5 (13)	11 (13)
Sulfasalazine	11 (24)	11 (28)	22 (26)
Concomitant use of oral NSAID, No (%)	40 (89)	33 (85)	73 (87)
Concomitant use of corticosteroid, No (%)	7 (16)	6 (15)	13 (15)
Mean scores on ASAS components:			
Spinal inflammation	67.5	62.9	65.4
Nocturnal and total back pain	60.0	56.1	58.2
Patient global assessment	65.6	63.4	64.6
Functional impairment	60.2	57.2	58.8
Mean BASDAI scores overall	61.0	58.6	59.9
BASDAI scores <40, No (%)	4 (9)	5 (13)	10 (12)
CRP level (mg/L) (median ±SD)	154	97	-
Spinal mobility measures (mean ±SD)			
Chest expansion, cm	3.3	3.9	[redacted]
Schober's test, cm	12.2	12.8	[redacted]
Occiput to wall measurement, cm	7.3	4.6	[redacted]

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation.
*p<0.05 vs placebo by analysis of variance (ANOVA).
**P<0.01 vs placebo (ANOVA).
†One patient in each group took more than one DMARD.

3.5.7.3 Results: study 311-EU/Calin, 2004

Primary efficacy outcome

ASAS 20

At Week 12, significantly more patients in the etanercept group were ASAS 20 responders compared with patients in the placebo group (26 [60%] vs 9 [23%], respectively; p<0.001; 95% confidence interval (CI) 17.4 to 56.4%). There was no significant effect with respect to the use of concomitant DMARDs on the primary efficacy endpoint (p=0.632). At Week 12, there was no interaction effect between etanercept and DMARDs on ASAS 20 (p=0.694).

Rapid clinical responses to therapy were evident in etanercept patients from Week 2 and were sustained through all assessment visits.

Secondary efficacy outcomes

There were significantly more responders at the ASAS 50 (p<0.01) level at all visits and at the ASAS 70 level at Weeks 2, 4 and 8 (p<0.05) in the etanercept group compared with the placebo group. By Week 12, 49% of the etanercept patients had 50% or more improvement on the ASAS criteria vs 10% of placebo patients while 24% of etanercept patients experienced 70% or more improvement compared with 10% of placebo patients. Etanercept patients had significantly improved mean scores on the individual ASAS components compared with placebo patients (p<0.01 for all components).

Patients treated with etanercept had statistically significant greater improvements in spinal inflammation, nocturnal and total pain, patient global assessment of disease activity and functional index scores compared with placebo treated patients at Week 12 (Table 34). Etanercept patients had significantly more improvement in nine out of ten types of functional abilities as assessed by the BASFI. In addition, etanercept patients had significantly greater improvements in BASDAI composite index, BASDAI fatigue component scores, and BASFI total score compared with the placebo group (Table 34). At Week 12, the percentage of etanercept patients with BASDAI scores <40 increased from 9% at baseline to 71% (post hoc analysis).

Table 34: Summary of secondary efficacy outcomes, study 311-EU/Calin, 2004

Response measure: assessment point	Etanercept (n=45)	Placebo (n=39)	p value†
ASAS: Spinal inflammation			
Baseline (mean)	67.5	62.9	
Week 12 (% change)	35.6 (43.3)	52.6 (15.9)	0.003
ASAS: Nocturnal and total pain			
Baseline (mean)	60.0	56.1	
Week 12 (% change)	31.0 (43.1)	51.2 (6.2)	0.000
ASAS: Patient global assessment			
Baseline (mean)	65.6	63.4	
Week 12 (% change)	38.4 (37.0)	54.1 (12.6)	0.011
ASAS: Functional impairment (BASFI)			
Baseline (mean)	60.2	57.2	
Week 12 (% change)	39.6 (35.4)	53.9 (3.4)	0.000
BASDAI: Composite score			
Baseline (mean)	61.0	58.6	
Week 12 (% change)	33.8 (43.6)	50.1 (13.6)	0.001
BASDAI: Fatigue score			
Baseline (mean)	68.2	59.0	
Week 12 (% change)	38.4 (42.6)	54.8 (-4.9)	0.000

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index.
†Differences in mean percentage change, baseline to Week 12 (Cochrane-Mantel-Haenszel test).

On examination of acute phase reactants, there was a significant decrease in levels of CRP and ESR in etanercept patients compared with placebo (p<0.0001). In terms of spinal

mobility, etanercept treated patients had a significant improvement in spinal flexion, as measured by Schober's test compared with placebo (p<0.01). While not significant, other measures of spinal mobility showed numerical improvements in etanercept patients compared with placebo patients (Table 35).

Table 35: Comparison of acute phase reactants and spinal mobility values at baseline and at Week 12, study 311-EU/Calin, 2004

Measurement: assessment point	Etanercept	Placebo	p value
CRP (mg/l)	n=45	n=39	
Baseline (median)	15.4	9.7	
Week 12 (median % change)	4.0 (69.5)	11.7 (0.0)	0.000
ESR (mm/1st h)	n=45	n=39	
Baseline (median)	27.0	26.0	
Week 12 (% change)	6.0 (79.6)	29.0 (0.0)	0.000
Schober's test of spinal flexion (cm)	n=44	n=39	
Baseline (median)	2.2	2.8	
Week 12 (% change)†	2.7 (-36.0)	2.7 (1.3)	0.009
Chest expansion (cm)	n=44	n=39	
Baseline (median)	3.3	3.9	
Week 12 (% change)	3.8 (-29.9)	4.1 (29.0)	0.870
Occiput to wall distance (cm)	n=32‡	n=22‡	
Baseline (median)	7.3	4.6	
Week 12 (% change)	6.2 (12.5)	4.0 (7.2)	0.065

Abbreviations: CRP, C reactive protein; ESR, erythrocyte sedimentation rate; SD, standard deviation.
*Differences in average percentage change, baseline to Week 12 (Cochrane Mantel-Haenszel test).
†Negative percentage change from baseline indicates improvement.
‡Includes only patients who had a value greater than zero at baseline.

3.5.7.4 Efficacy conclusion: study 311-EU/Calin, 2004

This randomised controlled trial demonstrates that compared with placebo, etanercept 25 mg, self-administered twice weekly produced rapid, significant and sustained improvement in clinical and biological measures of AS, including ASAS response criteria and composite and fatigue BASDAI scores.

3.5.8 Study: 312-EU/Dijkmans, 2009

3.5.8.1 Methodology: study 312-EU/Dijkmans, 2009

The source of the efficacy data for study 312-EU is Dijkmans et al, 2009 (120). This was a 96 week, open-label extension follow up study to assess the longer term safety and efficacy of etanercept in patients with AS who had previously enrolled in study 311-EU/Calin, 2004. A summary of the methodology for study 312-EU/Dijkmans, 2009 is provided in Table 36.

Table 36: Methodology: study 312-EU/Dijkmans, 2009

Study	312-EU/Dijkmans, 2009
Objectives	To evaluate the longer term efficacy and safety, including radiographic outcomes, in patients receiving etanercept for treatment of AS.
Study location	The study was conducted in 14 centres in eight countries in Europe: UK, Belgium, Finland, France, Germany, Italy, The Netherlands, and Spain.
Design details	Patients who had previously received etanercept in study 311-EU continued to receive etanercept for the duration of the extension study. Patients who had previously received placebo in study 311-EU switched to etanercept therapy. Efficacy assessments were conducted every 4 weeks for the first 12 weeks followed by every 12 weeks thereafter.
Interventions	Patients received 25 mg etanercept subcutaneous injection, twice weekly.
Key inclusion criteria	• Previously completed study EU-311/Calin, 2004 • Aged 18–70 • Diagnosis of active AS according to modified New York criteria
Key exclusion criteria	Detailed exclusion criteria are provided in Table 32 Table 32
Efficacy outcomes	Primary efficacy endpoint • The proportion of patients achieving an ASAS 20, ASAS 40 and ASAS 5/6 response Secondary efficacy endpoints • Radiographic progression assessed by the mSASSS • BASFI, BASDAI, nocturnal and total back pain, patient and physician global assessments, spinal mobility (modified Schober's test, chest expansion and occiput to wall distance), number of painful joints (70 joints) and swollen joints (68 joints), CRP levels, and the proportion of patients achieving partial remission.
Populations analysed	Patients receiving the study drug.
Statistical analysis	Missing data was handled using the LOCF approach. Separate and overall total mSASSS scores were calculated for the cervical spine and the lumbar spine. Only patients with total baseline and Week 48 mSASSS scores were included in the radiographic analysis. If >3 scoring vertebral units in the cervical or lumbar spine sections were missing, the total mSASSS score was assigned a missing value. If ≤3 vertebral units in a section were missing, the mean of the other scoring units was used as a substitute for the missing units. Basic summary statistics were used to report treatment sequence for all patients. Within group changes from baseline were tested using paired t-tests. The intraclass correlation coefficient between the two matched readings of nine patients was calculated to assess intra-reader variability for baseline and Week 48 images.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; LOCF, last observation carried forward; mSASSS, modified Stoke Ankylosing Spondylitis Spine Score.

3.5.8.2 Patient population: study 312-EU/Dijkmans

Patient disposition

From a total of 84 patients enrolled in study 311-EU, 82 patients (43 from the etanercept group and 39 from the placebo group) entered the extension study. Eighty one patients received the study drug and were included in the analysis of the extension study. A total of 61 patients (75%) completed the open-label extension study. Of the 20 patients who withdrew from the extension study, 15 discontinued due to AEs, one patient withdrew because of lack of efficacy and 4 patients discontinued due to other reasons.

Patient demographics and baseline characteristics

The majority of the patients who participated in the extension study were male (78%) with a mean age of 43 years. The mean disease duration was 12.5 years.

3.5.8.3 Results: study 312-EU/Dijkmans

Primary efficacy outcome

The proportion of patients achieving an ASAS 20 response in the group who had previously received etanercept in study 311-EU remained constant for the duration of the extension study. Throughout the extension study period, the proportion of patients achieving ASAS 20 in the group who had previously received placebo continued to increase. At Week 108, 79% of all patients achieved an ASAS 20 response. At Week 108, ASAS 40 and ASAS 5/6 responses were achieved in 59% and 30% of patients, respectively.

Secondary efficacy outcomes

Secondary efficacy outcomes are shown in Table 37. Both treatment groups showed continued improvement in the BASDAI score up to Week 108. At Weeks 60 and 108, overall BASDAI scores showed a 60% and 57% improvement from baseline, respectively. Continual improvement was seen in individual components of the ASAS criteria throughout the extension study in both treatment groups. After 108 weeks, less pain was reported in the patients continuing to receive etanercept group compared with those patients who had switched to etanercept from placebo. Early and sustained improvements were seen in two of the assessments of spinal mobility in the patients who continued to receive etanercept whereas improvements in all three components were observed in the group switching from placebo.

Table 37: Secondary efficacy outcomes: study 312-EU/Dijkmans, 2009

	Etanercept† **(n=42)**			**Placebo to etanercept†** **(n=39)**		
Outcome	**Baseline**	**Week 60**	**Week 108**	**Baseline**	**Week 60**	**Week 108**
BASDAI	60.9	21.3 (65.1)	23.0 (62.3)	58.6	27.0 (53.9)	29.1 (50.3)
ASAS components						
Patient global assessment	66.2	23.0 (65.3)	21.9 (67.0)	63.4	29.8 (53.0)	28.8 (54.5)
Total back pain	62.2	18.7 (70.0)	22.5 (63.9)	56.5	27.3 (51.7)	29.9 (47.1)
Nocturnal back pain	59.1	15.0 (74.6)	17.2 (70.9)	55.8	21.7 (61.0)	25.2 (54.9)
BASFI	60.1	29.3 (51.3)	30.4 (49.4)	57.2	33.7 (41.1)	34.6 (39.6)
Morning	67.9	25.5 (62.4)	25.3 (62.7)	62.9	25.7 (59.1)	27.9 (55.6)

stiffness						
Spinal mobility						
Modified Schober's test	2.2	2.7 (26.0)	2.6 (21.4)	2.8	2.9 (5.1)	3.0 (6.9)
Chest expansion	3.4	4.4 (27.7)	4.6 (33.6)	3.9	4.6 (16.3)	4.8 (21.8)
Occiput to wall	7.0	6.0 (14.1)	6.8 (3.2)	4.6	2.1 (53.4)	2.5 (44.7)

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index.
†Data are presented as mean scores (percentage change from baseline).

At Week 108, the mean improvements in the number of swollen joints were 76.1% and 78.4% in those patients continuing etanercept and in those switching from placebo, respectively. Improvements in the number of painful joints were 68.2% and 63.1% in the patients continuing etanercept and in the patients switching to etanercept, respectively.

There was no evidence of deterioration or disease progression from baseline as assessed by mSASSS for cervical and lumbar spine in either treatment arm (Table 38).

Table 38: mSASSS scores for cervical and lumbar spine: study 312-EU/Dijkmans, 2009

	Etanercept n=33	**Placebo to etanercept n=34**
Baseline, mean (SD)	18.27 (21.0)	11.95 (16.8)
Week 60, mean (SD)	18.63 (20.9)	11.79 (16.8)
Change, mean (95% CI)	0.36 (-0.1, 0.8)	-0.15 (-0.7, 0.4)
Change, median (interquartile range)	0.00 (0.0, 0.4)	0.00 (0.0, 0.0)

Abbreviations: CI, confidence interval; SD, standard deviation.

3.5.8.4 Efficacy conclusion: study 312-EU/Dijkmans,2009

This 96 week open-label extension study demonstrated that the early clinical benefits of etanercept in patients with AS (previously seen in study 311-EU/Calin, 2004) were sustained for up to two years. These included achievement of ASAS response, marked improvements in the individual components of the ASAS criteria and improvements in the BASDAI and in measures of spinal mobility.

3.5.9 Study 907-EU/Martin-Mola, 2010

3.5.9.1 Methodology: study 907-EU/Martin-Mola, 2010

The source of the efficacy data for study 907-EU is Martin-Mola et al, 2010 (122) (NCT00444340) and the clinical study report (163). This was a long-term, open-label, multicentre Phase IV extension study to assess the safety and efficacy of etanercept treatment in patients with AS who had previously completed the placebo-controlled study 311-EU and the open-label follow up study 312-EU. A summary of the methodology for study 907-EU/Martin-Mola, 2010 is provided in Table 39.

Table 39: Methodology: study 907-EU/Martin-Mola, 2010

Study	907-EU/Martin-Mola, 2010 (NCT00444340)
Objectives	To evaluate the long term safety (primary objective) and sustained efficacy (secondary objective) of etanercept in patients with AS who had previously completed Studies 311-EU and 312-EU.
Study location	The study was conducted in the UK, Belgium, Finland, France, Germany, Italy, The Netherlands and Spain.
Design details	Patients who completed study 311-EU and 312-EU were eligible to enrol in the extension study. Patient assessments were performed at baseline (immediately after the completion of the Week 96 visit of study 312-EU) and at approximately 3-month intervals for up to 3 years. Each patient received treatment for up to 156 weeks.
Interventions	Patients received etanercept 25 mg subcutaneous injections twice weekly.
Key inclusion criteria	• As described for study 311-EU (Table 32) • Completion of study 311-EU and the open-label follow up, study 312-EU
Key exclusion criteria	• Discontinuation from study 312-EU for safety or other reasons • Abnormal haematology or chemistry profile (e.g. white blood cell count, haemoglobin levels and platelet count) • Clinically relevant concurrent medical events including: cancer/history of cancer; serious infection; a diagnosis of any central demyelinating disease; uncompensated congestive heart failure; presence/history of blood dyscrasias
Efficacy outcomes	<u>Primary efficacy endpoints</u> • ASAS 20, ASAS 50 and ASAS 70 response • Change from baseline in individual ASAS components (patient global assessment, total and nocturnal pain, BASFI, morning stiffness) • Frequency of and time to partial remission <u>Secondary efficacy endpoints</u> • Change from baseline in physician global assessment, BASDAI, spinal mobility (modified Schober's test, chest expansion, occiput to wall distance), complete joint assessment and CRP levels <u>Other endpoints</u> • ASAS 40 and ASAS 5/6 • BASDAI 50
Populations analysed	Efficacy analyses were conducted on all patients who received at least one dose of study drug in the extension study.
Statistical analysis	Missing data was handled using the LOCF approach. As this was an open-label study, efficacy analyses were descriptive only. ASAS 40, ASAS 5/6 and the BASDAI 50 were performed as post-hoc analyses.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP,

C reactive protein; DMARD, disease modifying anti-rheumatic drug; LOCF, last observation carried forward; NSAID, non-steroidal anti-inflammatory drug.

3.5.9.2 Patient population: study 907-EU/Martin-Mola, 2010

Patient disposition

Out of a total of 61 eligible patients completing study 312-EU, 59 patients continued onto the extension study. All patients received at least one dose of study drug. Thirty seven patients (62.7%) completed the extension study. Fourteen patients (23.7%) discontinued the study. The most common reason for discontinuation was AEs (n=7; 11.9% of patients), followed by 'other reasons', including protocol violations and non-medical reasons (n=6; 10.2%).

Patient demographics and baseline characteristics

The majority of the patients continuing onto the extension study were male (76%), the mean patient age was 44 years and the mean disease duration was 14.1 years.

3.5.9.3 Results: study 907-EU/Martin-Mola, 2010

Primary efficacy outcome

At Week 156, there were no appreciable changes in the proportions of ASAS responders in each response category. Rather, by Week 156 of the extension study the proportion of patients achieving ASAS 20, ASAS 50 and ASAS 70 responses was sustained over time. Despite a fluctuation in the mean outcome values at specific time points during this study, the overall trend demonstrated a sustained treatment response on all outcome measures in comparison to baseline. The results observed for the individual ASAS components were consistent with the overall ASAS response results (Table 40).

Table 40: Primary efficacy outcomes: study 907-EU/Martin-Mola, 2010

	Etanercept 25 mg (n=59)		
Outcome	**Baseline Week 12 (study 312-EU)**	**Baseline Week 96 (study 907-EU)**	**Week 156 (study 907-EU)**
ASAS 20, n (%)	28/58 (48)	52 (88)	44 (75)
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
ASAS 40 (post hoc analysis), n (%)	23/58 (40)	44 (75)	40 (68)
ASAS 5/6 (post hoc analysis), n (%)	12/58 (21)	19 (32)	19 (32)
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society; BASFI, Bath Ankylosing Spondylitis Functional Index; †Values are presented as the mean (percentage) improvement from baseline.

Secondary efficacy outcomes

At Week 156, the efficacy results observed in the secondary efficacy clinical parameters were consistent with the overall ASAS response results in that responses were maintained over time (Table 41). Again, despite a fluctuation in the mean outcome values at specific time points during this study, the overall trend showed a sustained treatment response on all outcome measures in comparison to baseline.

Table 41: Secondary efficacy outcomes: study 907-EU/Martin-Mola, 2010

	Etanercept 25 mg (n=59)		
Outcome	**Study 312-EU baseline**	**Study 907-EU baseline**	**Study 907-EU Week 156**
Physician global assessments†	[redacted]	[redacted]	19.2 (66.4)
BASDAI†	[redacted]	[redacted]	26.6 (55.5)
BASDAI 50, n (%)	[redacted]	[redacted]	39 (66)
Spinal mobility measures†			
Modified Schober's test	[redacted]	[redacted]	2.9 (18.9)
Chest expansion	[redacted]	[redacted]	4.9 (32.7)
Occiput to wall	[redacted]	[redacted]	4.2 (26.6)
CRP levels, mg/L†	[redacted]	[redacted]	9.1 (64.1)

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein.
†Values are presented as the mean (percentage) improvement from study 311-EU baseline.

3.5.9.4 Efficacy conclusion: study 907-EU/Martin-Mola, 2010

The results of study 907-EU demonstrate that the positive efficacy outcomes achieved with etanercept 25 mg twice weekly in study 312-EU, including improvements in physical function and spinal mobility, were largely sustained in comparison to baseline throughout the duration of this 156 week extension study.

3.5.10 Study: SPINE trial/Dougados, 2011

3.5.10.1 Methodology: SPINE trial/Dougados, 2011

The source of the efficacy data for the SPINE trial is Dougados et al, 2011 (13), (NCT00420238). SPINE was a multicentre, 12-week, randomised, double-blind, placebo-controlled trial to evaluate the effect of etanercept in patients with advanced AS. The SPINE trial differs from the clinical trials of etanercept previously presented in this section as it focuses specifically on a patient population with advanced AS with an average disease duration of 21 years who have severe spinal ankylosis. The latter was an exclusion criterion in the other etanercept trials. The evidence from the SPINE trial therefore emphasises that the efficacy of etanercept has been studied across a broad spectrum of AS patients. A summary of the methodology and efficacy results for the SPINE trial is provided in Table 42.

Table 42: Methodology: SPINE trial/Dougados, 2011

Study	SPINE trial/Dougados, 2011 (NCT00420238)
Objectives	To assess the efficacy of etanercept on the symptoms and signs of advanced AS.
Study location	The study was conducted at 21 centres in France, Germany, Hungary and The Netherlands.
Design details	Patient assessments were conducted at screening, baseline and at Weeks 2, 4, 8 and 12.
Interventions	Patients were randomised to receive either: • Etanercept 50 mg once weekly, or • Placebo
Key inclusion criteria	• Aged 18–70 • Diagnosis of AS as defined by the modified New York criteria and presenting with one of the following three criteria defining advanced and severe spinal ankylosis: two intervertebral adjacent bridges and/or fusion at the lumbar spine; three intervertebral adjacent bridges and/or fusion at the thoracic spine; or two intervertebral adjacent bridges and/or fusion at the cervical spine. The radiological stage of AS was determined on a scale of 0–V[†]. • Active refractory disease as defined by a score ≥40 on the BASDAI despite NSAID treatment (≥2 NSAIDS at the maximum tolerated dose for ≥3 months) • Pain with axial involvement of the overall level of AS back, hip or neck for a score ≥30 on a VAS
Key exclusion criteria	• Previous therapy with an anti-TNF • A change in dose of NSAID within 2 weeks of baseline • A change in the dose of concomitant DMARD within 4 weeks of baseline • Significant concurrent medical disorders (e.g. serious infection, cancer) • Abnormal laboratory test results
Permitted concomitant medications	• Doses of any concomitant DMARDs and NSAIDs must remain stable throughout the 12-week study period • Analgesics (with or without codeine or dextropropoxyphen) were permitted in cases of painful episodes during the study
Efficacy outcomes	<u>Primary efficacy endpoint</u> • Improvement in the BASDAI from baseline to Week 12 <u>Secondary efficacy endpoints</u> • Proportion of patients achieving ASA20, ASAS 40 and ASAS 5/6 • Proportion of patients achieving partial remission • Changes from baseline in patient's global assessment of disease activity,

Study	SPINE trial/Dougados, 2011 (NCT00420238)
	BASFI, BASMI, CRP levels • Health outcomes assessments as measured by the MCII and PASS Other outcomes • Pulmonary function tests in accordance with American Thoracic Society standards
Populations analysed	Assessment of efficacy was based on the mITT population which included all randomised patients who received at least one dose of blinded study drug.
Statistical analysis	The population sample size of 40 patients per treatment arm was based on a power of 90% and a two-sided type 1 error of 5%. The primary endpoint was measured as the normalised net incremental AUC. For the primary endpoint, the linear trapezoidal method, divided by the number of days the patient remained in the study, was used to calculate the area between baseline and the patient global assessment curve as a function of time. The primary endpoint was analysed using an ANCOVA method with treatment as a factor and BASDAI at baseline as a covariate. Changes from baseline in the primary and secondary endpoints were analysed using a mixed model ANCOVA with an autoregressive correlation structure with baseline scores as a covariate and treatment groups, visits and their interaction as fixed factors. A generalised estimating equation model with an autoregressive correlation structure, a logit link and a binomial distribution was used for analysis of binary efficacy variables. The Kaplan-Meier method was used to estimate the time to achieve a sustained PASS with the log rank test used for comparison of survival curves. Missing data was handled using the LOCF approach.

Abbreviations: ANCOVA, analysis of covariance; AS, ankylosing spondylitis; AUC, area under the curve; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; LOCF, last observation carried forward; NSAID, non-steroidal anti-inflammatory drug; MCII, minimum clinically important improvement; PASS, patient acceptable symptom state; TNF, tumour necrosis factor; VAS, visual analogue scale.
†Stages I-III =moderate radiographic evidence of structural damage involving <50% (≤12 vertebrae) in two spinal segments. Stage IV = advanced AS involving 50–80% of the spine in >2 spinal segments (13–19 vertebrae). Stage V = advanced AS involving ≥80% of the spine.

3.5.10.2 Patient population: SPINE trial/Dougados, 2011

Patient Disposition

Out of a total of 95 screened patients, 82 were randomised to receive study drug (39 patients in the etanercept group and 43 patients in the placebo group). One patient in the etanercept group discontinued due to a pre-existing lung neoplasm while four patients discontinued in the placebo group; two patients for lack of efficacy, one patient lost to follow up and one due to withdrawal of consent.

Patient demographics and baseline characteristics

There were no significant differences in patient demographics and baseline disease characteristics between treatment arms. The majority of patients in the study were male (93%), the mean patient age was 47 years and the mean duration of disease was 21 years. A greater proportion of patients were classified as spinal radiological stage III (57%) (Table 43).

Table 43: Patient demographics and baseline characteristics: SPINE trial/ Dougados, 2011

Characteristics	Etanercept† (n=39)	Placebo† (n=43)
Age, years	46±11	48±10
Male, n (%)	37 (95)	39 (91)
Disease duration, years	19±10	23±11
HLA-B27 positive, n (%)	31 (79)	36 (86)
Spinal radiological stage, n (%)		
Stage II	7 (18)	3 (7)
Stage III	21 (54)	26 (60)
Stage IV	7 (18)	9 (21)
Stage V	4 (10)	5 (12)
Intervertebral radiological fusion (0–24)	9±6	10±6
Radiological mSASSS (0–72)	34±22	39±19
Radiological hip abnormalities (coxitis)‡	12 (32)	14(35)
BASDAI	64±12	58±15
BASFI	63±20	57±19
BASMI	5.7±1.4	5.8±1.3
Total back pain	70±16	61±20
CRP level, mg/L	25±31	17±19
ASDAS-CRP	3.90±0.71	3.63±0.76
Disease activity state, n (%)		
ASDAS-CRP <1.3 (inactive disease)	0 (0)	0 (0)
ASDAS-CRP >1.3–<2.1 (moderate disease activity)	0 (0)	1 (2.4)
ASDAS-CRP >2.1–<3.5 (high disease activity)	12 (30.8)	17 (41.5)
ASDAS-CRP ≥3.5 (very high disease activity)	27 (69.2)	23 (56.1)

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score – C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; HLA, human leucocyte antigen; mSASSS, Modified Stoke Ankylosing Spondylitis Spine Score.
†Data are presented as mean ± SD.
‡On pelvic x-ray assessment.

3.5.10.3 Results: SPINE trial/Dougados, 2011

Primary efficacy outcomes

Over 12 weeks of treatment, the mean (±SD) normalised net incremental AUC for the BASDAI was significantly greater in the etanercept group compared with the placebo group (-11±16 vs -20±17, respectively; p=0.019).

Secondary efficacy outcomes

The etanercept group showed significantly greater improvements in the absolute changes from baseline to Week 12 in the BASDAI, BASFI, BASMI, total back pain and CRP levels compared with the placebo group (Table 44). Significant changes in the BASDAI over time were observed from Week 8 and were sustained until Week 12 in the etanercept group compared with the placebo group (Table 44).

Table 44: Absolute changes from baseline to Week 12 in clinical parameters: SPINE trial/Dougados, 2011

	Adjusted changes (baseline to Week 12)		
Parameter	**Etanercept (n=39)**	**Placebo (n=43)**	**p value**
BASDAI	−26±20	−14±20	0.008
BASFI	−22±18	−10±18	0.004
BASMI	−0.57±0.65	−0.20±0.65	0.011
Total back pain	−29±24	−15±24	0.010
CRP level, mg/L	−16±14	−1±14	<0.0001
ASDAS-CRP	−1.51±0.87	−0.49±0.87	<0.0001

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score-C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index.

Significantly more patients in the etanercept group achieved ASAS 20, ASAS 5/6 and BASDAI 50 responses compared with the placebo group. There were no significant differences in ASAS 40 and ASAS partial remission between treatment arms. The etanercept group showed significant changes in ASDAS-CRP from baseline compared with the placebo group (Table 45).

Table 45: Secondary efficacy outcomes (responder criteria): SPINE trial/Dougados, 2011

Responder criteria	**Etanercept (n=39)**	**Placebo (n=43)**	**p value**
ASAS 20	25 (67%)	14 (33%)	0.003
ASAS 40	17 (44%)	10 (23%)	0.053
ASAS 5/6	8 (21%)	2 (5%)	0.044
ASAS partial remission	7 (18%)	2 (5%)	0.073
BASDAI 50	18 (46%)	10 (23%)	0.031
ASDAS-CRP (change from baseline)			
ASDAS ≥1.1 (minimally important improvement)	25 (64.1)	7 (17.1)	<0.0001
ASDAS ≥2.0 (major improvement)	15 (38.5)	1 (2.4)	<0.0001
ASDAS-CRP status at Week 12			
ASDAS <1.3	17 (43.6)	35 (85.4)	<0.001
ASDAS <2.1	25 (64.1)	41 (100)	<0.0001

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Score; CRP, C reactive protein.

According to the MCII, a significantly greater proportion of patients in the etanercept group showed “moderately” or “very improved” scores compared with the placebo group (63% vs 38%, respectively). At Week 12, a greater proportion of patients in the etanercept group compared with the placebo group considered their condition as “acceptable” according to the PASS (61% vs 41%, respectively; p=0.065). By Week 12, a significantly lower proportion of patients in the etanercept group had a pulmonary restrictive pattern compared with the placebo group (p=0.032) indicating improvement in respiratory function.

3.5.10.4 Efficacy conclusions: SPINE trial/Dougados, 2011

The results of the SPINE trial demonstrate that short term therapy with etanercept 50 mg once weekly in patients with advanced AS is associated with significant improvements in the clinical symptoms (BASDAI, total back pain and MCII scores), markers of inflammation (CRP), spinal mobility (BASMI) and pulmonary function. The efficacy results presented for the SPINE study are particularly important as the data add to and strengthen the current evidence base for the clinical efficacy of etanercept. The SPINE trial results, taken together with the results of the previously reported etanercept AS trials, demonstrate that etanercept is clinically effective throughout the entire disease spectrum.

3.5.11 Study: Dougados, 2012

3.5.11.1 Methodology and efficacy outcomes: Dougados, 2012

The source of the efficacy data is Dougados et al (121). This study was a 12 week open-label extension of the SPINE trial to assess the longer term efficacy of etanercept treatment of patients with advanced AS (13). A summary of the methodology for Dougados, 2012 is presented in Table 46.

Table 46: Methodology and efficacy outcomes: Dougados, 2012

Study	Dougados, 2012
Objectives	To evaluate the longer term efficacy of etanercept in patients with severe and advanced AS.
Study location	The study was conducted at 21 centres in France, Germany, Hungary and The Netherlands.
Design details	Patient assessments were conducted at baseline and at Weeks 2, 4, 8 and 12 of the double-blind study. Assessments were performed at Weeks 14, 18 and 24 of the open label extension study.
Interventions	Patients received subcutaneous etanercept 50 mg, once weekly for 12 weeks.
Key inclusion criteria	• Patients who had previously completed the 12 week SPINE study were eligible to enrol in the 12 week open label extension study • Full inclusion criteria are provided in Table 42
Key exclusion criteria	• Full exclusion criteria are provided in Table 42
Permitted concomitant medications	• Doses of any concomitant DMARDs and NSAIDs must remain stable throughout the 12-week study period • Analgesics (with or without codeine or dextropropoxyphen) were permitted in cases of painful episodes during the study
Efficacy outcomes	• Improvements in the BASDAI, BASFI, BASMI, and total back pain • Change in CRP levels and ASDAS-CRP score • ASAS 20, ASAS 40 and ASAS 5/6 response, • ASAS partial remission • BASDAI 50 response
Populations analysed	All patients who entered the open-label extension phase of the trial
Statistical analysis	Outcomes were assessed using descriptive statistics. Missing data was handled using the LOCF approach.

Abbreviations: AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; NSAID, non-steroidal anti-inflammatory drug;

3.5.11.2 Patient population: Dougados, 2012

Patient Disposition

A total of 77 patients were enrolled in the open-label extension study. Seventy four patients completed the 12-week extension study: 36 patients continuing with etanercept and 38 patients switching from placebo. Of the three patients who failed to complete the extension phase, two withdrew consent (one from each group) while one patient continuing with etanercept was lost to follow-up.

Patient demographics and baseline characteristics

The demographic and disease baseline characteristics of patients entering the open-label extension trial were similar to those in the 12 week double-blind period (Table 43). The majority of patients in the study were male (92%), the mean patient age was 47.5 years and the mean duration of disease was 16 years. A greater proportion of patients were classified as spinal radiological stage III (57%)

3.5.11.3 Results: Dougados, 2012

Efficacy outcomes

Patients continuing with etanercept in the open label extension study showed continuing improvement in BASDAI, BASFI and BASMI scores up to Week 24. Improvements in the BASDAI, BASFI and BASMI were also observed in patients switching from placebo to etanercept over the course of the extension study. Both treatment groups showed continued improvement in total back pain up to Week 24 (Table 47). CRP levels in the patients continuing etanercept remained stable during the extension study while patients previously receiving placebo achieved a reduction in CRP level when switching to etanercept in the open label period (Table 47).

Table 47: Absolute changes from baseline to Week 24 in clinical parameters: Dougados, 2012

	Etanercept† (adjusted change from baseline,95% CI)		Placebo to etanercept† (adjusted change from baseline,95% CI)	
Parameter	**Week 12 (n=38)**	**Week 24 (n=38)**	**Week 12 (n=39)**	**Week 24 (n=39)**
BASDAI	−27.4 (23.8)	−37.6 (22.4)	−15.0 (20.0)	−28.6 (24.3)
BASFI	−22.4 (19.5)	−28.9 (20.8)	−10.7 (18.5)	−21.8 (20.9)
BASMI	−0.6 (0.8)	−0.8 (0.8)	−0.2 (0.6)	−0.5 (0.7)
Total back pain	−32.7 (25.8)	−44.5 (25.1)	−13.8 (27.5)	−32.3 (31.8)
CRP level, mg/L	−18.5 (29.5)	−16.1 (33.0)	1.4 (9.2)	−9.6 (13.0)
ASDAS-CRP	−1.6 (1.1)	−1.9 (1.2)	−0.5 (0.8)	−1.7 (1.1)

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score-C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CI, confidence interval.
†Data are presented as mean (SD).

Secondary efficacy outcomes

The proportions of patients achieving ASAS 20 and ASAS 40 response in both treatment arms continued to increase up to Week 24 of the open label extension study. By Week 24, ASAS partial remission was achieved by 29% of patients continuing with etanercept treatment and 23% of patients switching from placebo to etanercept (Table 48). A BASDAI 50 response was achieved by 65.8% of patients continuing with etanercept and 48.7% of patients switching from placebo. More than 75% of patients continuing with etanercept achieved a clinically important improvement in the ASDAS-CRP while 26% of patients in this group reached inactive disease status. In the placebo to etanercept group, 63.2% of patients achieved a clinically important improvement in the ASDAS-CRP while 38.5% of patients achieved inactive disease status (Table 48).

Table 48: Secondary efficacy outcomes (responder criteria): Dougados, 2012

	Etanercept†		Placebo to etanercept†	
Parameter	**Week 12**	**Week 24**	**Week 12**	**Week 24**
ASAS 20	26 (68.4)	32 (84.2)	14 (35.9)	26 (66.7)
ASAS 40	17 (44.7)	26 (68.4)	10 (25.6)	21 (53.9)
ASAS 5/6	8 (21.0)	6 (15.8)	2 (5.1)	7 (18.0)
ASAS partial remission	7 (18.4)	11 (29.0)	2 (5.1)	9 (23.1)
BASDAI 50	18 (47.4)	25 (65.8)	10 (25.6)	19 (48.7)
ASDAS-CRP (change from baseline)				
ASDAS ≥1.1 (clinically important improvement)	7 (18.4)	29 (76.3)	7 (18.4)	25 (63.2)
ASDAS ≥2.0 (major improvement)	1 (2.6)	19 (50.0)	1 (2.6)	13 (34.2)
ASDAS-CRP inactive disease status (<1.3)	2 (5.1)	10 (26.3)	2 (5.1)	15 (38.5)

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score-C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index.
†Data are presented as number of patients (%).

3.5.11.4 Efficacy conclusions: Dougados, 2012

The results of this 12 week open label study showed that the clinical signs and symptoms of severe and advanced AS continued to improve after up to 24 weeks of etanercept therapy as evidenced by improvements in the BASDAI, BASFI, ASAS 20 and ASAS 40 and achievement of a BASDAI 50 response.

3.5.12 Study: Brandt, 2003

3.5.12.1 Methodology and efficacy outcomes: Brandt, 2003

The source of the efficacy data is Brandt et al, 2003 (89). Note that this was an independent study rather than a Pfizer sponsored trial.

Brandt, 2003 was a multicentre trial conducted in two phases to investigate the efficacy of etanercept in patients with active AS:

- Phase 1: placebo-controlled period of 6 weeks duration
- Phase 2: an observational phase of 24 weeks duration

A summary of the methodology and efficacy results for study Brandt, 2003 is provided in Table 49.

Table 49: Methodology and efficacy outcomes: Brandt, 2003

Study	Brandt, 2003
Objectives	To investigate the efficacy of etanercept in patients with active AS not currently receiving therapy with DMARDs or steroids.
Study location	Not reported in the publication.
Design details	An enrolment period of 4 months was followed by a treatment period of 6 weeks during which patients received etanercept or placebo. After Week 6, patients in the placebo group switched over to etanercept for the next 12 weeks while etanercept patients continued to receive etanercept for a further 6 weeks. The study remained blinded until Week 12. The follow up period was 24 weeks for the etanercept group and 30 weeks for the placebo group. Assessments were performed every 3 weeks during the observational period.
Interventions	During the first 6 weeks of the study, patients were randomised to receive either: • Etanercept 25 mg subcutaneous injection, twice weekly, or • Placebo 25 mg subcutaneous injection, twice weekly, switching to etanercept 25 mg twice weekly after Week 6.
Key inclusion criteria	• Patients fulfilling the modified New York criteria for AS • Active AS as defined by a BASDAI score ≥4 and spinal pain ≥4 on a 0-10 rating scale • Cessation of DMARD and oral corticosteroid use at least 4 weeks prior to screening
Key exclusion criteria	• Active tuberculosis within the last 3 years • Serious infection within the last 2 months • Lymphoproliferative disease or other malignancies within the last 5 years • Multiple sclerosis or a related disorder • Currently showing symptoms of severe disease
Permitted concomitant medications	• NSAID use (with no increase over the baseline dosage)
Efficacy outcomes	<u>Primary efficacy endpoint</u> • A ≥50% improvement in disease activity measured on the BASDAI between baseline and Week 6 <u>Secondary efficacy endpoints</u> • Improvements in scores for spinal pain, BASFI, BASMI, ASAS response criteria, SF-36, CRP and ESR levels <u>Other outcomes</u>

Study	Brandt, 2003
	• Time to relapse after cessation of etanercept treatment (only assessed in patients achieving at least a 20% improvement from baseline in the BASDAI after 3 months of etanercept treatment)
Populations analysed	Assessment of efficacy was based on the ITT population.
Statistical analysis	Power calculations were performed using one-sided Fisher's exact test. In this superiority study, a population size of 15 patients per treatment group was calculated to be sufficient to detect a difference between groups at a significance level of 0.05 with a power >88%. All other data were analysed using two-sided statistical tests. Fisher's exact test was used for comparisons in the response criteria. Response rates were calculated using 95% Clopper/Pearson confidence bounds. Analysis of covariance with the baseline value as covariable was used to compare means. Wilcoxon's unpaired rank sum test was used to compare values at baseline. Changes within single BASDAI items were compared using a paired t-test. Non-responders were patients violating the study protocol and patients who withdrew from the study.
Patient disposition	Out of a total of 33 eligible patients, 16 were randomised to receive etanercept while 17 were randomised to receive placebo. Two patients in the etanercept group and one in the placebo group were withdrawn from the study due to lack of compliance. One patient in the placebo group withdrew due to requiring surgery. In the etanercept group, 14 patients completed the study at Week 24. In the placebo group, 15 patients competed at Week 30.
Patient demographics and baseline characteristics	Baseline demographics were similar between treatment groups in terms of sex, disease duration, parameters of clinical disease, HLA-B27 status and radiographic changes in the spine. Patients in the placebo group were significantly younger compared with the etanercept group (p<0.05).
Efficacy outcomes	<u>Primary efficacy outcome</u> • At Week 6 (placebo-controlled phase), 8 patients (57%) treated with etanercept achieved ≥50% improvement in the BASDAI compared with 1 patient (6%) in the placebo group (p=0.004) • After 12 weeks of etanercept treatment, a 50% response was achieved in 71% of patients continuing with etanercept and in 56% of patients switching to etanercept after placebo <u>Secondary efficacy outcomes</u> • At Week 6 of the placebo-controlled phase, significant improvements in the BASDAI components, the BASFI, BASMI and CRP levels were observed in the etanercept group • At Week 6, an ASAS 20 response was achieved in 78.6% of etanercept patients versus 25% of placebo patients (p<0.01) while 42.9% of etanercept patients achieved an ASAS 50 response compared with 12.5% of placebo patients (p<0.01)

	Etanercept† (n=14)		**Placebo†** (n=16)		
Outcome	**Baseline**	**Week 6**	**Baseline**	**Week 6**	**p value**
BASDAI	6.5 ± 1.2	3.5 ± 1.9	6.6 ± 1.0	5.8 ± 2.0	0.003
BASDAI component					
Fatigue	6.9 ± 1.2	4.5 ± 1.7	NR	NR	0.0002
Spinal pain	7.6 ± 1.3	3.9 ± 2.1	NR	NR	<0.0001
Peripheral joint pain	5.4 ± 2.8	3.1 ± 2.4	NR	NR	0.002

Study	Brandt, 2003

Entheseal pain	7.0 ± 1.7	3.6 ± 2.4	NR	NR	<0.0001
Morning stiffness	5.5 ± 2.1	2.5 ± 1.5	NR	NR	0.0003
BASFI	6.2 ± 1.8	4.3 ± 2.3	5.3 ± 2.3	5.1 ± 2.4	0.008
BASMI	4.1	2.6	3.8 ± 2.1	3.5 ± 2.3	0.01
CRP level, mg/L	19 ± 17	<6 ± 4	NR	NR	0.0001

†Values are mean ± SD

- After 12 weeks of etanercept (i.e. Week 12 in the patients continuing etanercept and Week 18 in patients switching from placebo, 10/30 patients (33.3%) achieved partial remission
- Between baseline and Week 6, NSAID use was stopped in 38% of etanercept patients vs 13% of placebo patients
- In the first 6 weeks of treatment, 62% of etanercept patients reduced their NSAID dose by 50% compared with 7% of placebo patients (p=0.016)
- By Week 6, etanercept patients showed improvement in the physical component score of the SF-36 with no improvements seen in the placebo group (p=0.026).
- Patients switching to etanercept from placebo after Week 6 showed similar responses to those observed in etanercept patients in all efficacy outcome measures. Response was increased or sustained in both groups throughout the treatment period with etanercept.

Follow up

- Within the 3 month follow up period, 18/24 patients (75%) experienced a relapse after cessation of etanercept treatment (mean ± SD time to relapse: 6.2 ± 3.0 weeks) while the remaining 25% patients relapsed after the follow up period.

Efficacy conclusion	This independent study demonstrates that short term therapy (3 months) with etanercept 25 mg twice weekly is associated with significant and sustained improvements in the treatment of AS in patients receiving NSAIDs but not currently receiving DMARD or steroid therapy. Following discontinuation of therapy, the majority of patients experienced a relapse within 3 months.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; ESR, erythrocyte sedimentation rate; HLA, human leucocyte antigen; ITT, intention-to-treat; NR, not reported; NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation; SF-36, short form-36.

3.5.13 Study: Brandt, 2005

3.5.13.1 Methodology and efficacy outcomes: Brandt, 2005

The source of the efficacy data is Brandt et al, 2005 (116). This was a 54 week, open-label, observational extension study to assess the longer term safety and efficacy of etanercept in patients previously treated with etanercept in Brandt, 2003. A summary of the methodology and efficacy outcomes is provided in Table 50.

Table 50: Methodology and efficacy outcomes: Brandt, 2005

Study	Brandt, 2005
Objectives	To investigate the longer term efficacy and safety of etanercept over one year, with inclusion of discontinuation and re-administration.
Study location	The study was conducted at two centres in Berlin and Herne, Germany.
Design details	Patients who had previously received etanercept in Brandt, 2003 had discontinued etanercept therapy for a mean of 26.8 weeks prior to screening. All patients had relapsed since discontinuation of etanercept (defined as a score ≥4 on the BASDAI and a score ≥4 on a 0–10 rating scale for physician global assessment according to the recently proposed ASAS Working Group consensus statement on the use of anti-TNF agents in patients with AS (2). Patient assessments were conducted every 3 weeks for a duration of 54 weeks.
Interventions	Patients received etanercept 25 mg subcutaneous injection, twice weekly for 54 weeks
Key inclusion criteria	• Detailed inclusion criteria are provided in Table 49 • All patients who had completed Brandt, 2003 were eligible • A definite diagnosis of high disease activity after cessation of previous etanercept therapy (defined as a score ≥4 on the BASDAI and a score ≥4 on a 0–10 rating scale for spinal pain despite NSAID therapy)
Key exclusion criteria	• Not reported
Permitted concomitant medications	• NSAID use (with no increase over the baseline dosage)
Efficacy outcomes	<u>Primary efficacy endpoint</u> • A ≥50% improvement on the BASDAI between baseline and Week 54 <u>Secondary efficacy endpoints</u> • Improvements in mean scores for BASFI, BASMI, ASAS 20 and ASAS 40 response criteria, ASAS 5/6, SF-36, CRP and ESR levels
Populations analysed	Assessment of efficacy was based on the ITT population and all patients who entered the extension study were included in the analysis.
Statistical analysis	The 95% CIs for frequencies were calculated using the Wilson scoring method. The paired t-test and the paired Wilcoxon rank sum test were used for comparisons between visits. A significance level of 0.05 was used with no adjustments for repeated significance testing. Non-responders were patients who withdrew from the study and the last observation carried forward approach was used to include non-responders in the calculation of mean values.
Patient disposition	Of the 29 patients who completed Brandt, 2003, 26 were enrolled in the extension study. Three patients discontinued treatment; one due to an AE, one due to lack of efficacy and one because of non-compliance.
Patient demographics	Baseline demographics were as reported for Brandt, 2003 (Table 49). More than two thirds of the patients starting the extension study had relapsed 12 weeks after

Study	Brandt, 2005
and baseline characteristics	previous cessation of etanercept therapy while almost all had relapsed after 24 weeks.
Efficacy outcomes	<u>Primary efficacy outcome</u> • At Weeks 3 and 6, 30.8% and 46.2% of patients, respectively achieved a BASDAI 50% response. This improvement in disease activity was sustained up to Week 54 <u>Secondary efficacy outcomes</u> • An ASAS 20 response was achieved in 73.1% (95% CI:54–56) of patients while 62.5% (95% CI:17–50) of patients achieved ASAS 40 • ASAS 5/6 was achieved in 65% (95% CI:46–81) of patients • At Week 54, 8 patients (31%) were in partial remission • At Week 54, significant improvements in the BASDAI total score, BASDAI components, the BASFI, BASMI, CRP and ESR levels were observed compared to baseline.

	Etanercept†		
Outcome	**Baseline**	**Week 54**	**p value**
BASDAI, mean ± SD	6.4± 1.6	2.8 ± 2.3	<0.0001
Fatigue	7.1 ± 1.5	3.7 ± 2.7	<0.0001
Spinal pain	7.3 ± 1.6	3.3 ± 2.8	<0.0001
Peripheral joint pain	5.0 ± 3.0	2.1 ± 2.4	0.0001
Entheseal pain	6.5 ± 1.8	2.6 ± 2.6	<0.0001
Morning stiffness	5.9 ± 2.1	2.2 ± 2.3	<0.0001
BASFI	6.0 ± 2.1	3.5 ± 2.8	<0.0001
BASMI	3.3 ± 2.0	2.5 ± 2.1	0.004
Number of swollen joints, mean	0.7 ± 4.0	0.0 ± 1.8	0.043
Number of enthesitic sites	2.4 ± 2.6	0.6 ± 1.8	0.002
CRP level, mg/L, mean	14.4 ± 8.5	6.3 ± 2.7	<0.0001
ESR level, mm/h, mean	22.8 ± 15.3	9.5 ± 8.3	<0.0001

†Values are mean ± SD unless otherwise stated

• Improvements in all components of the SF-36 were observed between baseline and Week 54

Efficacy conclusion	This study demonstrates that re-administration of etanercept 25 mg twice weekly in patients with active AS who relapsed following cessation of prior etanercept therapy is associated with rapid improvements in the majority of clinical endpoints which were then sustained over 54 weeks. Specifically, significant improvements were observed in all components of the BASDAI with patients remaining at low disease activity over the one year study duration period.

Abbreviations: AE, adverse event; AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CI, confidence interval; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; ITT, intention-to-treat; NSAID, non-steroidal anti-inflammatory drug; SD, standard deviation; SF-36, short form-36; TNF, tumour necrosis factor.

3.5.14 Study Baraliakos, 2013

3.5.14.1 Methodology: Baraliakos, 2013

The source of the efficacy data is Baraliakos et al, 2013 (9) (NCT01289743). This was an open label extension of Brandt, 2003. This study was excluded by the clinical systematic review performed in support of this submission since it only included 26 patients, and the review criteria excluded studies of less than 30 participants. However, it was identified by Pfizer as relevant to the submission since it reports an extension of the Brandt 2003 study (identified by the systematic review) which included more than 30 participants, and represents the longest open-label extension of etanercept treatment in AS (7 years) conducted to date. Patients who experienced clinical relapse were enrolled in the open-label extension phase investigating the sustained efficacy of etanercept treatment over a 7 year period. A summary of the methodology for Baraliakos, 2013 is provided in Table 51.

Table 51: Methodology: Baraliakos, 2013

Study	Baraliakos, 2013 (NCT01289743)
Objectives	To assess the long-term efficacy and safety of etanercept for the treatment of active AS.
Study location	Not reported in the publication.
Design details	Patients who completed the Brandt, 2003 study protocol and experienced clinical relapse following etanercept discontinuation were enrolled in the extension study, in which patients were treated with etanercept for a period of 7 years. Clinical and patient-reported outcomes were used to assess efficacy parameters, including disease activity, patient function and mobility, and HRQoL.
Interventions	All patients received etanercept 25 mg subcutaneous injections, twice weekly for 7 years
Key inclusion criteria	Patients who had completed the placebo-controlled phase of the trial (Brandt, 2003), and had discontinued treatment, were eligible to enter the open-label extension phase if they had experienced clinical relapse, defined as: • A BASDAI score of ≥ 4 and • A score of ≥ 4 on a 0 to 10 rating scale
Key exclusion criteria	Detailed exclusion criteria are provided in Table 49
Permitted concomitant medications	• NSAID use (with no increase over the baseline dosage)
Efficacy outcomes	<u>Primary efficacy endpoint</u> • Percentage of patients in ASAS partial remission or inactive disease status at Year 7 <u>Secondary efficacy endpoints</u> • Proportion of patients who met the ASAS 40 and the ASAS 5/6 criteria after 7 years • Mean improvements in BASDAI, ASDAS, CRP, BASFI, BASMI and SF-36 scores after 7 years <u>Other efficacy outcomes</u> • Proportion of patients who achieved mean improvements in BASDAI and ASDAS score after 2 years • Proportion of patients who met the ASAS 40 criteria after 2 years • Disease activity and improvements in BASDAI, ASDAS, BASFI and CRP values in patients who did not reach a BASDAI score of <3 or moderate

Study	Baraliakos, 2013 (NCT01289743)
	disease activity at the end of the 7 year treatment period
Populations analysed	• All patients who completed the 7 year etanercept treatment period were included in the analysis of the primary and secondary efficacy outcomes • A sub-analysis was conducted using the population of patients who did not reach a BASDAI score of <3 or ASDAS moderate disease activity at the end of the 7 year treatment period (n=4)
Statistical analysis	The Mann-Whitney U test was used for comparisons between subgroups of patients and the Wilcoxon test was used for the comparison of variables between different time points, Fischer's exact test was used to compare categorical data. A significance level of 5% was used.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; ASDAS; Ankylosing Spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drugs; HRQoL, health-related quality of life; NSAID, non-steroidal anti-inflammatory drug.

3.5.14.2 Patient population: Baraliakos, 2013

Patient Disposition

Of the 26 patients who were randomised in the initial placebo-controlled study (Brandt, 2003), 21 patients (81%) completed the 2 year treatment period and 16 patients (62%) completed the 7 year study treatment period.

Patient demographics and baseline characteristics

Demographics and baseline characteristics of patients who completed the 7 year treatment period and those who discontinued treatment are shown in Table 52. Fourteen male and two female patients completed the 7 year treatment period. Baseline disease characteristics of patients who completed 7 years of etanercept treatment and those who discontinued treatment were similar, with no significant differences between the groups. The mean BASFI score at baseline was lower in patients who completed the treatment period than those who did not, although this difference was not significant.

Table 52: Patient demographics and baseline characteristics: Baraliakos, 2013

Characteristics	Completed† (n=16)	Discontinued† (n=10)
Age, years	36.3±7.5	38.4±11.0
Male, n (%)	14 (87.5)	6 (60)
Disease duration, years	13±7.7	15±10.8
HLA positive, n (%)	13 (81.3)	100 (100)
BASDAI (0 to 10)	6.3±0.9	7.0±1.4
BASFI (0 to 10)	5.3±1.9	6.8±1.6
BASMI (0 to 10)	3.9±2.2	3.7±1.3
CRP level, mg/L	20.8±17.7	19.3±16.7
ASDAS-CRP	3.9±0.7	4.3±0.9

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score – C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; HLA, human leucocyte antigen.
†Data are presented as mean ± SD.

3.5.14.3 Results: Baraliakos, 2013

Primary efficacy outcome

By the end of the 7 year treatment period, five patients (31.3%) were in ASAS partial remission and seven patients (43.8%) showed an ASDAS inactive disease status.

Secondary efficacy outcomes

Mean ASDAS, BASDAI, BASFI and BASMI scores improved rapidly with treatment and remained stable over the 7 year study period (Table 53); there were no significant changes observed in these mean values from Year 2 to Year 7. A BASDAI score of <3 units was achieved by 8 patients (50%) after two years and by 11 patients (68.8%) at Year 7. An ASDAS major clinical improvement was achieved in nine patients (56.3%) after 2 years and ten patients (62.5%) after 7 years, while ASDAS moderate disease activity was achieved by eight (50%) and 11 (68.8%) patients after 2 and 7 years, respectively.

Both components of the SF-36 improved significantly (p<0.001) from baseline to Year 7 (Table 53).

Table 53: Secondary efficacy outcomes (absolute changes from baseline to Year 7 in clinical parameters), Baraliakos, 2013

Parameter	Baseline†	Year 7†	p value
BASDAI	6.3±0.9	2.5±2.2	NR
BASFI	5.3±1.9	NR	NR
BASMI	3.9±2.2	NR	NR
CRP level, mg/L	20.8±17.7	NR	NR
ASDAS-CRP	3.9±0.7	1.6±0.8	<0.001
SF-36 MCS	40.6±11.5	46.7±11.2	<0.001
SF-36 PCS	30.8±6.8	43.1±11.5	<0.001

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score – C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; MCS, Mental Component Score; NR, not reported; PCS, Physical Component Score; SF-36 short form 36.
†Data are presented as mean ± SD.

Nine patients (56.3%) achieved an ASAS 40 response after 2 years of etanercept treatment, and 11 patients (68.8%) achieved ASAS 40 after 7 years (Table 54).

Four patients (25%) did not reach a BASDAI of <3 or ASDAS moderate disease activity at the end of the 7 year treatment period. Analysis showed three out of four of these patients had been in ASDAS high disease activity or ASDAS moderate disease at the end of each study year. For these four patients, at the end of Year 7, BASDAI scores ranged between 3.6 and 8.1, ASDAS scores ranged between 2.1 and 3.0, and BASFI scores ranged between 4.4 and 7.8. Despite being elevated prior to treatment, the mean CRP value in this group of patients improved and was normalised during the study period (4.2 mg/L at the end of Year 7).

Table 54: Secondary efficacy outcomes, Baraliakos, 2013

Parameter	Year 2 (n=16) n (%)	Year 7 (n=16) n (%)
ASAS 40	9 (56.3)	11 (68.8)
BASDAI <3 units	8 (50.0)	11 (68.8)
ASDAS (moderate disease activity)	8 (50.0)	11 (68.8)
ASDAS (major improvement)	9 (56.3)	10 (62.5)

Abbreviations: ASDAS-CRP, Ankylosing Spondylitis Disease Activity Score – C reactive protein; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index.

3.5.14.4 Efficacy conclusions: Baraliakos, 2013

The results of this 7 year open label extension study demonstrate that long-term etanercept therapy is efficacious for the treatment of AS. The clinical benefits of etanercept treatment on disease activity, and patient function and mobility (as evidenced by improvements on the BASDAI, BASFI and achievement of ASAS 40) were sustained over the 7 year treatment period following rapid, initial improvement. After 7 years of continuous etanercept therapy, one third of patients were in partial remission. Long term etanercept therapy also had beneficial effects on patients HRQoL which were sustained up to 7 years, as evidenced by significant improvements in the SF-36.

An important observation to note is that in several of the longer-term follow-up extension studies presented in this section, some of the efficacy endpoints reported for etanercept in patients with AS appear to show some numerical fluctuations at certain time points. However, on closer scrutiny, there is evidence which demonstrates that over the course of these longer-term efficacy evaluations of etanercept, a natural fluctuation occurs in the various treatment response measures (e.g. ASAS response rates and improvements in the BASDAI).

In particular, the mean values for a specific outcome may vary at different time-points over the course of the analysis, although the overall trend at study endpoint is one of maintenance of a sustained response. For example, an analysis of data from patients receiving anti-TNF agents for the treatment of AS taken from the Czech national registry showed that the percentage of etanercept patients achieving BASDAI <4 was significantly greater compared with patients receiving infliximab from 3 to 36 months. However, despite this overall trend displayed by the end of the study, at specific time points during the analysis, the etanercept data showed a fluctuation in mean BASDAI values (164).

Based on the data from the Czech study and the long-term five year data from study 907-EU and seven year data from Baraliakos 2013 (122, 163) (9) (presented in sections 3.5.9 and 3.5.14), it can be concluded that when observing efficacy outcomes over a longer term, a random variation in the outcome values may be observed at specific time-points, despite an overall trend of a sustained response provided by etanercept over the total study duration.

Etanercept versus sulfasalazine

3.5.15 Study 402-WW/Braun, 2011

3.5.15.1 Methodology: study 402-WW/Braun, 2011

The source of the efficacy data for study 402-WW is Braun et al, 2011 (14) (NCT00247962). Note that while it is generally accepted that the comparator in this study, sulfasalazine (SSZ), is now known not to be an effective therapy for the axial symptoms of AS, at the time that this trial was conducted, information regarding its effectiveness in improving spinal symptoms was contradictory.

Study 402-WW was a multicentre, randomised, double-blind head-to-head trial to compare the efficacy and safety of etanercept with SSZ in patients with active AS. A summary of the methodology for study 402-WW/Braun, 2011 is provided in Table 55.

Table 55: Methodology study 402-WW/Braun, 2011

Study	402-WW/Braun, 2011 (NCT00247962)
Objectives	To compare the efficacy and safety of etanercept and SSZ in patients with AS.
Study location	The study was conducted at 85 centres in 23 countries in Asia, Australia, Europe, Latin America and the Middle East.
Design details	A four week screening period was followed by a 16 week treatment period during which patients received etanercept or SSZ. Study medications were visually identical to maintain blinding. Patients were assessed at screening and at Weeks 2, 4, 8, 12 and 16.
Interventions	Patients were randomised to receive either: • Etanercept 50 mg administered subcutaneously once weekly, or matching placebo to maintain blinding, or • An initial dose of oral SSZ 0.5 gm, or matching placebo to maintain blinding in Week 1, followed by increasing doses by increments of 0.5 gm each week until the target daily dose of 3 gm was achieved
Key inclusion criteria	• Aged ≥18 years • Active AS, diagnosed according to modified New York criteria for AS. Active disease was diagnosed if the patient had an overall score ≥30 for each of the following: BASDAI score; duration and intensity of morning stiffness (on a 100-mm VAS); and on at least two of three other clinical parameters (BASFI score, VAS scores for nocturnal and total back pain and patients global assessment) • Previously failed treatment with at least one NSAID, taken for at least 3 months at the recommended or tolerated dose
Key exclusion criteria	• Complete spinal ankylosis • Pre-specified abnormal laboratory findings • Significant concurrent medical conditions or other comorbidities • Previous therapy with TNF-α inhibitors or other biologic agents • Use of SSZ within 6 months prior to screening • Use of any corticosteroids within 4 weeks prior to screening • Previous use of intravenous bisphosphonates • Failure to respond adequately to > DMARD • Receiving >1 DMARD or NSAID at baseline • Previous intolerance or lack of response to SSZ
Efficacy outcomes	<u>Primary efficacy endpoint</u> • The proportion of patients achieving ASAS 20 at Week 16

Study	402-WW/Braun, 2011 (NCT00247962)
	Secondary efficacy endpoints • The proportion of ASAS 20 responders at pre-specified assessments up to Week 16 • The proportion of patients achieving ASAS 40 • The proportion of patients achieving 20% improvement in ASAS 5/6 at pre-specified visits up to Week 16 Other outcomes • Patients achieving partial remission, BASMI score, BASDAI score, BASFI score, patient and physician global assessment of disease activity, total and nocturnal back pain scores, spinal mobility (modified Schober's test), tender joint counts (70 joints), swollen joint counts (68 joints), patient's global assessment of peripheral joint arthritis, CRP levels
Populations analysed	Assessment of efficacy was based on the modified ITT population, defined as all randomised patients who received at least one dose of study drug and had at least one assessment post baseline. The safety population included all randomised patients who received at least one dose of study drug.
Statistical analysis	The study assumed at least 90% power based on the enrolment of 540 patients (etanercept, n=360, SSZ, n=180). ASAS 20 response rates were estimated to be 55% and 40% in the etanercept and SSZ groups, respectively. A Fisher's two-sided exact test for superiority was used for the assessment of the primary endpoint with a significance level of 0.05. Descriptive statistics were used for the assessment of all other efficacy endpoints missing data in the modified ITT population was handled using a last observation carried forward approach.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; ITT, intention-to-treat; NSAID, non-steroidal anti-inflammatory drug; SSZ, sulfasalazine; TNF, tumour necrosis factor; VAS, visual analogue scale.

3.5.15.2 Patient population: study 402-WW/Braun, 2011

Patient disposition

A total of 566 patients were enrolled in the study and were randomised in a 2:1 ratio to receive at least one dose of etanercept (n=379) or SSZ (n=187). The mean dosage of SSZ administered was 2.8 gm/day and 166 patients (89%) received 3 gm/day. In the etanercept group, 353 patients completed the study and 26 discontinued prematurely. In the SSZ group, 168 patients completed the study and 19 withdrew prematurely. The most common reason for discontinuation in both treatment arms was AEs.

Patient demographics and baseline characteristics

Baseline demographics and disease characteristics were similar between patient groups. The majority of patients randomised into the study were male (74%) and the mean disease duration was 7.6 years (range 0–45 years). A summary of patient baseline demographics and disease characteristics is provided in Table 56.

Table 56: Patient demographics and baseline characteristics: study 402-WW/Braun, 2011

Characteristic	Etanercept (n=379)	SSZ (n=187)
Age, mean ± SD years	40.7 ± 11.7	40.9 ± 12.2
Disease duration, mean ± SD years	7.5 ± 9.5	8.0 ± 8.9
Male (%)	279 (73.6)	140 (74.9)
White (%)	331 (87.3)	162 (86.6)
BASMI†	3.7 ± 2.25	3.4 ± 2.19
BASDAI†	59.3 ± 16.3	59.1 ± 15.9
BASFI†	55.0 ± 20.2	55.1 ± 20.8
Physician's global assessment of disease activity†	60.1 ± 17.2	60.1 ± 17.2
Patient's global assessment of disease activity†	65.2 ± 18.7	65.6 ± 17.8
Total back pain†	63.1 ± 20.32	61.6 ± 20.34
Nocturnal back pain†	63.4 ± 22.63	63.2 ± 23.13
Spinal mobility on modified Schober's test, mean ± SD cm	3.7 ± 2.3	3.8 ± 2.2
Tender joint count (of 70 joints), mean ± SD	4.6 ± 8.6	4.7 ± 7.2
Swollen joint count (of 68 joints), mean ± SD	1.1 ± 3.0	1.2 ± 2.4
Patients with ≥1 swollen and ≥1 tender joint (%)	117 (30.9)	58 (31.0)
Patients with peripheral arthritis (≥1 swollen or tender joint) (%)	268 (70.7)	138 (73.8)
CRP, mean ± SD mg/L	17.0 ± 20.8	15.5 ± 18.
Prior NSAID‡ use (%)	335 (88.4)	159 (85.0)
Concomitant NSAID‡ use (%)	324 (85.5)	154 (82.4)
Prior corticosteroid use (%)	41 (10.8)	17 (9.1)
Concomitant corticosteroid use (%)	38 (10.0)	20 (10.7)
Prior DMARD use (%)	106 (28.0)	53 (28.3)
Prior SSZ use (%)	67 (17.7)	39 (20.9)

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; SD, standard deviation; SSZ, sulfasalazine.
†Values presented as mean ± SD mm on VAS.
‡Only patients who had received NSAIDs within 28 days of screening were included.

3.5.15.3 Results: study 402-WW/Braun, 2011

Primary efficacy outcome

At Week 16, a significantly greater proportion of etanercept patients achieved an ASAS 20 response compared with SSZ patients (75.9% and 52.9%, respectively; p<0.0001). A significantly greater proportion of patients in the etanercept group achieved an ASAS 20 response at Week 2 compared with the SSZ group (p<0.0001) and this difference was sustained over the 16 weeks of the study.

Secondary efficacy outcomes

In the etanercept group, a significantly greater proportion of etanercept patients achieved ASAS 40 and ASAS 5/6 responses at all time points up to Week 16 compared with the SSZ group ($p<0.0001$). At Week 16, 59.8% and 45.5% of etanercept patients achieved ASAS 40 and ASAS 5/6, respectively, compared with 32.6% and 21.2% of SSZ patients, respectively ($p<0.0001$). At Week 16, a significantly greater proportion of etanercept patients achieved partial remission compared with SSZ patients (33.3% and 15.5%, respectively; $p<0.0001$). At Week 16, etanercept patients achieved a 54.1% improvement in the BASDAI compared with 33.3% of SSZ patients (Table 57).

[redacted]

Patients receiving etanercept showed significantly greater improvements in mobility (BASMI) and disease activity (BASFI) than patients receiving SSZ at all time points. Etanercept was associated with significantly greater improvements versus SSZ in all other secondary outcome measures including total back pain, CRP levels and spinal mobility. A summary of secondary efficacy outcomes is presented in Table 57.

Table 57: Summary of secondary efficacy outcomes: study 402-WW/Braun, 2011

Outcome	Etanercept† (n=379)	SSZ† (n=187)	Tx Difference‡	p value§
BASDAI (mean, % change from baseline)	27.24 [redacted]	39.41 [redacted]	-	[redacted]
BASMI¶	2.78 (2.56, 3.01)	3.21 (2.87, 3.55)	-0.64 (-0.87, -0.40)	<0.0001
BASFI¶	28.72 (26.14, 31.29)	39.35 (35.56, 43.14)	-10.70 (-14.43, -6.97)	<0.0001
Physician's global assessment of disease activity¶	22.31 (20.37, 24.26)	34.56 (31.20, 37.91)	-12.30 (-15.80, -8.80)	<0.0001
Patient's global assessment of disease activity¶	29.57 (27.01, 32.13)	44.87 (40.88, 48.85)	-15.18 (-19.57, -10.79)	<0.0001
Total back pain¶	28.00 (25.38, 30.61)	41.91 (37.87, 45.96)	-14.57 (-19.01, -10.12)	<0.0001
Nocturnal back pain¶	25.42 (22.81, 28.04)	41.19 (37.17, 45.21)	-15.88 (-20.40, -11.35)	<0.0001
Spinal mobility (modified Schober's test)	4.38 (4.14, 4.62)	4.05 (3.73, 4.38)	0.45 (0.17, 0.73)	0.0018
Tender joint count (baseline >0)	3.11 (2.35, 3.88)	4.26 (2.71, 5.81)	-1.31 (-2.40, -0.22)	0.0187
Swollen joint count (baseline >0)	1.33 (0.82, 1.85)	2.82 (0.98, 4.66)	-1.44 (-2.79, -0.09)	0.0371
CRP level, mg/L	6.42 (5.39, 7.44)	13.88 (11.49, 16.28)	-7.86 (-9.84, -5.89)	<0.0001
Patient's assessment of peripheral joint arthritis††	23.89 (20.75, 27.03)	35.59 (30.37, 40.81)	-13.82 (-19.13, -8.50)	<0.0001

Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; SSZ, sulfasalazine.
†Values presented as the mean (95% CI) unless otherwise indicated.
‡Values presented as the least squares mean (95% CI)
§Between group differences determined using analysis of covariance with baseline, treatment and pooled site as predictors.
¶mm on VAS
††Calculated only for patients with >0 tender or swollen joints at baseline.

3.5.15.4 Efficacy conclusion: study 402-WW/Braun, 2011

This randomised, active-controlled trial demonstrates that compared with SSZ, etanercept 50 mg once weekly demonstrated rapid, significant and sustained improvement in measures of physical function, spinal mobility, and in axial symptoms of AS.

Etanercept versus infliximab

3.5.16 Giardina, 2010

3.5.16.1 Methodology: Giardina, 2010

The source of the efficacy data is Giardina et al, 2010 (104). This study was a multicentre, randomised, 2 year open-label, head-to-head trial to compare the efficacy and safety of etanercept and infliximab in patients with active AS. Note that this was an independent study rather than a Pfizer sponsored trial. A summary of the methodology and efficacy outcomes for Giardina, 2010 is provided in Table 58.

Table 58: Methodology: Giardina, 2010

Study	Giardina, 2010
Objectives	To compare the long-term efficacy and safety of etanercept and infliximab in patients with AS.
Study location	Not reported in the publication.
Design details	Efficacy and safety evaluations were performed at Weeks 2, 4, 8, and 12 and then every 3 months up to 2 years.
Interventions	Patients were randomised in a 1:1 ratio to receive either: • Etanercept 50 mg administered subcutaneously once weekly, or • Infliximab 5 mg/kg at Weeks 0, 2, 6 and every 6 weeks for 102 weeks
Key inclusion criteria	• AS diagnosed according to modified New York criteria for AS • Active AS for at least 3 months • Non responder to oral NSAIDs and naïve for DMARDs or other anti-TNF agents • BASDAI >4 • VAS for spinal pain score >4
Key exclusion criteria	• Complete spinal ankylosis
Efficacy outcomes	• Proportion of patients achieving ASAS 20 and ASAS 40 at Weeks 2, 12, 54 and 104 • Improvement in the BASDAI, BASFI and BASMI • Health assessment questionnaire (HAQ) • Spinal mobility (Schober's test) • Spinal pain
Populations analysed	All enrolled, randomised patients were assessed.
Statistical analysis	An unpaired t-test was used to compare mean differences between time points (week 0 vs week 12, week 54 and week 104). The corresponding non-parametric test (Wilcoxon test) was used In the case of skewed distributions (CRP and ESR). The McNemar test was used to compare paired proportions.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C reactive protein; DMARD, disease modifying anti-rheumatic drug; ESR, erythrocyte sedimentation rate; HAQ, health assessment questionnaire; NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor; VAS, visual analogue scale.

3.5.16.2 Patient population: Giardina, 2010

Patient disposition

A total of 50 patients were enrolled in the study and were randomised in a 1:1 ratio to receive at least one dose of etanercept (n=25) or infliximab (n=25). No patients discontinued prematurely.

Patient demographics and baseline characteristics

Baseline demographics and disease characteristics were similar between patient groups with no statistically significant differences observed for any of the reported variables. The average age of patients was 32.2 years (±8.0 years). The majority of patients randomised into the study were male (79%) and the mean disease duration was 15.6 years (±8.7 years). A summary of patient baseline demographics and disease characteristics is provided in Table 59

Table 59: Patient demographics and baseline characteristics: Giardina, 2010

Characteristic	Etanercept (n=25)	Infliximab (n=25)
Age, mean ± SD years	32.6 ± 6.8	31.9 ± 9.2
Disease duration, mean ± SD years	15.7 ± 6.5	15.4 ± 10.6
Male (%)	20 (80)	19 (76)
BASMI score, 0-10, mean ± SD	3.9 ± 1.7	3.7 ± 1.6
BASDAI score, 0-10, mean ± SD	6.6 ± 1.1	6.5 ± 1.2
BASFI score, 0-10, mean ± SD	6.5 ± 1.1	6.1 ± 0.9
Patient's global assessment of disease activity†	6.7 ± 1.4	6.4 ± 1.4
CRP, mean ± SD mg/L	22.9 ± 10.5	25.0 ± 12.1
Chest expansion, mean ± SD cm	3.1 ± 0.9	2.9 ± 0.8
Spinal pain, mean ± SD 0-10	6.3 ± 1.2	6.0 ± 1.5
HLA-B27 positive (%)	24 (96)	23 (92)
ESR, mean ± SD mm/hr	29.6 ± 13.7	32.1 ± 14.6
HAQ, mean ± SD	1.4 ± 0.5	1.5 ± 0.5

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; CRP, C-Reactive protein; ESR, erythrocyte sedimentation rate; HLA, human leukocyte antigen; HAQ, health assessment questionnaire; hr, hour; SD, standard deviation.
†Values presented as 0-10 VAS ± SD.

3.5.16.3 Results: Giardina, 2010

At Week 2, there were more ASAS 20 responders in patients treated with infliximab compared with etanercept patients. However, this difference was not significant and was not maintained at Week 12.

At Week 12, 60% and 43% of patients treated with etanercept and 75% and 55% of patients treated with infliximab were ASAS 20 and ASAS 40 responders, respectively. At Week 12, both groups achieved a reduction in BASDAI and BASFI (3.5 vs 5.6 and 3.5 vs 5.0, both $p<0.005$, for infliximab and etanercept, respectively).

During the observation period, both patient groups had significant improvements from baseline in CRP and ESR levels, BASMI score, spinal pain and HAQ with no differences between the two groups. A significant improvement in spinal mobility, as measured by Schober's test, was observed in both treatment groups.

3.5.16.4 Efficacy conclusion: Giardina, 2010

Both etanercept and infliximab produced rapid, significant, and sustained improvement in multiple clinical and laboratory measures of AS. Although there were some between-treatment differences on some outcomes at 12 weeks, these were not sustained over the course of the study, and no significant differences between etanercept and infliximab were observed at the end of the study. At Week 12, both groups achieved a reduction in BASDAI and BASFI.

In general, improvements with both etanercept and infliximab were observed as early as Week 2 and were sustained until the end of observation at Week 104. Patients were treated continuously throughout the study and showed no reduction in efficacy with either treatment. The results of this study suggest that both etanercept and infliximab are effective therapies for the treatment of AS.

3.6 Phase II RCTs and related open-label extension studies: etanercept for the treatment of AS

One Phase II RCT evaluating the efficacy of etanercept in patients with AS is presented in this section. A brief summary of data from an open-label extension study of this Phase II trial is also presented.

3.6.1 Study 16.0626/Gorman, 2002

3.6.1.1 Methodology: study 16.0626/Gorman, 2002

The source of the efficacy data for study 16.0626 is Gorman et al, 2002 (93). This was a randomised, double-blind, placebo-controlled trial to assess the efficacy of etanercept in patients with AS. A summary of the methodology and efficacy outcomes in study 16.0626/Gorman, 2002 is provided in Table 60.

Table 60: Methodology and efficacy outcomes: study 16.0626/Gorman, 2002

Study	16.0626/Gorman, 2002
Objectives	To evaluate the efficacy of etanercept for the treatment of AS.
Study location	Patients were recruited from rheumatology practices in California, US.
Design details	Patients were randomised to receive either etanercept or placebo for 4 months. Clinical and laboratory assessments were conducted at screening and on Days 1, 28, 56, 84 and 112.
Interventions	Patients were randomised to receive either: • Etanercept 25 mg subcutaneous injection, twice weekly, or • Placebo 25 mg subcutaneous injection, twice weekly
Key inclusion criteria	• Aged ≥18 years • Active AS as defined by the modified New York criteria and despite receiving accepted treatments
Key exclusion criteria	• Any spondylitis other than AS • Complete ankylosis of the spine • History of recurrent infections or cancer • A serious liver, renal, haematologic or neurologic disorder
Permitted concomitant medications	• Treatments for AS providing the doses had not been altered for at least 4 weeks prior to randomisation and remained stable throughout the study • NSAIDs (≤10 mg/day) • Gold injections (≤50 mg/month) • Methotrexate (≤20 mg/week) • SSZ ≤3 g/day)
Efficacy outcomes	<u>Primary efficacy outcome</u> • A 20% or greater improvement in at least three out of five measures of disease activity according to ASAS Working Group recommendations: duration of morning stiffness, nocturnal back pain, BASFI, patient's global assessment and joint swelling score. <u>Secondary efficacy outcomes</u> • Physicians global assessment of disease activity, spinal mobility scores, enthesitis scores, peripheral joint tenderness scores and CRP and ESR levels • [redacted]

Study	16.0626/Gorman, 2002
Populations analysed	Two populations were analysed; one analysis based on the ITT principle and a second analysis which included only patients who completed the study.
Statistical analysis	The study was designed with 80% power to detect response rates of 71% vs 27% in the etanercept and placebo groups, respectively. The target population was 40 patients to allow for the discontinuation of up to two patients per treatment arm. For the main analysis, data from each last patient visit was compared with data from the baseline visit. For continuous measures, the Mann-Whitney test for two independent groups was used to compare mean changes in the two treatment arms. The two-tailed Fisher's exact test was used to compare categorical and ordinal data. All statistical analyses tests were two-sided with a significance level of ≤0.05.
Patient disposition	A total of 40 patients were enrolled in the study; 20 patients in each group. One patient in the etanercept group withdrew (for reasons unrelated to the study) and two patients in the placebo group discontinued because of lack of efficacy.
Patient demographics and baseline characteristics	The majority of patients were male (78%), the mean age was 39 years (range 20–66) and the mean disease duration was 13 years (range 1–40). Patients in the etanercept group had a statistically significant lower mean SF-36 score for physical functioning (p=0.01) and lower mean haemoglobin levels (p=0.04) compared with the placebo group which may have indicated the presence of more severe disease in the etanercept group.
Efficacy Results	<u>Primary efficacy outcome</u> • In the ITT population, there was a rapid and significantly greater response rate in the etanercept group compared with the placebo group at 4 months (80% and 30%, respectively; p=0.004) which was sustained throughout the study <u>Secondary efficacy outcomes</u> • At 4 months, etanercept patients had significantly greater improvements in all secondary outcome measures with the exception of peripheral joint tenderness scores, modified Schober's test scores and occiput to wall measurements

	Etanercept† (n=20)		Placebo† (n=20)		
Outcome	**Baseline**	**Month 4**	**Baseline**	**Month 4**	**p value‡**
Physicians global assessment	54.5±19.6	23.0±10.6	48.0±16.4	55.5±22.7	<0.001
Spinal mobility measures					
Chest expansion, cm	2.6±1.6	3.5±1.9	3.1±1.7	2.9±1.7	0.006
Modified Schober's test, cm	12.5±1.5	13.4±1.6	13.5±1.5	13.4±1.5	0.26
Occiput to wall§, cm	5.7±7.9	4.7±7.6	2.0±1.5	2.7±4.4	0.11
Modified Newcastle Enthesitis Index¶	4.5±8.4	0.0±3.0	3.0±7.9	1.5±8.0	0.001
Peripheral joint tenderness††	3.5±10.5	1.0±2.5	2.5±11.8	2.5±23.1	0.07
ESR level, mm/hr	34.5±23.1	8.5±12.8	20.0±16.3	16.5±18.7	<0.001
CRP level, mg/dL	2.0±1.8	0.7±1.1	1.5±1.2	2.0±2.8	0.003

†Values are median ± SD unless otherwise stated.
‡p values are the difference in change from baseline between the two groups.
§Value is the mean ± SD.
¶17 entheses were scored on a four-point pain scale (0=none; 1=mild; 2=moderate; 3=eliciting a wince or withdrawal).
††66 joints were scored on a four-point tenderness scale (from 0=none to 3=severe)

Study	16.0626/Gorman, 2002
	• [redacted]
Efficacy conclusion	The results of the placebo-controlled trial demonstrate that etanercept 25 mg for 4 months is associated with rapid, significant and sustained improvement in the clinical symptoms of AS. Patients on etanercept demonstrated significant improvement in enthesitis compared with patients on placebo.

Abbreviations: AS, ankylosing spondylitis; ASAS, Assessment of Spondyloarthritis international Society; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; ITT, intention to treat; NSAID, non-steroidal anti-inflammatory drug; SSZ, sulfasalazine.

3.6.2 Study: Davis, 2004

3.6.2.1 Methodology: Davis, 2004

The source of the efficacy data for Davis 2004 is Davis et al, 2004 (117) and the clinical study report (103). Davis, 2004 was an open-label extension of study 16.0626. Thirty eight of the 40 patients who had completed study 16.0626 enrolled in the extension study and received etanercept 25mg, subcutaneously twice weekly for 6 months.

The primary efficacy outcome was achievement of ASAS 20. Other efficacy outcome measures included ASAS 50 and ASAS 70 responses; morning stiffness duration; nocturnal back pain, BASFI scores, patient global assessment, swollen joint count, spinal mobility, joint pain and Enthesopathy Index. Clinical and laboratory assessments were performed at Weeks 20, 28, and 40.

3.6.2.2 Efficacy outcomes: Study - Davis, 2004

A total of 36 patients completed the extension study.

[redacted]

Of the 17 patients who had previously received etanercept in study 16.0626, 16 patients (94%) achieved ASAS 20 after 6 months. At Week 40, ASAS 20 was achieved in 16/19 patients (84%) who had previously received placebo in study 16.0626 (Table 61).

Table 61: ASAS 20 results at open-label baseline and Weeks 20, 28 and 40

	Etanercept† N=19	Placebo/Etanercept‡ N=19	Total N=38
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ASAS, Assessment of Spondyloarthritis international Society.
†17 of 19 patients completed the open-label extension through Week 40 in the etanercept group. The two patients who did not complete Week 40 are omitted from the summary at Week 40.
‡Placebo/etanercept patients began receiving etanercept after the Week 16 evaluation.

At Week 40, ASAS 50 and ASAS 70 were achieved in 81% and 42%, respectively of patients continuing etanercept. In the original placebo group, 84% and 47% of patients achieved ASAS 50 and ASAS 70, respectively (Table 62).

Table 62: ASAS 50 and ASAS 70 results at open-label baseline and Weeks 20, 28 and 40

	Etanercept† N=19	Placebo/Etanercept‡ N=19	Total N=38
ASAS50			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
ASAS70			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ASAS, Assessment of Spondyloarthritis International Society.
†17 of 19 patients completed the open-label extension through Week 40 in the etanercept group. The two patients who did not complete Week 40 are omitted from the summary at Week 40.
‡Placebo/etanercept patients began receiving etanercept after the Week 16 evaluation.

At Week 40, both treatment groups showed mean improvements from baseline in duration of morning stiffness, nocturnal back pain, BASFI scores, and patient global assessment (Table 63).

Table 63: Additional efficacy results at Week 40

	Etanercept‡ N=19	Placebo/Etanercept† N=19	Total N=38
Morning stiffness			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
Nocturnal back pain			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
BASFI			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
Patient global assessment			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index.

†17 of 19 patients completed the open-label extension through Week 40 in the etanercept group. The two patients who did not complete Week 40 are omitted from the summary at Week 40.
‡Placebo/etanercept patients began receiving etanercept after the Week 16 evaluation.

Patients continuing etanercept showed 59.6%, 26.2% and 48.1% improvements from baseline in chest expansion, the modified Schober's test and for the occiput to wall measurement, respectively. In the original placebo group, mean improvements from baseline in chest expansion, modified Schober's test and the occiput to wall measurement were 48.8%, 7.8% and 24.8%, respectively.

During the open-label extension period, patients in the original etanercept group maintained the treatment benefits observed in joint pain and scores of the Enthesopathy Index and placebo patients achieved similar benefits after treatment with etanercept.

At the end of the extension period, 66% of the total patient cohort had either discontinued at least one of their AS medications or reduced the dosage.

3.6.2.3 Efficacy conclusion: study: Davis, 2004

This 6 month open-label extension study demonstrated that etanercept treatment was associated with significant improvements in the clinical symptoms of AS. Patients previously treated with placebo showed improvement from baseline of the open-label extension to Week 40 comparable to that observed in patients who were treated with etanercept throughout study 16.0626.

3.7 Safety profile of etanercept

SAFETY

The safety profile of etanercept is well established with over 15 years of global post-marketing experience (33). There is considerable clinical evidence from RCTs and longer term follow-up studies which demonstrate that etanercept has an established long-term safety profile and is generally well tolerated with a favourable safety profile across the spectrum of AxSpA.

- Studies reported that adverse reactions were largely mild to moderate in intensity and usually resolved without discontinuation of treatment
- Data from longer term extension studies demonstrated that etanercept was generally well tolerated for up to 7 years of treatment with a low incidence of discontinuation due to AEs
- Etanercept discontinuation rates due to adverse events were similar to placebo in AS patients (10).

Etanercept for the treatment of nr-axSpA

- Etanercept treatment was well tolerated in this patient population over 48 weeks of treatment. There were no new safety signals reported during the 1031 trial.
- **There were no serious infections reported for ETN during double blind (week 12), or in the open-label phase (week 24).**

[redacted]

Etanercept for the treatment of AS

- The most common adverse events reported during clinical trials of etanercept in patients with AS were injection site reaction (8–37% of patients), upper respiratory tract infection, (8–65% of patients) and headache (7–28% of patients)
- There were no significant differences in the rates or incidence of adverse events reported in patients receiving etanercept compared with patients receiving placebo
- Serious adverse events (SAEs) generally occurred as isolated events and were not considered to be related to etanercept therapy
    - Discontinuation due to an SAE occurred in 1.6–11.5% of patients treated with etanercept in trials
- The majority of patients presenting with uveitis had a previous history of the condition while a new case of inflammatory bowel disease lead to discontinuation in one study (120)
        - Uveitis events did not result in discontinuation from any of the studies

Safety profile of etanercept based on pooled analyses

- Data from pooled analyses indicated that occurrences of uveitis in patients receiving etanercept for active AS tended to be mild to moderate in severity (18)
- Etanercept has an established safety profile in AS patients with uveitis or at risk for uveitis; in pooled analysis of double-blind RCTs and open-label extensions, reported rates of uveitis were low in patients treated with etanercept (18)
- A pooled analysis of etanercept-treated patients with active AS showed that etanercept was associated with low rates of serious infections, opportunistic infections, malignancies, non-malignant skin cancers, and inflammatory bowel disease (165)

3.7.1 Introduction

The safety profile of etanercept has been characterised by more than 15 years of post-marketing experience (33). AS was approved by the European Medicines Agency (EMA) in 2004 and is approved for use by the US Food and Drug Administration (FDA). Safety data based on the experience of 1,653 patients from RCTs (966 patients) and open-label extension trials (687 patients) demonstrates that etanercept is generally well tolerated with an acceptable safety profile. The most commonly reported adverse reactions are injection site reactions (e.g. pain, swelling and itching), infections (e.g. bronchitis, upper respiratory tract infection) and fever. An integrated pooled analysis of AEs examined the collective safety of etanercept in 49 RCTs in the database as of May 2006 across all approved indications, including AS (total patients N=13,977). Rates of serious infections in all etanercept trials were 3.01 per 100 patients-years and similar in both groups: AS patients (N=695) and all patients (N=13,877) (6). Exposure-adjusted serious infection rates in AS patients were similar between etanercept and placebo in a 192-week study (118). Other more serious undesirable effects have been reported with the use of etanercept including life-threatening infections and tumours of the skin, lungs and lymph glands (8). However, life threatening infections such as cellulitis and sepsis are uncommon in patients receiving etanercept, affecting ≥1/1000 to <1/100 patients while myocobacterial infections and legionella are rare (i.e. affecting ≥1/10,000 to <1/1000 patients). While the occurrence of malignancies in patients receiving etanercept therapy is uncommon (i.e. affecting ≥1/1000 to <1/100 patients) with respect to non-melanoma skin cancers or rare (i.e. affecting ≥1/10,000 to <1/1000 patients) with respect to lymphoma and melanoma, caution is advised before commencing therapy in patients with a history of malignancy or when considering continuing treatment in patients who develop a malignancy (8).

Etanercept therapy is contraindicated in patients with sepsis or the risk of sepsis and treatment should not be initiated in patients with infection. Patients should be closely monitored for signs and symptoms of infection whilst taking etanercept and treatment should be discontinued if serious infection develops (8). Reactivation of latent tuberculosis (TB) or new TB infections have also been reported in etanercept patients. TB incidence with etanercept is low, and the BSR registry found that rates of TB in Rheumatoid Arthritis patients receiving anti-TNF alphas in the UK was 40 per 100 000 person-years (166, 167).

A small retrospective study of 48 patients with positive purified protein derivative (PPD) tests for TB who were treated with etanercept 25 mg twice weekly reported that none of the study

population developed active TB while receiving etanercept for an average of 17 months (15% of the study population were receiving etanercept for the treatment of AS) (168). In addition, studies have reported that the risk of TB reactivation is lower in patients receiving etanercept for the treatment of inflammatory diseases including AS and RA than in those receiving other anti-TNF therapies (166, 167).

In addition to the spinal symptoms of AS, many patients will also have auto-immune disorders (AuIDs) related to the disease. The most commonly occurring AuID in patients with AS is acute anterior uveitis which occurs in approximately 25–30% of patients (35). Other AuIDs associated with AS include psoriasis and inflammatory bowel disease (IBD) which affects 10–25% of patients and 5–10% of patients, respectively (35). The occurrence of AuIDs was recorded in the safety results for the majority of the RCTs and open-label follow-up studies discussed in section 3.3 to section 3.6. Therefore, the frequencies of any reported AuIDs of disease are presented in section 3.7.4.

An adverse event is “any untoward medical occurrence in a patient or clinical–trial subject administered a medical product and which does not necessarily have a causal relationship with the treatment (169). It is not always easy to establish precise causality of AEs, i.e. whether due to disease, as a result of a received treatment or due to other causes. In clinical trials all AEs are recorded. An adverse reaction is a “noxious and unintended” response to a medical product where the causality is” a reasonable possibility (170, 171).

Safety data for the clinical trials and the longer term follow-up studies summarised in sections 3.3–3.6 is presented in the following sections together with the results from three pooled analyses.

3.7.2 Adverse events

Etanercept for the treatment of nr-axSpA

In the 1031 clinical trial of etanercept in patients with nr-axSpA, etanercept therapy over 48 weeks was well tolerated in this population. The frequency and type of AEs reported by patients receiving etanercept in study 1031 were consistent with the safety profile of etanercept reported in clinical trials of etanercept in AS. During the double-blind period, there were no clinically meaningful differences in AEs between patients receiving etanercept and patients receiving placebo; AEs were reported in 57% of etanercept patients versus 45% of placebo patients.
[redacted]
(Table 64). In the double blind period of the 1031 trial, three patients receiving etanercept discontinued treatment due to an AE versus one placebo patient (Table 64). Auto-immune disorders are discussed in section 3.7.4.

[redacted]
64[redacted]
64[redacted]

[redacted]

Etanercept for the treatment of a mixed early AxSpA population

In the ESTHER trial of a mixed early AxSpA population, upper respiratory tract infection was the most frequently reported AE in patients receiving etanercept (details not provided in the publication) (15).

Etanercept for the treatment of AS

The majority of adverse events (AEs) reported during RCTs and related extension studies of etanercept in patients with AS were of moderate or mild intensity. Data from longer term extension studies demonstrated that etanercept was generally well tolerated for up to 7 years of treatment with a low incidence of discontinuation due to AEs. In the placebo-controlled trials, there were no significant differences in the rates or incidence of AEs in patients receiving etanercept compared with patients receiving placebo. In all studies, the most common adverse events (AEs) reported in patients receiving etanercept were injection site reaction (8-37% of patients), upper respiratory tract infection (8–65% of patients), headache (7–28% of patients) and diarrhoea (12–18% of patients). In placebo-controlled trials, the proportion of patients experiencing injection site reaction was greater in patients receiving etanercept compared with patients receiving placebo (8–33% vs 9–15% of patients, respectively). Table 64 lists the most commonly reported AEs that occurred in ≥5% of patients and the number of discontinuations due to AEs in AS patients treated with etanercept in clinical trials.

Withdrawals from treatment due to an AE were reported in 8/16 studies (9, 10, 12, 14, 104, 118-120, 163). The proportion of discontinuations from etanercept treatment as a result of an AE reported in all included studies ranged from 1.4–18.5% of patients. Only two trials reported injection site reaction as a reason for discontinuation from the study (118, 119). Other reasons for discontinuation included rash and infection.

Grade 3/4 abnormal laboratory results were reported in six trials, none of which lead to an AE resulting in discontinuation of etanercept (10-12, 119, 120, 122). Incidences of patients developing antibodies to etanercept were reported in 3/15 trials: study 16.0037, n=3 (10); study 314-EU, n=4 (12); and study 16.0626 (93). In study 16.0037 and study 314-EU antibodies to etanercept were reported as non-neutralising.

Table 64: Summary of most commonly reported AEs and discontinuations due to AEs in patients receiving etanercept in clinical trials

Study	Study design	Population (N)	Most common AEs occurring in ≥5% of patients N (%)	Discontinuation due to AEs (N)
nr-axSpA				
1031 (68, 141)	12 week, double blind, placebo-controlled RCT with a 48 week open-label extension period	**Double blind period, Week 12:** ETN: n=111 Placebo: n=113 **Open-label period, Week 48:** n=208	**Double blind period: etanercept** [redacted]**Double blind period: placebo** [redacted]**Open-label period, Week 48:** • [redacted]	**Double blind period, Week 12: etanercept** [redacted]**Double blind period, Week 12: placebo** [redacted]**Open-label period, Week 48:** • [redacted]
Early AxSpA				
ESTHER/Song, 2011 (15)	Open-label RCT, 48 weeks	ETN: n=40 SSZ: n=36	• Upper respiratory tract infection (N and treatment group not specified in publication)	• Not reported
AS				
314-EU/van der Heijde, 2006 (12)	Double-blind, placebo-controlled, non-inferiority RCT, 12 weeks	ETN 50 mg: n=155 ETN 25 mg: n=150 Placebo: n=51	• Injection site reaction ○ ETN 50 mg, 32 (20.7); ETN 25mg, 34 (22.7) ○ Placebo, 6 (11.8) • Upper respiratory tract infection ○ ETN 50 mg, 12 (7.7); ETN 25mg, 12 (8.0) ○ Placebo, 7 (13.7) • Back pain ○ ETN 50 mg, 1 (0.6); ETN 25mg, 1 (0.7) ○ Placebo, 3 (5.9)	• ETN 50 mg (n=6). AEs not specified. • ETN 25 mg (n=8). AEs not specified. • Placebo (n=0)
16.0037/Davis, 2003 (10)	Double-blind,	ETN: n=	**Etanercept:** • Injection site reaction, 41 (30)	**Etanercept:** • GI haemorrhage secondary to haemorrhoids (n=1)

	placebo-controlled RCT, 24 weeks	138 Placebo: n=139	• Upper respiratory tract infection, 28 (20) • Headache, 19 (14) • Accidental injury, 17 (12) **Placebo** • Injection site reaction, 13 (9) • Upper respiratory tract infection, 16 (12) • Headache, 16 (12) • Accidental injury, 6 (4)	• Ileitis secondary to Crohn's disease (n=1)
16.0040/Davis, 2005 (119)	72 week, open-label extension study of 16.0037	n=257	• Injection site reaction, 53 (21) • Headache, 39 (15) • Diarrhoea, 32 (12) • Abdominal pain, 30 (12) • Accidental injury, 27 (11)	• Injection site reaction (n=1) • Dizziness, shortness of breath and increased heart rate (n=1) • Dermatitis and urticaria (n=1) • Rash (n=2) • Liver steatosis (n=1)
Davis, 2008 (118)	168 week, second open-label extension study of 16.0037	n=257	• Upper respiratory tract infection, (45) • Injection site reaction, 57 (22.2) • Headache, 52 (20.2) • Diarrhoea, 45 (17.5) • Sinusitis, 41 (16) • Flu syndrome, 39 (15) • Pain, 37 (14.4) • Bronchitis, 28 (11)	• Infection event (n=4) • Injection site reaction (n=1) • Rash (n=2)
311-EU/Calin, 2004 (11)	Double-blind, placebo-controlled RCT, 12 weeks	ETN: n=45 Placebo: n=39	**Etanercept** • Injection site reaction, 15 (33) • Headache, 6 (13) • Nausea, 3 (7) **Placebo** • Injection site reaction, 6 (15) • Headache, 4 (10) • Nausea, 4 (10)	• None
312-EU/Dijkmans, 2009 (120)	96 week, open-label extension study of 311-	n=81	• Upper respiratory tract infection, 43 (53) • Injection site reaction, 30 (37) • Headache, 16 (20) • Abdominal pain, 14 (17) • Diarrhoea, 12 (15)	• n=15. AEs not specified.

	EU			
907-EU/Martin-Mola, 2010 (163)	156 week, second open-label extension study of 312-EU	n=59	• Upper respiratory tract infection, 16 (27.1) • Flu syndrome, 16 (27.1) • Accidental injury, 12 (20.3) • Hypertension, 11 (18.6) • Diarrhoea, 9 (15.3) • Injection site reaction, 7 (11.9)	• Discontinuation from any AE (n=7) • Discontinuation from infection event (n=2)
SPINE/Dougados, 2011 (13)	Double-blind, placebo-controlled RCT, 12 weeks	ETN: n=39 Placebo: n=43	**Etanercept** • Injection site reaction, 3 (8)	• Not reported
Dougados, 2012 (121)	12 week, open-label extension study of the SPINE trial	n=77	• Nasopharyngitis, 4 (5.2) • Injection site reaction, 2 (5.1)	• None
Brandt, 2003 (89)	6 week placebo-controlled RCT followed by 24 week observational phase	ETN: n=14 Placebo: n=16	**Etanercept:** • Injection site reaction, 2 (14) • Upper respiratory tract infection, 6 (43) **Placebo:** • Upper respiratory tract infection, 6 (37.5)	• None
Brandt, 2005 (116)	54 week, open-label, observational	n=26	• Upper respiratory tract infection, 17 (65) • Bronchitis, 5 (19) • Diarrhoea, 4 (15)	• None

	extension study of Brandt, 2003			
Baraliakos,2013 (9)	7 year, second, open-label, observational extension study of Brandt, 2003	n=26	• Not reported	• Reactivation of Crohn's disease (n=1) • Development of Crohn's disease (n=1) • Heart discomfort (n=1)
402-WW/Braun, 2011 (14)	Double-blind, head-to-head RCT, 16 week	ETN: n=379 SSZ: n=187	**Etanercept:** • Injection site reaction, 41 (10.8) • Upper respiratory tract infection, 31 (8.2) • Headache, 29 (7.7) • Nausea, 25 (6.6) **SSZ:** • Upper respiratory tract infection, 17 (9.1) • Headache, 21 (11.2) • Nausea, 18 (9.6)	• ETN, n=15 (headache, nausea, rash) • SSZ, n=12 (headache, nausea, rash)
Giardina, 2010 (104)	2 year, open-label, head-to-head RCT	ETN: n=25 Infliximab: n=25	**Etanercept:** • Injection site reaction, 5 (25) • Headache, 7 (28) • Tachycardia, 8 (32) • Arthralgia, 3 (12) • Hypertension, 2 (8) **Infliximab:** • Headache, 8 (32) • Tachycardia, 12 (48) • Hypertension, 4 (16) • Arthralgia, 4 (16) • Diarrhoea, 2 (8)	• None
16.0626/Gor	Double-	ETN:	**Etanercept:**	• None

man, 2002 (93)	blind, placebo-controlled, Phase II RCT, 4 months	n=20 Placebo: n=20	• Injection site reaction, 5 (25) • Upper respiratory tract infection, 10 (50) • Diarrhoea, 3 (15) **Placebo:** • Upper respiratory tract infection, 12 (60)	
Davis, 2004 (103, 117)	6 month, open-label extension study of 16.0626	n=38	**Placebo/Etanercept:** • Injection site reaction, 3 (16) • Upper respiratory tract infection, 3 (16) • Rash, 3 (16) • Diarrhoea, 2 (11) • Headache, 2 (11) **Etanercept:** • Upper respiratory tract infection, 4 (21) • Headache, 3 (16) • Fever, 3 (16) • Diarrhoea, 2 (11)	• None

Abbreviations: AE, adverse event; AS, ankylosing spondylitis; nr-axSpA, non-radiographic axial spondyloarthritis; ETN, etanercept; GI, gastrointestinal; SSZ, sulfasalazine.

3.7.3 Serious adverse events

Etanercept for the treatment of nr-axSpA

There were no serious infections reported for ETN during double blind (week 12), or in the open-label phase (week 24). Serious adverse events (SAEs) were reported in 2 patients receiving etanercept during the double-blind period of the 1031 trial.

[redacted]

Etanercept for the treatment of a mixed early AxSpA population

In the ESTHER trial, SAEs were reported in 3 patients receiving etanercept, only one of which was considered as potentially related to treatment (details not provided in the publication) (15).

Etanercept for the treatment of AS

The majority of SAEs reported in clinical trials of etanercept in AS patients occurred in single patients or were considered in the opinion of the investigator to be unrelated to etanercept therapy. Five of the 15 etanercept trials summarised in section 3.5 and section 3.6 reported patient discontinuations due to SAEs (9, 10, 116, 119, 163) with discontinuation occurring in 1.6–11.5% of patients. These included depression, fever related to injection site reaction, ulcerative colitis, hepatitis B, lung carcinoma and Crohn's disease (Table 65). SAEs which did not result in discontinuation of therapy included migraine, colitis, arthritis, ptosis, mild to moderate neutropenia, acute myocardial infarction and infectious diarrhoea.

Table 65: Summary of SAEs resulting in discontinuations in patients with AS receiving etanercept in clinical trials

Study	Study Design	Population (N)	Discontinuation due to SAE (N)
16.0037/Davis, 2003 (10)	Double-blind, placebo-controlled RCT, 24 weeks	n=138	• fever related to injection site (n=1) • ulcerative colitis (n=1) • intestinal obstruction (n=1) • bone fracture (n=2)
16.0040/Davis, 2005 (119)	72 week, open-label extension study of 16.0037	n=257	• hepatitis B (n=1) - considered unrelated to treatment • Crohn's disease (n=1) - considered unrelated to treatment • asthmatic bronchitis (n=1) - considered unrelated to treatment • worsening depression (n=1) - considered unrelated to treatment
907-EU/Martin-Mola, 2010 (163)	60 week, open-label extension study of 312-EU	n=59	• bilateral pneumonia and sepsis - considered possibly related to treatment (n=1)
Brandt, 2005 (116)	54 week, open-label, observational extension study of Brandt, 2003	n=26	• Crohn's disease – considered possibly related to treatment (n=1)

Baraliakos, 2013 (9)	7 year, second, open-label, observational extension study of Brandt, 2003	n=26	• reactivation of Crohn's disease (n=1) • development of Crohn's disease (n=1) • lung carcinoma (n=1)

Abbreviations: AS, ankylosing spondylitis; SAE, serious adverse event.

3.7.4 Other clinically relevant AEs and auto-immune disorders

Etanercept for the treatment of nr-axSpA

[redacted]

Etanercept for the treatment of AS

TB was reported in 3/16 trials in patients receiving etanercept for AS. One patient in study 16.0040 had a positive PPD test for TB but no symptoms of active disease (119). Fifteen patients in study 314-EU received prophylactic treatment for TB, but none developed active disease (12). One patient in Davis, 2008 had a TB event following 2.8 years of etanercept treatment and subsequently discontinued treatment (118). Two cases of herpes zoster were reported in patients enrolled in study 16.0040/Davis, 2005 (119).

Four trials reported malignancies occurring in patients receiving etanercept. Study 312-EU reported malignancies in 4 (5%) of patients, none of which resulted in discontinuation of the study (120). Carcinomas considered to be related to treatment and resulting in withdrawal from the trial were reported in 3 patients in study 907-EU (163). One patient in the SPINE trial withdrew due to a lung neoplasm (13) while one patient in the 7-year Baraliakos study discontinued due to lung carcinoma (9).

Auto-immune disorders were reported in 9/16 etanercept trials in patients with AS. Uveitis or iritis and IBD flares were reported in nine studies (Table 66). With the exception of studies 907-EU, Davis, 2008 and Baraliakos, 2013, the majority of patients presenting with uveitis/iritis had a previous history of the condition. There were no discontinuations due to uveitis/iritis reported in any of the studies. Discontinuation of etanercept due to a new case of IBD was reported in one study (120). Other extra-articular disease manifestations (EAMs) reported in clinical trials of etanercept in patients with AS included psoriasis and balanitis.

Table 66: EAMs occurring in AS patients receiving etanercept

Study	Study Design	Extra-articular disease manifestation
16.0037/Davis, 2003 (10)	Double-blind, placebo-controlled RCT, 24 weeks	• inflammatory bowel disease (n=3) • uveitis or iritis (n=3)
16.0040/Davis, 2005 (119)	72 week, open-label extension study of 16.0037	• uveitis or iritis (n=26) • colitis (n=5) • ileitis (n=2) • psoriasis (n=1) • balanitis (n=1)
Davis, 2008 (118)	168 week, second open-label	• uveitis (flares, n=62; new incidence,

	extension study of 16.0037	n=10) • inflammatory bowel disease (n=23)
312-EU/Dijkmans, 2009 (120)	96 week, open-label extension study of 311-EU	• inflammatory bowel disease (n= 4) • uveitis (n=10)
907-EU/Martin-Mola, 2010 (163)	156 week, second open-label extension study of 311-EU	• Crohn's disease (n=2) • uveitis (n=10) • iritis (n=3)
402-WW/Braun, 2011 (14)	Double-blind, active comparator RCT, 16 weeks	• uveitis or iritis (n=10) • colitis (n=1) – considered not related to study drug
Giardina, 2010 (104)	2 year, open-label, head-to-head RCT	• uveitis (n=2)
Davis, 2004 (103, 117)	6 month, open-label extension study of 16.0626	• uveitis (n=2)
Baraliakos,2013 (9)	7 year, second, open-label, observational extension study of Brandt, 2003	• recurrent uveitis (n=7/16)

Abbreviations: RCT, randomised controlled trial.

3.7.5 Mortality

Etanercept for the treatment of nr-axSpA

[redacted]

Etanercept for the treatment of a mixed early AxSpA population

No deaths were reported in patients enrolled in the ESTHER trial.

Etanercept for the treatment of AS

Patient death was reported in 3/16 trials of etanercept in patients with AS. Two patients died during study 907-EU. The first patient had chronic lung disease and had a myocardial infarction followed by cardiac arrest, none of which were considered to be related to treatment. The second patient had sepsis complicated by multi-organ failure (considered possibly related to treatment) and myocardial infarction and cardiogenic shock (possibly not related to treatment) (163). One patient death due to accidental injury was reported in study 314-EU/van der Heijde, 2006 (12). One patient in the 7-year Baraliakos study died, possibly due to a heart attack (9).

3.7.6 Safety data from pooled analyses

3.7.6.1 Study: Sieper, 2010

Methodology and safety outcomes: Sieper, 2010

The source of the safety data for Sieper, 2010 is Sieper et al, (18). This study pools safety data across eight etanercept trials; Gorman et al, 2002 (93), Davis et al, 2003 (10), Calin et al, 2004 (11), van der Heijde et al, 2006 (12), Braun et al, 2008 (ASCEND trial) (127), Davis et al, 2008 (118), Dijkmans et al, 2009 (120) and Martin Mola et al, 2008 (172). Note that the Braun et al, 2008, and Martin-Mola et al, 2008, studies were conference abstracts which have since been superseded by full publications.

A summary of the methodology and safety results for Sieper, 2010 is provided in Table 67.

Table 67: Methodology and safety outcomes: Sieper, 2010

Study	Sieper, 2010
Summary	An observational, retrospective, analysis of the safety of etanercept in patients with active AS. Eight studies were examined for reports of uveitis: • Four placebo-controlled trials (10-12, 93) • One active-controlled trial (127) • Three open-label trials (118, 120) (172)
Objectives	To assess the incidence of uveitis (including iritis and iridocyclitis) in clinical trials of etanercept in patients with active AS.
Study location	The trials were conducted in the USA, Europe, Latin America and Asia.
Design details	Eight clinical trials of etanercept in patients with active AS were examined for reports of uveitis: • Four placebo-controlled trials (3-6 month duration) • One active-controlled trial (4 month duration) • Three open-label trials (6-60 month duration)
Interventions	In the placebo controlled trials, patients were randomised to receive: • Etanercept 25 mg subcutaneous injection, twice weekly, or • Etanercept 50 mg subcutaneous injection, once weekly, or • Placebo In the active comparator trials, patients were randomised to receive: • Etanercept 50 mg subcutaneous injection, once weekly, or • SSZ up to 3 g orally, per day In the open-label trials, patients received: • Etanercept 25 mg subcutaneous injection, twice weekly
Key inclusion criteria	Enrolment into one of the eight etanercept trials (see the full publications of each trial for more detail on inclusion criteria).
Permitted concomitant medications	• SSZ • Hydroxychloroquine • Methotrexate
Safety outcomes	The key safety outcome of this analysis was to assess the incidence of uveitis (including iritis and iridocyclitis) in clinical trials of etanercept. Adverse events in the clinical trials were coded using the Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART). Study data were filtered by prespecified TEAEs using the COSTART preferred terms ‘uveitis’, ‘iritis’ or ‘iridocyclitis’, then included in the analysis. Events were classified as new onset in patients with no history, or flare in patients with history/prior event during trials. Patients receiving placebo who had an event during the double-blind study were included in the placebo group; if they had an event after switching to etanercept they were included in the etanercept group. Severity (mild/moderate/severe), duration of symptoms (where available) and treatment were recorded.
Populations analysed	All randomised patients who received at least one dose of treatment with etanercept, SSZ or placebo.
Patient disposition	A total of 1,323 patients were included in the analysis.
Statistical analysis	Rates of uveitis events per 100 patient years were calculated for each arm of the double-blind studies. Pooled rates for etanercept and placebo were also calculated. A long-term rate per 100 patient years was calculated for etanercept

Study	Sieper, 2010
	which included all events occurring in both the double-blind and open-label extension studies. For all studies, between-group differences (including confidence intervals (CIs) in the rates of uveitis) were calculated. Baseline demographics and disease characteristics for all patients included in this analysis were pooled and calculated using descriptive statistics. Between-group differences with CIs were calculated for each study and for the pooled population.
Patient demographics and baseline characteristics	Baseline demographics were similar between treatment groups in terms of age, sex, disease duration, parameters of clinical disease and HLA-B27 status.
Safety results	<u>Incidence of uveitis</u> • In placebo-controlled trials, rates of uveitis events with etanercept were lower than placebo • In the active comparator trial, rates of uveitis events with etanercept were similar to SSZ • Mean disease duration was longer in patients who reported uveitis events than in their respective treatment groups • The subgroups of patients who experienced an event were 91–100% HLA-B27 positive compared with the 79–82% seen in the pooled groups.

<u>Incidence and rates of uveitis treatment-emergent events, pooled results</u>

	Etanercept DB	Etanercept AC	Etanercept DB + OL	SSZ	Placebo
Number of patients	508	379	1,074	187	249
New events	0	6	21	3	1
Flares	9	5†	55	4§	10
Total events	12	12‡	136	8	16
Exposure patient-years	139.1	111.9	1,136.9	54.3	83.0
Rate/100 patient-years (95% CI)	8.6 (4.5-14.2)	10.7 (15.5-17.6)	12.0 (10.0-14.1)	14.7 (6.4-26.5)	19.3 (11.0-29.8)
p value	0.031	0.486	-	-	-

Abbreviations: AC, active comparator; CI, confidence interval; DB, double-blind; OL, open-label; SSZ, sulfasalazine.
†Includes one patient with an inflammatory eye event with onset during the study and subsequent flares.
‡Active comparator trial data. Includes two patients with an event during screening and one patient with an event during the post-study follow up.
§Includes one patient with an event during screening and an event after the study.

- The overall uveitis rate for etanercept was lower than that for placebo, but similar to that for SSZ
- Of the total 1,323 patients, 78 reported at least one uveitis event during the studies; 25 experienced new events and 69 reported one or more flares
- In double-blind, placebo-controlled trials, the uveitis rate with etanercept was lower than with placebo (8.6 and 19.3 per 100 patient years, respectively; p=0.03)
- No uveitis event led to a study discontinuation or was considered to be a serious adverse event by the study investigator
- Events were mostly (96%) mild to moderate in severity and resolved with no treatment or topical therapy

Study	Sieper, 2010
	• The objective of the open-label studies was to determine the long-term uveitis rate in patients receiving etanercept for between 6 and 60 months • In the pooled, open-label analysis, 1,074 subjects received etanercept for a total exposure of 1,136.9 patient years. Seventy-six patients receiving etanercept reported 136 uveitis events resulting in a rate of 12.0 events per 100 patient years • The long-term pooled uveitis rate (per 100 patient years) for etanercept was similar to the double-blind rates. These data indicate that in patients with active AS, the rate of uveitis was not increased with etanercept use and did not change with longer treatment duration
Safety conclusion	The pooled analysis indicates that etanercept is a safe therapy for the long-term treatment of active AS (open-label results for up to 60 months). These results also indicate that patients who experience uveitis during treatment with etanercept can expect the event to be mild to moderate in severity and to resolve with no treatment or topical therapy.

Abbreviations: AC, active-control; AS, ankylosing spondylitis; CI, confidence intervals; COSTART, Coding Symbols for a Thesaurus of Adverse Reaction Terms; DB, double-blind; HLA, human leucocyte antigen OL, open-label.

3.7.6.2 Study: Gottlieb, 2011

Methodology and safety outcomes: Gottlieb, 2011

The source of this safety data for Gottlieb, 2011 is Gottlieb et al, 2011 (6). This study pools safety data on serious/opportunistic infections, malignancies and mortality across 49 US and non-US trials of etanercept across AS, rheumatoid arthritis (RA), juvenile idiopathic arthritis (JIA), psoriatic arthritis (PsA) and moderate-to-severe plaque psoriasis (PsO) indications. A summary of the methodology and safety results for Gottlieb, 2011 is provided in Table 68.

Table 68: Methodology and safety outcomes: Gottlieb, 2011

Study	Gottlieb, 2011
Summary	An observational, retrospective, pooled analysis of the safety of etanercept in patients with active AS, RA, JIA, PsA, and PsO. A total of 49 US and non-US clinical trials of etanercept were included, and data were extracted from 1) the controlled part of double-blind studies, and 2) uncontrolled double-blind and open-label studies. In total 13,877 patients were included from all trials, and by indication this involved the following: • AS: n=695 • JIA: n=486 • PsA: n=1362 • PsO: n=4361 • RA: n=6973 In the AS indication, six studies, were examined for reports of serious infections/opportunistic infections, malignancies, and mortality including: • Four placebo-controlled trials (10-12, 93) • Two open-label extensions (118, 119)
Objectives	To examine the rates of serious infections/opportunistic infections, malignancies, and mortality during treatment with etanercept and placebo in patients with active AS, and across other indications of interest (JIA, RA, PsA, and PsO).
Study location	The clinical trials were conducted in the US, and non-US locations (the latter

Study	Gottlieb, 2011
	unspecified).
Design details	Forty-nine trials (both double-blind and open-label extensions of double-blind trials, and open-label studies). Of these, in the AS indication there were: • Four placebo-controlled trials • Two open-label extensions Data from these studies was pooled for safety analysis.
Interventions	For AS, the post-hoc analysis included all patients randomised to either: • Etanercept 50 mg subcutaneous injection, weekly, or 25 mg twice weekly • Placebo
Key inclusion criteria	• Controlled portions of double-blind studies • Uncontrolled double-blind and open-label studies. • Only registries managed by Amgen or Wyeth • Studies focusing solely on pharmacokinetic parameters were excluded
Permitted concomitant medications	Patients could receive concomitant DMARDs
Safety outcomes	1) **Serious infections**: Met regulatory definition, i.e. fatal, life-threatening, requiring in-patient hospitalisation or prolongation of existing hospitalisation, persistent or significant disability or incapacity/congenital anomaly or birth defect, or another significant medical hazard. Identified cases were reviewed to determine the type of microorganism and infection site 2) **Opportunistic infections**: The US Centers for Disease Control and Prevention (CDC) definition of opportunistic infections was applied. 3) **Malignancies:** Identified malignancies were independently adjudicated by three physicians. For questionable events, narratives were reviewed. Only the incident cancer was counted, and metastatic disease and recurrences were counted as one case. Malignancies at different sites, without being metastatic disease, were counted as separate events. 4) **Mortality:** Deaths that occurred within 30 days after the last dose of study drug were included and reviewed. The underlying causes of death were then placed into categories used by the CDC and defined according to codes from the International Classification of Diseases,
Populations analysed	**Serious/opportunistic infections:** All events that occurred during the double-blind phase with onset dates between the first dose and the last dose of study drug were included in the analyses. Serious infectious events that occurred within 30 days after the last dose in the double-blind phase were captured in the analyses of all patients who continued beyond the double-blind phase and received at least one dose of etanercept in clinical studies. **Malignancies:** All events that occurred while the patient was taking the study drug or within 30 days after the last dose were tabulated. For patients who discontinued after the double-blind phase of a study, events that occurred 30 days after the last dose were also included. **Mortality:** Deaths that occurred within 30 days after the last dose of study drug were included and reviewed.
Statistical analysis	Review of identified events and cases was performed to classify deaths by cause, serious infections by site and type, and malignancies by type. Data were reported by trial, indication, and dosage. Standardised incidence ratios (SIRs) and standardised mortality ratios (SMRs) were calculated (observed/expected), taking sex and age distribution into consideration. All exposure-adjusted rates were presented as the number of events per 100 patient years. A corresponding 95% CI was calculated by the exact method based on the Poisson distribution method. Analyses were performed with SAS STAT system version 9.1 (SAS Institute Inc., Cary, NC).

Study	Gottlieb, 2011
Patient disposition	Double-blind analysis of included trials included 4,580 patients administered etanercept and 2,270 patients receiving placebo/concomitant DMARDs. The total analysis across all trial periods (double-blind and open-label) included 13,977 patients.
Patient demographics and baseline characteristics	In all clinical trials demographic characteristics were similar (data not shown). By indication patients ranged between 41 and 53 years of age (mean 10 years for JIA). The majority of patients in the RA (78%) and JIA (73%) indications were female, whereas in AS (75%) and PsO (64%) they were male.
Safety outcomes	**Serious/opportunistic infections** • Incidence of serious infections in AS patients (all clinical trials) was 3.01/100 patient-years (exposure-adjusted rate) which was identical to the rate in the overall (across-indication) etanercept trials analysis • Pooled double-blind analysis across indications indicated a rate ratio of 0.76 (95%CI 0.50, 1.18) for etanercept vs placebo, which was not statistically significant • For all clinical trials there were no significant differences in rates of serious infections in the different dosing groups within each indication • Most common events in AS were infections (pathogen class unspecified) with common sites being skin/subcutaneous tissue (0.90) and gastrointestinal tract (0.60) • In the double-blind trials analysis, no opportunistic infections were reported in patients from the AS, PsA and PsO studies • In the analysis of all clinical trials, no opportunistic infections were reported in the AS studies. ○ Two cases of tuberculosis were reported across all indications (SIR 1.08 (95% CI 0.13, 3.91) **Malignancies** • Rate ratio of malignancies in double-blind analysis across all indications for etanercept vs placebo was 0.88 (95%CI 0.41, 1.99) • No cases of lymphoma or leukaemia were reported in AS patients • Across all indications, rates of melanoma and basal cell carcinoma were not significantly different compared to the general population rates **Mortality** • No deaths were reported in the AS studies • Across indications, etanercept had an SMR of 0.46 (95%CI 0.36, 0.59), and was similar for US and non-US studies
Safety conclusion	This pooled analysis found rates of serious infectious events were generally similar between etanercept and control patients in the double-blind analysis of the clinical trials. The overall rates of opportunistic infections in the etanercept groups were low and were similar to controls. There were no cases of leukaemia or lymphoma in the AS studies. The results of the analysis also suggest no increase in mortality rates associated with etanercept use compared with control populations. Overall the analysis indicates that etanercept is generally well tolerated, with a low risk of serious AEs, including infections and malignancies across the indications (including AS) assessed.

Abbreviations: AE, adverse event; AS, ankylosing spondylitis; CI, confidence interval; DMARD, disease modifying anti-rheumatic drug; JIA, juvenile idiopathic arthritis; PsA, psoriatic arthritis; PsO, plaque psoriasis; RA, rheumatoid arthritis; SIR, standardised incidence ratio; SMR, standardised mortality ratio.

3.7.6.3 Study: van der Heijde, 2014

Methodology and safety outcomes: van der Heijde, 2014

The source of the safety data for van der Heijde, 2014 is van der Heijde et al, 2014 (165). This study pools safety data on serious infections, inflammatory bowel disease (IBD), malignancies and non-malignant skin cancers across eight etanercept trials reported in nine publications;Gorman et al, 2002 (93), Davis et al, 2003 (10), Calin et al, 2004 (11), van der Heijde et al, 2006 (12), Braun et al, 2008 (ASCEND trial) (127), Davis et al, 2008 (118), Dijkmans et al, 2009 (120) and Martin Mola et al, 2008 (172). Note that the Braun et al, 2008, and Martin-Mola et al, 2008, studies were conference abstracts which have since been superceded by full publications.

A summary of the methodology and safety results for van der Heijde, 2014 is provided in Table 69.

Table 69: Methodology and safety outcomes: van der Heijde, 2014

Study	van der Heijde, 2014
Summary	An observational, retrospective, analysis of the safety of etanercept in patients with active AS. Eight studies, reported in nine publications, were examined for reports of serious infections, IBD, malignancies, and non-malignant skin cancers including: • Four placebo-controlled trials (10-12, 93) • One active-controlled trial (127) • Three open-label trials (93, 118, 120, 172)
Objectives	To examine the rates of serious infections, IBD, malignancies, and non-malignant skin cancers during treatment with etanercept, SSZ and placebo in patients with active AS.
Study location	The clinical trials were conducted in the USA, Europe, Latin America and Asia.
Design details	Eight clinical trials of etanercept in patients with active AS were examined for reports of serious infections, IBD, malignancies, and non- malignant skin cancers: • Four placebo-controlled trials • One active-controlled trial • Three open-label trials Data from these studies was pooled for safety analysis.
Interventions	The post-hoc analysis included all patients randomised to either: • Etanercept 50 mg subcutaneous injection, weekly, or • SSZ up to 3 g orally, per day, or • Placebo
Key inclusion criteria	All randomised patients who received at least one dose of treatment with etanercept, SSZ or placebo were included in the study.
Permitted concomitant medications	Not reported in publication.
Safety outcomes	The objective of this post-hoc analysis was to examine events of serious infections, opportunistic infections, and some types of malignancies in patients with active AS who were treated with etanercept, sulfasalazine or placebo in the eight clinical trials. Study data were filtered by prespecified TEAEs using COSTART preferred terms and/or search stems. Serious infections were defined as those meeting regulatory criteria or those requiring hospitalisation and/or the need for parenteral anti-infectives. IBD events included new onset and flares. Malignancies were

Study	van der Heijde, 2014
	consistent with those included in the Surveillance, Epidemiology and End Results (SEER) database and excluded in situ cancers with the exception of bladder and non- melanoma skin cancers.
Populations analysed	All randomised patients who received at least one dose of treatment with etanercept, SSZ or placebo.
Statistical analysis	Exposure was calculated from the first to the last dose with the exclusion of time gaps between double-blind studies and open-label extensions. Endpoints included exposure-adjusted rates and rate ratios (RRs), with 95% CIs. Fisher's exact binomial test was used to analyse AE rates. AEs occurring within 30 days after the final dose of a double-blind study or an open-label study were analysed as occurring in those studies. THE SEER database was used to calculate standardised incidence ratios (SIRs) for malignancies. The SEER, Arizona basal cell carcinoma, Arizona squamous cell carcinoma, and Minnesota squamous cell carcinoma databases were used to calculate expected numbers of cancers.
Patient disposition	A total of 1,323 patients were included in the pooled dataset (etanercept, n=1,074; placebo, n=249; SSZ, n=187). Total exposure to study drugs was 1,520 patient-years, including 1,131 patient-years of exposure to etanercept.
Patient demographics and baseline characteristics	At baseline, disease duration, age, gender, weight, race, HLA-B27 positive status and treatment with corticosteroids were similar between treatment groups in the five double-blind studies (p>0.05). However, the following baseline characteristics were significantly different between treatment groups: • medical history of IBD was significantly higher in the placebo group (p=0.01) • treatment with DMARDs was significantly higher in the etanercept group (p=0.001) • treatment with NSAIDs was significantly lower in the placebo group (p=0.001).
Safety outcomes	<u>Incidence of serious infections and IBD-related events</u> Rates of serious infections, IBD-related events, and malignancies during the double-blind studies were comparable between the etanercept and SSZ/placebo groups. Patients treated with etanercept had numerically higher RRs of serious infections than with placebo (not significant). No opportunistic infections were reported during the trials. <u>Incidence of malignant and non-melanomas skin cancers</u> No cancers were reported during the double-blind studies. In the open-label extensions, six malignancies and two non-melanoma skin cancers were reported. The SIRs did not indicate a significantly elevated rate of malignancy or non-malignant skin cancer.

	Placebo + SSZ (n=436)	Etanercept DB (n=887)	Etanercept OL (n=1,074)
Study duration weeks/ total exposure (patient - years)	12-24/137.27	12-24/250.95	12-264/1,131.28
Serious infectious events			
Events	1[‡]	4	21
Rate[†]	0.73	1.59	1.86
RR (95% CI)	2.19 (0.22-107.79)	2.19 (0.22-107.79)	N/A
IBD-related events			
Events	1[‡§]	2[¶]	20[††]

Study	van der Heijde, 2014			
	Rate, total (new events, flares)†	0.73 (0.73, 0)	0.80 (0.40, 0.40)	1.77 (0.53, 1.24)
	RR (95% CI)	1.09 (0.06-64.56)	1.09 (0.06-64.56)	N/A
	Malignancies and non-melanoma skin cancers			
	Events	0	0	8
	Rate†	0	0	0.71
	RR (95% CI)	N/A	N/A	N/A

Abbreviations: CI, confidence intervals; DB, double-blind; IBD, inflammatory bowel disease; OL, open-label; RR, rate ratio; SSZ, sulfasalazine.
†Rate = number of events/100 patient years
‡Patient was treated with placebo
§No events were flares
¶One event was a flare
††Fourteen events were flares, occurring in 10 patients.

- The rate of IBD flare was higher for etanercept during the open-label extensions (1.24/100 patient-years) than during the double-blind studies (0.4/100 patient-years)
- Most patients (77%) with long-term exposure to etanercept and a history of IBD did not experience a flare during follow-up
- Based on a SIR analysis, the rate of malignancies in etanercept-treated patients (double-blind and open-label extensions) was not elevated relative to the general population (compared with the SEER database)
- Similarly, the etanercept rate of basal cell and squamous cell skin cancer was not elevated compared with reference databases of the general population.

Safety conclusion	This pooled analysis represents the largest safety dataset of etanercept-treated patients in active AS trial settings available to date. In the studies analysed, etanercept was associated with low, unadjusted rates of serious infections, opportunistic infections, malignancies, non-malignant skin cancers, and disease-related IBD. Rates of serious infectious events were similar between the double-blind studies and the open-label extensions and were comparable to other studies. Also, long-term exposure to etanercept was not associated with higher rates of serious infections. Although opportunistic infections, including tuberculosis, have been reported previously, none were recorded in these nine publications.

Abbreviations: AS, ankylosing spondylitis; CI, confidence intervals; COSTART, Coding Symbols for a Thesaurus of Adverse Reaction Terms; DB, double-blind; DMARD, disease modifying anti-rheumatic drug; HLA, human leucocyte antigen; IBD, inflammatory bowel disease; NSAID, non-steroidal anti-inflammatory drug; OL, open-label; RRs, rate ratios; SEER, Surveillance, Epidemiology and End Results; SIR, standardised incidence ratios; SSZ, sulfasalazine; TEAEs, treatment emergent adverse events.

3.7.7 Summary and conclusions

The safety data presented in this section demonstrates that etanercept has an established safety profile in AS patients and is generally well tolerated for the treatment of nr-axSpA and AS. There is extensive evidence from placebo-controlled trials which shows similar discontinuation rates for patients receiving etanercept and for patients receiving placebo. In addition, data from long term extension studies shows that etanercept is generally well tolerated for up to 7 years of therapy with a low frequency of treatment discontinuation due to adverse reactions.

Data from the pivotal 1031 trial provides evidence that the safety profile of etanercept for the treatment of patients with nr-axSpA is consistent with that reported in studies of etanercept

in patients with active AS. Adverse reactions reported in RCTs and open label extension studies of etanercept in patients with active AS were generally mild to moderate in intensity. In addition, any SAEs which were reported largely occurred as a single, isolated event or were considered to be unrelated to etanercept therapy. There were no new safety signals during the 1031 study.

No incidences of uveitis reported in any of the studies in nr-axSpA and AS patients resulted in patient discontinuation. In addition, a pooled analysis investigating the incidence of uveitis in trials of patients receiving etanercept for active AS reported that any occurrences of the condition tended to be mild to moderate in severity and were generally resolved with topical treatments. Also, the largest pooled safety evidence dataset presented for etanercept to date showed that in patients with active AS, etanercept was associated with low rates of serious infections, opportunistic infections, malignancies, non-malignant skin cancers, and disease-related IBD

In conclusion, there is extensive clinical evidence which shows that etanercept is well tolerated across the spectrum of AxSpA.

4. Comparative efficacy and safety

4.1 Summary

In the absence of direct head-to-head evidence, NICE recommends conducting a network meta-analysis to investigate all relevant direct and indirect treatment comparisons (173). Therefore a Bayesian network meta-analyses (NMA) was conducted separately for the AS and nr-axSpA populations to compare within their licensed populations: 1) direct comparisons of active treatments versus placebo (representing conventional care, NSAIDs), and 2) indirect comparisons of etanercept versus the other active treatments of interest.

Network meta-analysis

Methods:

The NMA was performed on difference in mean change from baseline in BASFI and BASDAI to provide comparative estimates of efficacy for use in the economic model. Complimentary estimates of efficacy based on an objective endpoint, CRP (AS) and ASDAS-CRP (nr-axSpA), were performed. Where robust analyses were possible, results for both fixed effect and random effects models are presented, and do not differ significantly in their goodness of fit.

Results:

- **AS (fixed effect and random effects model)**

Anti-TNFs demonstrate a benefit over PBO (conventional care) and have comparable clinical efficacy

  - The fixed effects model indicated all anti-TNFs analysed had statistically significant improvements on the BASFI, BASDAI and CRP outcomes versus PBO. These results are generally supported by the random effects model which show a similar numerical differences, although statistical significance is not reached by CZP in BASFI, and for BASDAI only maintained for ETN versus PBO.
  - There were no statistically significant differences between ETN and the other comparator treatments on BASFI or BASDAI (fixed effects and random effects models)
    - When using the fixed effects model, CZP was associated with significantly less improvement in BASDAI total score versus INF
  - ETN and ADA had statistically significant improvements in CRP versus GOL; there were no statistically significant differences for the other treatment comparisons on this outcome (fixed effects model)

- **Nr-axSpA (fixed effects model)**

Consistent with the available trial evidence, the NMA supports the conclusion that anti-TNFs provide a significant benefit over conventional care.

  - ETN, ADA and CZP had statistically significant improvements on the BASFI and BASDAI versus PBO

- Due to lack of data, it was not possible to perform an analysis on ASDAS-CRP.

Extensive heterogeneity in nr-axSpA trial populations exists in important baseline characteristics (e.g. baseline CRP, BASDAI and BASFI, age, disease duration) and disease classification. These differences may impact outcomes analysed in the NMA, and confound the comparisons between anti-TNFs, making the results of the NMA potentially unreliable.

In the absence of sufficient information for meta-regression, heterogeneity was addressed using matching adjusted indirect comparison (MAIC) and simulated treatment comparison (STC) analyses of patient level data available to Pfizer. The results of both analyses were consistent with each other, and suggest that if adjustments are made to address the differences in baseline characteristics and disease classification, ETN has comparable efficacy with CZP and ADA.

4.2 Efficacy: Meta-analysis and indirect comparisons

4.2.1 Methods

A detailed description of the systematic review methodology, data extraction methodology, literature search results and related flow diagrams for the etanercept and comparator RCTs are reported in section 2 and Appendix A.

4.2.1.1 Statistical methods

Data extraction and handling of missing data

Study characteristics, patient characteristics, treatment regimen and outcome data were extracted from the relevant RCTs. Information regarding the following patient characteristics at baseline were extracted and used to assess comparability of studies: age (mean); sex (% male); disease duration (years); % with prior NSAIDs; % with prior DMARDs; % on concomitant DMARDs; % on concomitant NSAIDs; % on concomitant sulfasalazine; % on concomitant methotrexate; % on concomitant corticosteroids; % fulfilling modified New York criteria; % with history of arthritis; % with history of psoriasis; % with a history of uveitis; % with a history of IBD; % HLA B27 positive; CRP, mg/dl (mean); ESR, mm/h (mean); Total Back score; BASFI, cm (mean); BASDAI, cm (mean); BASMI, cm (mean); Treatment experienced or naïve (n, %), SPARCC score (mean).

The primary objective of the NMA was to provide comparative estimates for the outcomes used to drive the economic model (BASDAI and BASFI). In line with the most recent recommendations on measuring treatment benefit (see section 3.2.1) an analysis of an acute phase reactant was also performed to provide an estimate of relative efficacy using an objective measure. For this outcome, CRP was selected in accordance with recent publications (29). For the outcomes of interest, BASFI, BASDAI and CRP, the mean change from baseline (CFB) was extracted whenever available, along with corresponding sample size, standard deviation (SD) and measures of uncertainty (i.e. standard error (SE), 95% confidence intervals [CI], and p value) for all relevant intervention groups. If CFB was not provided in the full text reports, the measure at baseline and at follow-up was used to calculate CFB. The corresponding SE was calculated by assuming a conservative 0.50

correlation between the measures at follow-up and baseline. If the SE was not reported, but other measures of variance were available, the SE was calculated using the following hierarchy: SD by intervention group along with the sample size, 95% CI by intervention group; 95% CI of the difference between intervention groups.

For studies for which no variance measure was reported, imputation was performed. First, the standard deviations were calculated for all remaining studies (by dividing the SE by the square root of the number of patients in each intervention group), the mean of these SDs was then applied to the studies with missing SE, and then the SDs were divided by the square root of the number of patients to obtain an estimate of the SE for each intervention group.

NMA models

Analyses were performed in the Bayesian framework and involve data, a likelihood distribution, a model with parameters, and prior distributions. The plausibility of model assumptions should guide model choice and, in general, the assumptions of random effects models are more plausible than those of the fixed effect model. However, the evidence networks consisted of a limited number of studies and some connections in networks consisted of only single trials. A limited number of studies in a network may result in unstable heterogeneity estimates when using random effects models. Accordingly, both fixed and random effects models were used for the NMA. See Appendix C for details regarding model specification.

For both models the goodness-of-fit of model predictions to the observed data was measured by calculating the posterior mean residual deviance, $\overline{D}$ $\overline{D}$ (174). Subsequently the deviance information criterion (DIC) was obtained which provides a measure of the trade-off between model fit and parsimony (the fixed effect model being the most parsimonious). The model with the better trade-off between fit and parsimony has a lower DIC. A difference in DIC of about 3 points can be considered meaningful. Since model fit is not a proxy for the plausibility of model assumptions we did not want to solely rely on the DIC to select fixed or random effects models. When the results of the random effects model were unstable (i.e. very large credible intervals), only the results obtained with the fixed effects model were presented. When both models provided meaningful results, results obtained with both models were presented along with the DIC. When an evidence network for a particular analysis consisted of only single RCTs for each direct comparison (i.e. one trial per connection) in the network, it was not feasible to estimate the heterogeneity parameter and only the fixed effects model was used.

Software

The parameters of the different models were estimated using a Markov Chain Monte Carlo (MCMC) method as implemented in the OpenBUGS software package (175, 176). A first series of iterations from the OpenBUGS sampler were discarded as ‘burn-in’ and the inferences were based on additional iterations using three chains. Convergence of the chains was confirmed by the Gelman-Rubin statistic. All analyses were performed using R version 3.0.3 (http://www.r-project.org/) and OpenBugs version 3.2.3 (OpenBUGS Project Management Group). The OpenBUGS code is presented in Appendix C.

Presentation of results

The primary outputs of the Bayesian NMA are posterior distributions of the relative treatment effects between all interventions in the networks, i.e. differences in CFB BASFI, CFB

BASDAI, and CFB CRP. In addition, the expected CFB for each of these outcomes by treatment was modelled ("modelled outcomes") by combining the relative treatment effect estimates versus placebo obtained from the NMA with the average outcome with placebo from the placebo controlled trials. The posterior distributions of relative treatment effects and modelled outcomes were summarised by the median and 95% credible intervals (CrIs), which were constructed from the 2.5th and 97.5th percentiles.

Evaluation of consistency between direct and indirect comparisons

Prior to the NMA, the consistency between direct and indirect comparisons was evaluated for networks that consisted of closed loops. For each of the comparisons (i.e. contrasts) that were part of a closed loop made up of more than one RCT, the available studies were split into direct and indirect information. For each contrast in question, two (pooled) relative treatment effect estimates were obtained, one with independent-means (or independent-effects) models using only the studies providing direct comparisons, and one based on an NMA of the remaining studies providing only indirect evidence. This iterative technique is called edge-splitting (177). The difference in estimates generated by the two sets of evidence was evaluated with the Bucher test for inconsistency (175). There were only sufficient studies in the AS BASFI and BASDAI networks included as sensitivity analyses to enable evaluation of consistency using this method for the placebo-etanercept-infliximab closed loop. The results of these analyses are presented in Appendix F as part of the sensitivity analyses of the NMA

4.2.2 Studies eligible for meta-analysis

4.2.2.1 Ankylosing Spondylitis (AS) RCT evidence

The original and updated electronic database searches identified 29 publications (11, 12, 14, 41, 69-71, 89, 91, 93-101, 104, 106-111, 114, 133, 178, 179) and the conference abstract searches identified 4 abstracts (129-131, 135) reporting on a total of 20 AS RCTs. Of these publications, 18 reporting on 14 RCTs (11, 69, 70, 91, 95, 101, 104, 106-108, 178-184) included outcomes data that appeared relevant for the NMA, specifically, at least one of the following:

- CRP at 12 weeks,
- BASDAI at 12 weeks
- BASFI at 12 weeks

The study design, characteristics and patient baseline disease status in these 18 publications (14 RCTS) are reported in Table 70 and Table 71.

Table 70: Characteristics of etanercept and comparator treatment RCTs in the AS population eligible for inclusion in the NMA

Publication & trial name	Study design	Treatment arms, dose regimen N= number randomised	Treatment duration (weeks)	Population	Age in Years, mean (SD)	Duration of disease, years; mean (SD)	Symptoms, years; mean (SD)	Prior NSAIDs use, n (%)	Concomitant NSAIDs use, n (%)
Indication: Ankylosing Spondylitis									
Etanercept									
16.0037/ Davis, 2003 (full paper) (CSR-48803) (10)	Multicentre, double-blind, RCT	ETN 25 mg, twice weekly N= 138 PBO N=139	24	Adult patients ≥18 years mNY criteria† BASFI ≥ 30‡ Total back pain or patient global assessment ≥30‡ Morning stiffness of ≥30‡	ETN arm: 42.1 (9.7) PBO arm: 41.9 (11.1)	ETN arm: 10.1 (8.2) PBO arm: 10.5 (9)	ETN arm: NR PBO arm: NR	ETN arm: 126 (91) PBO arm: 128 (92)	ETN arm: NR PBO arm: NR
16.0626/ Gorman, 2002 (full paper) (CSR-48704) (93)	Multicentre, double-blind, RCT	ETN 25 mg, twice weekly N= 20 PBO N=20	16	Adult patients ≥18 years mNY criteria† Inflammatory back pain Morning stiffness of ≥45 mins	ETN arm: [redacted] PBO arm: [redacted]	ETN arm: [redacted] PBO arm: [redacted]	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: [redacted] PBO arm: [redacted]
402-WW/ Braun, 2011	Multicentre, double-blind, RCT	ETN 50 mg, weekly N=	16	Adult patients ≥18 years	ETN arm: 40.69 (11.69)	ETN arm: [redacted]	ETN arm: [redacted]	ETN arm: 335 (88.39)	ETN arm: 324 (85.5)

ASCEND (full paper) (CSR-72737) (14)		379 SSZ 3 mg, daily, N=187		mNY criteria† BASDAI ≥30‡ PGADA VAS ≥30 average nocturnal and total pain VAS ≥30 BASFI VAS ≥30 at least 1 NSAID failed	SSZ arm: 40.9 (12.23)	SSZ arm: [redacted]	SSZ arm: [redacted]	SSZ arm: 159 (85.03)	SSZ arm: 154 (82.4)
311-EU/Calin, 2004 (full paper) (CSR-47687) (11)	Multicentre, double-blind, RCT	ETN 25 mg, twice weekly N= 45 PBO N=39	12	Adult patients ≥18 years mNY criteria† BASDAI ≥30‡ Inflammatory back pain ≥30 by VAS PGADA VAS ≥30 Back pain VAS ≥30 Physical function VAS ≥30	ETN arm: 45.3 (9.5) PBO arm: 40.7 (11.4)	ETN arm: 15 (8.8) PBO arm: 9.7 (8.2)	ETN arm: NR PBO arm: NR	ETN arm: [redacted] PBO arm: [redacted])	ETN arm: 40 (89) PBO arm: 33 (85)

314-EU/van der Heijde, 2006 (full paper) (CSR-59339) (12)	Multicentre, double-blind, RCT	ETN 50 mg, weekly N= 155 ETN 25 mg, twice weekly N= 150 PBO N=51	12	Adult patients ≥18 years mNY criteria† BASFI VAS ≥30 PGADA VAS ≥30 average nocturnal and total pain VAS ≥30 BASFI VAS ≥30 Morning stiffness of ≥30‡	ETN 50 mg arm: 41.46 (11.02) ETN 25 mg arm: 39.83 (10.71) PBO arm: 40.06 (10.93)	ETN 50 mg arm: 8.96 (8.65) ETN 25 mg arm: 9.97 (9.05) PBO arm: 8.48 (6.83)	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: [redacted] ETN 25 mg arm: [redacted]) PBO arm: [redacted]	ETN 50 mg arm: 124 (80) ETN 25 mg arm: 127 (84.7) PBO arm: 40 (78.4)
314-EU/ Braun 2007 (full paper) (CSR-59339) (98)	Multicentre, double-blind, RCT	ETN 50 mg, weekly N= 155 ETN 25 mg, twice weekly N= 150 PBO N=51	12	Adult patients ≥18 years mNY criteria† BASFI or BASDAII VAS ≥30 Morning stiffness of ≥30‡ PGADA VAS ≥ 30 average nocturnal and	ETN 50 mg arm: 41 ETN 25 mg arm: 40 PBO arm: 40	ETN 50 mg arm: 9 ETN 25 mg arm: 10 PBO arm: 8	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: 124 (80) ETN 25 mg arm: 127 (85) PBO arm: 40 (78)

				total pain VAS ≥30					
SPINE trial/ Dougados, 2011 (full paper) (13)	Multicentre, double-blind, RCT	ETN 50 mg, weekly N= 39 PBO N=43	12	Adult patients ≥18 years mNY criteria† BASDAI VAS ≥40	ETN arm: 46 (11) PBO arm: 48 (10)	ETN arm: 19 (10) PBO arm: 23 (11)	ETN arm: NR PBO arm: NR	ETN arm: 39 (100) PBO arm: 43 (100)	ETN arm: NR PBO arm: NR
Cantini, 2013 (full paper) (101)	RCT, blinding NR	ETN 50 mg, weekly N= 21 ETN 50 mg, every other week N= 22	NR	Adult patients ≥18 years mNY criteria†	ETN (weekly) arm:38 (median) ETN (every other week) arm:37 (median)	ETN (weekly) arm: 12(median) ETN (every other week) arm: 13(median)	ETN (weekly) arm: NR ETN (every other week) arm: NR	ETN (weekly) arm: NR ETN (every other week) arm: NR	ETN (weekly) arm: NR ETN (every other week) arm: NR
Giardina, 2010 (full paper) (104)	Open-label, RCT	ETN 50 mg, weekly N= 25 INF 5 mg, 0, 2, 6 and every 6 weeks N= 25	104	Adult patients ≥18 years mNY criteria† BASDAI >4 Spinal pain >4§ non-responders to oral NSAID	ETN arm: 32.6 (6.8) INF arm: 31.9 (9.2)	ETN arm: 15.7 (6.5) INF arm: 15.4 (10.6)	ETN arm: NR INF arm: NR	ETN arm: 25 (100) INF arm: 25 (100)	ETN arm: NR INF arm: NR
Navarro-Sarabi, 2011 (full paper) (110) LOADET	Multicentre, double-blind, RCT	ETN 50 mg, twice weekly N= 54 ETN 50 mg, once weekly N= 54	12	Adult patients ≥18 years mNY criteria† BASDAI ≥4 Spinal pain ≥4	ETN (twice weekly) arm: 40.22 (10.36) ETN (once weekly) arm: 42.63 (10.66)	ETN (twice weekly) arm:7.03 (6.83) ETN (once weekly) arm:7.28	ETN (twice weekly) arm:11.87 (8.45) ETN (once weekly) arm:14.35	ETN (twice weekly) arm: 54 (100) ETN (once weekly) arm:	ETN (twice weekly) arm: 45 (83.33) ETN (once weekly) arm: 48 (88.89)

				by VAS Patient's global disease assessment ≥4		(7.06)	(10.52)	54 (100)	
Infliximab									
Braun, 2002 ASSERT (full paper) (95)	Multicentre, double-blind, RCT	INF 5 mg, 0, 2, 6 weeks N=35 PBO N=35	12	Adult patients ≥18 years mNY criteria† BASDAI ≥4 Spinal pain ≥4§	INF arm: 40.6 (8) PBO arm: 39 (9.1)	INF arm: 16.4 (8.3) PBO arm: 14.9 (9.3)	INF arm: NR PBO arm: NR	INF arm: NR PBO arm: NR	INF arm: NR PBO arm: NR
Inman, 2010, CANDLE (full paper) (106)	Multicentre, double-blind, RCT	INF 3 mg, 0, 2, 6 weeks N=39 PBO N=37	12	Adult patients ≥18 years mNY criteria† BASDAI ≥4	INF arm: 42.9 (10.4) PBO arm: 39.3 (9)	INF arm: 11.7 (10.6) PBO arm: 11.1 (10.3)	INF arm: 18.7(11.3) PBO: 18.6(9.8)	INF arm: NR PBO arm: NR	INF arm: NR PBO arm: NR
Marksymowych, 2010 CANDLE (full paper) (108)	Multicentre, double-blind, RCT	INF 3 mg, 0, 2, 6 weeks N=18 PBO N=18	12	Adult patients ≥18 years mNY criteria† BASDAI ≥4	INF arm: 43.6 (11.8) PBO arm: 41.7 (9.3)	INF arm: 12(11.2) PBO arm: 14.3(12.0)	INF arm: 20 (11.8) PBO arm: 20.2 (11.5)	INF arm: 18 (100) PBO arm: 18 (100)	INF arm: NR PBO arm: NR
Certolizumab Pegol									
Landewe, 2012 RAPID-axSpA (full paper) (90)	Multicentre, double-blind, RCT	CZP 200 mg, every two weeks, N=65	48	Adult patients ≥18 years mNY criteria†	CZP 200 mg arm : 41 (10.8)	CZP 200 mg arm : 7.4 (7.8)	CZP 200 mg arm : 10.5 (8.8)	CZP 200 mg arm : 52 (80)	CZP 200 mg arm : NR

		CZP 400 mg, every four weeks, N=56 PBO N=51		BASDAI >4 Spinal pain >4 §	CZP 400 mg arm : 41.9 (11.5) PBO arm: 41.6 (12.8)	CZP 400 mg arm : 7.5 (7.8) PBO arm: 9.9 (8.5)	CZP 400 mg arm : 12.4 (10.1) PBO arm: 13.1 (10.8)	CZP 400 mg arm : 44 (79) PBO arm: 55 (97)	CZP 400 mg arm : NR PBO arm: NR
Adalimumab									
Lambert 2007 (full paper) (107) M03-606	Multicentre, double-blind, RCT	ADA 40 mg, every other week, N=39 PBO, N=44	24	Adult patients ≥18 years mNY criteria† BASDAI ≥4 Total back pain ≥4‡ Morning stiffness of ≥1 hour Failed treatment ≥1 NSAIDs	ADA arm: 41.9 PBO arm: 40	ADA arm: 14.5 (9) PBO arm: 12.1 (8.7)	ADA arm: NR PBO arm: NR	ADA arm: 38 (100) PBO arm: 44 (100)	ADA arm: NR PBO arm: NR
Van der Heidje 2006 (full paper) (91) ATLAS	Multicentre, double-blind, RCT	ADA 40 mg, every other week, N= 208 PBO N=107	24	Adult patients ≥18 years mNY criteria† BASDAI ≥4 Total back pain ≥4§	ADA arm: 41.7 (11.69) PBO arm: 43.4 (11.32)	ADA arm: 11.3 (9.99) PBO arm: 10 (8.34)	ADA arm: NR PBO arm: NR	ADA arm: 208 (100) PBO arm: 107 (100)	ADA arm: 166 (79.8) PBO arm: 84 (78.5)

				Morning stiffness of ≥1 hour					
Revicki 2008 (full paper) (70) ATLAS	Multicentre, double-blind, RCT	ADA 40 mg, every other week N=208 PBO N=107	24	Adult patients ≥18 years mNY criteria† intolerance to ≥1NSAIDs	ADA arm: 41.7 (11.7) PBO arm: 43.4 (11.3)	ADA arm: 11.3 (10) PBO arm: 10 (8.3)	ADA arm: NR PBO arm: NR	ADA arm: 208 (100) PBO arm: 107 (100)	ADA arm: NR PBO arm: NR
Golimumab									
Inman, 2008 GO-RAISE (full paper) (69)	Multicentre, double-blind, RCT	GOL 50 mg, every four weeks, N=138 GOL 100 mg, every four weeks, N=140 PBO N=78	24	Adult patients ≥18 years mNY criteria† BASDAI ≥4 Spinal pain ≥4§ inadequate response to current or previous NSAIDs or DMARD	GOL 50 mg arm:38(median) GOL 100 mg arm: 38(median) PBO arm: 41 (median)	GOL 50 mg arm: 5.15(median) GOL 100 mg arm: 5.20(median) PBO arm: 7.25(median)	GOL 50 mg arm: 11 (median) GOL 100 mg arm: 9.5 (median) PBO arm: 16 (median)	GOL 50 mg arm: NR GOL 100 mg arm: NR PBO arm: NR	GOL 50 mg arm: 124 (89.9) GOL 100 mg arm: 123 (87.9) PBO arm: 72 (92.3)

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CSR, Clinical Study Report; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; mNY, Modified New York; NSAID, non-steroidal anti-inflammatory drug; PBO, placebo; PGADA, Patient global assessment of disease activity; RCT, randomised controlled trial; SD, standard deviation; SSZ, sulfasalazine; VAS, visual analogue scale.
† Active AS based on modified New York criteria
‡ Scores on a 0-100-mm visual analogue scale
§ Scores on a 0-10-cm visual analogue scale

Table 71: Baseline disease status in etanercept and comparator treatment RCTs in the AS population eligible for inclusion in the NMA

Publication & trial name	Treatment arms	Smoking Status	HLA B27 positive: n (%)	Elevated CRP at baseline: n (%)	Baseline BASDAI:	Baseline BASFI:	mSASSS scores
Indication: Ankylosing spondylitis							
Etanercept							
16.0037/ Davis, 2003 (full paper) (CSR-48803) (10)	ETN 25 mg, PBO	ETN arm: NR PBO arm: NR	ETN arm: 108 (84) PBO arm: 109 (84)	ETN arm: NR PBO arm: NR	ETN arm: 5.81(0.15)‡ PBO arm: 5.96 (0.14)‡	ETN arm: 5.17(0.18)‡ PBO arm:5.63(0.17)‡	ETN arm: NR PBO arm: NR
16.0626/ Gorman, 2002 (full paper) (CSR-48704) (93)	ETN 25 mg, PBO	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: [redacted] PBO arm: [redacted]	ETN arm: NR PBO arm: NR
402-WW/ Braun, 2011 ASCEND (full paper) (CSR-72737)(14)	ETN 50 mg, SSZ 3 mg,	ETN arm: NR SSZ arm: NR	ETN arm: [redacted] SSZ arm: [redacted]	ETN arm: NR SSZ arm: NR	ETN arm: 5.93 (1.63)† SSZ arm: 5.91 (1.59) †	ETN arm: 5.5 (SD, 2.02) † SSZ arm: 5.5 (2.08) †	ETN arm: NR SSZ arm: NR
311-EU/Calin, 2004 (full paper) (CSR-47687) (11)	ETN 25 mg, PBO	ETN arm: NR PBO arm: NR	ETN arm: [redacted] PBO arm: [redacted])	ETN arm: NR PBO arm: NR	ETN arm: 6.1 (1.57) † PBO arm: 5.86 (1.32) †	ETN arm: 6.02 (2.07) † PBO arm: 5.72 (1.62) †	ETN arm: NR PBO arm: NR
314-EU/van der Heijde, 2006 (full paper) (CSR-59339) (12)	ETN 50 mg, ETN 25 mg, PBO	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: [redacted] ETN 25 mg arm: [redacted] PBO arm: [redacted]	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: 6.2(1.69) † ETN 25 mg arm: 5.9(1.67) † PBO arm: 6.1(1.3) †	ETN 50 mg arm: 6.05(2.02) † ETN 25 mg arm: 5.77(2.01) † PBO arm: 5.96(1.93) †	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR
314-EU/ Braun 2007 (full paper) (CSR-59339) (98)	ETN 50 mg, ETN 25 mg,	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: [redacted] ETN 25 mg arm: [redacted]) PBO arm:	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR	ETN 50 mg arm: 6.2¶ ETN 25 mg arm: 5.9¶ PBO arm: 6.1¶	ETN 50 mg arm: 6.1¶ ETN 25 mg arm: 5.8¶ PBO arm: 6¶	ETN 50 mg arm: NR ETN 25 mg arm: NR PBO arm: NR

	PBO		[redacted]				
SPINE trial/ Dougados, 2011 (full paper) (13)	ETN 50 mg, PBO	ETN arm: NR PBO arm: NR	ETN arm: 31(79) PBO arm: 36(86)	ETN arm: NR PBO arm: NR	ETN arm: 6.4(1.2) † PBO arm: 5.8(1.5) †	ETN arm: 6.3(2) † PBO arm: 5.7(1.9) †	ETN arm: 34 (22) † PBO arm: 39 (19) †
Cantini, 2013 (full paper) (101)	ETN 50 mg, (weekly) ETN 50 mg, (every other week)	ETN (weekly) arm: NR ETN (every other week) arm: NR	ETN (weekly) arm: NR ETN (every other week) arm: NR	ETN (weekly) arm: NR ETN (every other week) arm: NR	ETN (weekly) arm: 2.3(0.4) † ETN (every other week) arm: 2.4 (0.3) †	ETN (weekly) arm: 2.5 (1.3) † ETN (every other week) arm: 2.6(1.4) †	ETN (weekly) arm: NR ETN (every other week) arm: NR
Giardina, 2010 (full paper) (104)	ETN 50 mg, INF 5 mg,	ETN arm: NR INF arm: NR	ETN arm: 24(96) INF arm: 23(92)	ETN arm: NR INF arm: NR	ETN arm: 6.6(1.1) † INF arm: 6.5(1.2) †	ETN arm: 6.5(1.1) † INF arm: 6.1(0.9) †	ETN arm: NR INF arm: NR
Navarro-Sarabi, 2011 (full paper) (110) LOADET	ETN 50 mg, twice weekly ETN 50 mg, once weekly	ETN twice weekly arm: NR ETN (once weekly) arm: NR	ETN twice weekly arm: 48 (90.57) ETN (once weekly) arm: 40 (74.07)	ETN twice weekly arm: NR ETN (once weekly) arm: NR	ETN twice weekly arm: 6.1(1.7) † ETN (once weekly) arm: 6.48 (1.61) †	ETN twice weekly arm: 5.87(1.96) † ETN (once weekly) arm: 5.66(2.04) †	ETN twice weekly arm: NR ETN (once weekly) arm: NR
Infliximab							
Braun, 2002 ASSERT(full paper) (95)	INF 5 mg, PBO	INF arm: NR PBO arm: NR	INF arm: 31 (91) PBO arm: 27 (88)	INF arm: NR PBO arm: NR	INF arm: 6.5(1.2) † PBO arm: 6.3(1.4) †	INF arm: 5.4 (1.8) † PBO arm: 5.1(2.2) †	INF arm: NR PBO arm: NR
Inman, 2010, CANDLE (full paper) (106)	INF 3 mg PBO	INF arm: NR PBO arm: NR	INF arm: 28 (72) PBO arm: 27 (73)	INF arm: NR PBO arm: NR	INF arm: NR PBO arm: NR	INF arm: NR PBO arm: NR	INF arm: NR PBO arm: NR
Marksymowych, 2010 CANDLE (full paper) (108)	INF 3 mg PBO	INF arm: NR PBO arm: NR	INF arm: 14 (77.8) PBO arm: 13 (72.2)	INF arm: NR PBO arm: NR	INF arm: 6.6(1.2) † PBO arm: 6.8(1.4) †	INF arm: 6.9 (1.4) † PBO arm: 6.1 (2) †	INF arm: NR PBO arm: NR

Certolizumab pegol							
Landewe, 2012 RAPID-axSpA (full paper) (90)	CZP 200 mg, CZP 400 mg, PBO	CZP 200 mg arm : NR CZP 400 mg arm : NR PBO arm: NR	CZP 200 mg arm : 53 (81.5) CZP 400 mg arm : 44 (78.6) PBO arm: 48 (84.2)	CZP 200 mg arm : NR CZP 400 mg arm : NR PBO arm: NR	CZP 200 mg arm : 6.5(1.7) † CZP 400 mg arm : 6.2(1.3) † PBO arm: 6.4(1.9) †	CZP 200 mg arm : 5.6(2.3) † CZP 400 mg arm : 5.7(2.3) † PBO arm: 6 (2.0) †	CZP 200 mg arm : NR CZP 400 mg arm : NR PBO arm: NR
Adalimumab							
Lambert 2007 (full paper) (107) M03-606	ADA 40 mg, PBO	ADA arm: NR PBO arm: NR	ADA arm: 33(86.8) PBO arm: 36(81.8)	ADA arm: NR PBO arm: NR	ADA arm: 6.2(1.7) † PBO arm: 6.5(1.6) †	ADA arm: 5.3(2.0) † PBO arm: 5.6(2.2) †	ADA arm: NR PBO arm: NR
Van der Heidje 2006 (full paper) (91) ATLAS	ADA 40 mg, PBO	ADA arm: NR PBO arm: NR	ADA arm:163(78.4) PBO arm: 85(79.4)	ADA arm: NR PBO arm: NR	ADA arm: 6.3(1.7) † PBO arm: 6.3(1.7) †	ADA arm: 5.2(2.2) † PBO arm: 5.6(2.2) †	ADA arm: NR PBO arm: NR
Revicki 2008 (full paper) (70) ATLAS	ADA 40 mg, PBO	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR
Golimumab							
Inman, 2008 GO-RAISE (full paper) (69)	GOL 50 mg, GOL 100 mg, PBO	GOL 50 mg arm: NR GOL 100 mg arm: NR PBO arm: NR	GOL 50 mg arm: 112 (81.8) GOL 100 mg arm: 118 (84.3) PBO arm: 66 (84.6)	GOL 50 mg arm: NR GOL 100 mg arm: NR PBO arm: NR	GOL 50 mg arm: 6.6†† GOL 100 mg arm: 7†† PBO arm: 6.6††	GOL 50 mg arm: 5†† GOL 100 mg arm: 5.4†† PBO arm: 4.9††	GOL 50 mg arm: NR GOL 100 mg arm: NR PBO arm: NR

Abbreviations: ADA, adalimumab; ; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C-reactive protein; CSR, Clinical Study Report; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; FDA, Food and Drug Administration; HLA B27, Human Leukocyte Antigen B27; IFN, infliximab; mSASSS, modified Stoke Ankylosing Spondylitis Spine Score; NR, not reported; PBO, placebo; SD, standard deviation; SE , Standard Error; SSZ, sulfasalazine;
† Mean (SD); ‡ Mean (SE); § Mean (Median); ¶ Mean; †† Median

Etanercept

Of 11 publications, ten (11, 12, 14, 93, 98, 101, 104, 110, 178, 179) reported on nine RCTs of etanercept. The eleventh publication (89) reported on a 10th RCT which did not include outcomes relevant to the NMA. All nine RCTS included double-blind phases with the exception of two studies (101, 104) where blinding was not reported and an open-label design was used, respectively. Three RCTs compared etanercept 25 mg administered twice weekly with placebo (11, 93, 178). Two publications reported on one RCT comparing etanercept 25 mg twice weekly versus etanercept 50 mg once weekly versus placebo (12, 98). In a fifth RCT, etanercept 50 mg versus placebo was investigated in advanced AS patients with spinal ankylosis (179). The sixth and seventh RCTs investigated etanercept 50 mg weekly versus: 1) 50 mg every other week (101), and 2) 50 mg twice weekly (110). Finally, two head-to-head RCTs compared etanercept 50 mg once weekly to infliximab 5 mg (104) and sulfasalazine 3 mg (14), respectively.

Infliximab

Three publications reported on two double-blind RCTs of infliximab versus placebo (95, 106, 108). One study compared infliximab 5 mg administered at weeks 0, 2, 6, and every 6 weeks thereafter to placebo (95), while the other two publications reported on an RCT comparing low-dose infliximab 3 mg at weeks 0, 2 and 6 to placebo (106, 108).

Adalimumab

For adalimumab, three publications reported on two double-blind RCTs versus placebo (70, 91, 107), with both studies investigating adalimumab 40 mg administered every other week.

Golimumab

One study publication reported on a double-blind RCT comparing golimumab 50 mg versus 100 mg administered every 4 weeks versus placebo (69).

4.2.2.2 Axial spondyloarthritis (including AS and nr-axSpA population sub-groups): RCT evidence

The original and updated electronic database searches identified two publications (15, 90) and the conference searches identified two abstracts (131, 134) reporting on a total of three RCTs. An FDA report was also identified that included additional data on one of these RCTs (140). Of these, one publication and the FDA report, both reporting on the same RCT, included outcomes data that appeared relevant to the NMA (90, 140). An etanercept RCT (15) in this population did not report results for AS or nr-axSpA sub-groups for the randomised part of the study, although subgroup analyses were conducted on a subset of the study population. This study was therefore excluded from consideration for the NMA, but is included in the clinical efficacy and safety sections in this MTA (section 3).

Certolizumab pegol

One double-blind RCT investigated certolizumab pegol 200 mg administered every two weeks versus certolizumab pegol 400 mg every four weeks versus placebo in patients with axial spondyloarthritis (axSpA), including patient subgroups with AS and nr-axSpA (90). This study was eligible for inclusion in the NMA since it provided data for the specific AS and nr-axSpA populations in the main study (Table 70 and Table 72. Additional outcome data for the nr-axSpA population subgroup of this trial was also sourced from an FDA report (Table 72 identified during the systematic review hand search (140).

4.2.2.3 Non-radiographic axial spondyloarthritis (nr-axSpA): RCT evidence

The original and updated electronic database searches identified two publications (105, 112) and the conference searches identified six abstracts (126, 128, 136-138, 185) reporting on a total of three RCTs. In addition, one FDA report identified in the systematic review hand search reported additional data on the adalimumab ABILITY-1 RCT identified in the database search (139).

Of the total sources identified, the two publications, three abstracts (126, 128, 185) and the FDA report included outcomes data relevant to the NMA.

- ASDAS-CRP at 12 weeks,
- BASDAI at 12 weeks,
- BASFI at 12 weeks

The key study design, characteristics, and patient baseline disease status features of the trials reporting nr-axSpA populations are presented in Table 72 and Table 73.

Table 72: Characteristics of etanercept and comparator treatment RCTs in the nr-axSpA population eligible for inclusion in the NMA

Publication & trial name	Study design	Treatment arms, dose regimen N= number randomised	Treatment duration in weeks	Population	Age in years, mean (SD)	Duration of disease, years; mean (SD)	Symptoms, years; mean (SD)	Prior NSAIDS use, n (%)	Concomitant NSAIDs use, n (%)
Non-radiographic axial spondyloarthritis									
Etanercept studies									
Study1031, (CSR-205791) (141) Sep 2013 Primary analysis of primary outcome	Multicentre, double-blind, RCT	ETN 50 mg, weekly N=111 PBO N=114	12	Adult patients ≥18 years ASAS classification criteria† BASDAI ≥4 failed at least 2 NSAIDs (including current NSAID)	ETN arm: 31.6 (7.8) PBO arm: 32 (7.8)	ETN arm: 2.44 (1.91) PBO arm: 2.49 (1.72)	ETN arm: 2.4 (1.9) PBO arm: 2.5 (1.7)	ETN arm: 111 (100) PBO arm: 114 (100)	ETN arm: [redacted] PBO arm: [redacted]
Study1031, (CSR 205791- (161) Feb 2014 Updated analysis on open-label extension	Multicentre, double-blind, RCT	ETN 50mg, weekly N=111 PBO N=114	12	Adult patients ≥18 years ASAS classification criteria† BASDAI ≥4 failed ≥2 NSAIDs (including current NSAID)	ETN arm: 31.6 (7.8) PBO arm: 32 (7.8)	ETN arm: 2.4 (1.9) PBO arm: 2.5 (1.8)	ETN arm: NR PBO arm: NR	ETN arm: 111 (100) PBO arm: 114 (100)	ETN arm: [redacted] PBO arm: [redacted]
Dougados 2013 (abstract)	Double-blind, RCT	ETN 50 mg, weekly N=NR	12	Adult patients ≥18 years	ETN arm: NR	ETN arm: NR	ETN arm: NR	ETN arm: (100)	ETN arm: NR

(128) Primary analysis data-cut		PBO N=NR		ASAS classification criteria† BASDAI ≥4 failed ≥ 2 NSAIDs (including current NSAID)	PBO arm: NR	PBO arm: NR	PBO arm: NR	PBO arm: (100)	PBO arm: NR
Dougados 2013 (abstract) (185) Primary analysis data-cut	Double-blind, RCT	ETN 50 mg, weekly N=NR PBO N=NR	12	Adult patients ≥18 years ASAS classification criteria† BASDAI ≥4 failed ≥ 2 NSAIDs (including current NSAID)	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: (100) PBO arm: (100)	ETN arm: NR PBO arm: NR
Dougados 2013 (abstract) (126) Primary analysis data-cut	Double-blind, RCT	ETN 50mg, weekly N=106 PBO N=109	12	Adult patients ≥18 years ASAS classification criteria† BASDAI ≥4 failed ≥2 NSAIDs (including	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: 106 (100) PBO arm: 109 (100)	ETN arm: NR PBO arm: NR

				current NSAID)					
Adalimumab studies									
Sieper, 2013 ABILITY-1 (112)	Multicentre, double-blind, RCT	ADA 40 mg, every other week, N=91 PBO, N=94	12	Adult patients ≥18 years ASAS classification criteria† BASDAI ≥4 Total back pain ≥4‡ Intolerant to one or more NSAIDs	ADA arm: 37.6 (11.3) PBO arm: 38.4 (10.4)	ADA arm: 2.7 (4.2) PBO arm: 3.0 (3.8)	ADA arm: 10.1 (9) PBO arm: 10.1 (8.8)	ADA arm: 91 (100) PBO arm: 94 (100)	ADA arm: 72 (79) PBO arm: 74 (79)
FDA 2013 ABILITY-1 (139)	Double-blind, open label RCT	ADA 40 mg, every other week, N=NR PBO, N=NR	12	Adult patients ≥18 years ASAS classification criteria† Intolerant to one or more NSAIDs	ADA arm: NR PBO arm: NR	ADA arm: 3(4) PBO arm: 3(4)	ADA arm: 10(9) PBO arm: 10(9)	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR
Haibel, 2008 (105)	Multicentre, double-blind, RCT	ADA 40 mg, every other week, N=22 PBO, N=24	12	Adult patients ≥18 years Classification criteria meeting ASAS classification† BASDAI ≥4	ADA arm: 38 PBO arm: 37	ADA arm: 7 PBO arm: 8	ADA arm: NR PBO arm: NR	ADA arm: 22 (100) PBO arm: 24 (100)	ADA arm: NR PBO arm: NR

				Prior failure of ≥1 NSAID					
Certolizumab Pegol studies									
Landewe, 2012 RAPID-axSpA (full paper) (90)	Multicentre, double-blind, RCT	CZP 200 mg, every two weeks, N=46 CZP 400 mg, every four weeks, N=51 PBO N=50	48	Adult patients ≥18 years ASAS classification criteria† BASDAI >4 Spinal pain ≥4‡ Intolerance to ≥1 NSAID	CZP 200 mg arm: 36.6 (13) CZP 400 mg arm: 37.5 (10.8) PBO arm: 38 (11.8)	CZP 200 mg arm: 4.6 (6.7) CZP 400 mg arm: 5.9 (6.8) PBO arm: 4.2(4.7)	CZP 200 mg arm: 8.5 (9.1) CZP 400 mg arm: 8.1 (6.8) PBO arm: 9.2 (10)	CZP 200 mg arm: 39 (85) CZP 400 mg arm: 47 (92) PBO arm: 40 (80)	CZP 200 mg arm: 38 (83) CZP 400 mg arm: 44 (86) PBO arm: 41 (82)
FDA 2013 RAPID-axSpA (140)	Multicentre, double-blind, RCT	CZP 200 mg, every other week, N=46 CZP 400 mg, every four weeks, N=51 PBO N=50	12	NR	CZP 200 mg arm: 36.6 (13) CZP 400mg arm: 37.5 (10.8) PBO arm: 38 (11.8)	CZP 200 mg arm: 4.6(6.7) CZP 400mg arm: 5.9(6.8) PBO arm: 4.2(4.7)	CZP 200 mg arm: 8.5(9.1) CZP 400mg arm: 8.1(6.8) PBO arm: 9.2 (10)	CZP 200 mg arm: 39 (85) CZP 400mg arm: 47 (92) PBO arm: 40 (80)	CZP 200 mg arm: 38 (83) CZP 400 mg arm : 44 (86) PBO arm: 41 (82)

Abbreviations: ADA, adalimumab; ASAS, Assessment of Spondyloarthritis international Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CSR, Clinical Study Report; CZP, certolizumab pegol; ETN, etanercept; FDA, Food and Drug Administration; IFN, infliximab; NR, not reported; nr-axSpA, non-radiographic axial spondyloarthritis; NSAID, non-steroidal anti-inflammatory drug; PBO, placebo; PGADA, Patient global assessment of disease activity; RCT, randomised controlled trial; SD, standard deviation; VAS, visual analogue scale.

† While criteria used in this study are identical to ASAS criteria they are not specified as ASAS criteria

‡ Scores on a 0-10-cm visual analogue scale

Table 73: Baseline disease status in etanercept and comparator treatment RCTs in the nr-axSpA population eligible for inclusion in the NMA

Publication & trial name	Treatment arms	Smoking Status	HLA B27 positive: n (%)	Elevated CRP at baseline: n (%)	Baseline BASDAI:	Baseline BASFI:	MRI SI positive: n (%)	SPARCC MRI SIJ score
Non-radiographic axial spondyloarthritis								
Etanercept Studies								
Study1031, (CSR-205791) (141) Sep 2013 Primary analysis of primary outcome	ETN arm PBO arm	ETN arm: NR PBO arm: NR	ETN arm: 71(66.98) PBO arm: 83 (76.15)	ETN arm: 48(45.28) PBO arm: 44 (40.37)	ETN arm: [redacted] † PBO arm: [redacted] †	ETN arm: [redacted] † PBO arm [redacted] †	ETN arm: 87 (82.08) PBO arm: 87 (79.82)	ETN arm: [redacted] † PBO arm: [redacted] †
Study1031, (CSR 205791-(161) Feb 2014 Updated analysis on open-label extension	ETN arm PBO arm	ETN arm: NR PBO arm: NR	ETN arm: 71(66.98) PBO arm: 83(76.15)	ETN arm: 48(45.28) PBO arm: 44 (40.37)	ETN arm: [redacted] † PBO arm: [redacted] †	ETN arm: [redacted] † PBO arm: [redacted] †	ETN arm: 87 (82.08) PBO arm: 87 (79.82)	ETN arm: [redacted] † PBO arm: [redacted] †
Dougados 2013 (abstract) (128) Primary analysis data-cut	ETN arm PBO arm	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR
Dougados 2013 (abstract) (185) Primary analysis data-cut	ETN arm PBO arm	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: 6(1.8) † PBO arm: 6(1.9) †	ETN arm: 4.2(2.5) † PBO arm: 3.9(2.5) †	ETN arm: NR PBO arm: NR	ETN arm: 8.0 (9.7) † PBO arm: 7.8(10.2) †
Dougados 2013 (abstract) (126) Primary analysis data-cut	ETN arm PBO arm	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR	ETN arm: NR PBO arm: NR
Adalimumab studies								
Sieper, 2013 ABILITY-1 (112)	ADA arm PBO arm	ADA arm: NR PBO arm: NR	ADA arm: 75(82) PBO arm: 70(74)	ADA arm: 29 (32) PBO arm: 37	ADA arm: 6.4 (1.5) † PBO arm: 6.5	ADA arm: 4.5 (1.9) † PBO arm: 4.9 (2.3)	ADA arm: NR PBO arm: NR	ADA arm: 5.1 (9.5) † PBO arm: 4.7

				(39)	(1.6) †	†		(9.9) †
FDA 2013 ABILITY-1 (139)	ADA arm PBO arm	ADA arm: NR PBO arm: NR	ADA arm: 75 (82) PBO arm: 70 (74)	ADA arm: 36(39.6) PBO arm: 36(38.3)	ADA arm: 6.4(1.5) † PBO arm: 6.5(1.6) †	ADA arm: NR PBO arm: NR	ADA arm: NR PBO arm: NR	ADA arm: 64(70) ‡ PBO arm:64(68) ‡
Haibel, 2008 (105)	ADA arm PBO arm	ADA arm: NR PBO arm: NR	ADA arm: 13 (59) PBO arm: 18 (75)	ADA arm: NR PBO arm: NR	ADA arm: 6.5(1.2) † PBO arm: 6.2 (1.3) †	ADA arm: 5.4 (2) † PBO arm: 4.9 (1.6) †	ADA arm: 12(55) PBO arm: 18 (75)	ADA arm: NR PBO arm: NR
Certolizumab pegol studies								
Landewe, 2012 RAPID-axSPA (full paper) (90)	CZP 200 mg arm CZP 400 mg arm PBO arm	CZP 200 mg arm: NR CZP 400 mg arm: NR PBO arm: NR	CZP 200 mg arm: 34(73.9) CZP 400 mg arm: 37(72.5) PBO arm: 39(78)	CZP 200 mg arm: NR CZP 400mg arm: NR PBO arm: NR	CZP 200 mg arm: 6.5(1.4) † CZP 400 mg arm: 6.6(1.6) † PBO arm: 6.4(1.5) †	CZP 200 mg arm: 4.8(2.2) † CZP 400 mg arm: 5.1(2.4) † PBO arm: 4.9(2.2) †	CZP 200 mg arm: NR CZP 400 mg arm: NR PBO arm: NR	CZP 200 mg arm: NR CZP 400 mg arm: NR PBO arm: NR
FDA 2013 RAPID-axSPA (140)	CZP 200 mg arm CZP 400mg arm PBO arm	CZP 200 mg arm: NR CZP 400mg arm: NR PBO arm: NR	CZP 200 mg arm: 34(73.9) CZP 400 mg arm: 37(72.5) PBO arm: 39(78)	CZP 200 mg arm: 99 (89.2) CZP 400mg arm: 94 (87.9) PBO arm: 94 (87.9)	CZP 200 mg arm: 6.4(1.4) † CZP 400 mg arm: 6.6(1.6) † PBO arm: 6.4(1.4) †	CZP 200 mg arm: 4.8(2.2) † CZP 400 mg arm: 5.1(2.4) † PBO arm: 4.9(2.2) †	CZP 200 mg arm: NR CZP 400mg arm: NR PBO arm: NR	CZP 200 mg arm: NR CZP 400mg arm: NR PBO arm: NR

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C-reactive protein; CSR, Clinical Study Report; CZP, certolizumab pegol; ETN, etanercept; FDA, Food and Drug Administration; HLA B27, Human Leukocyte Antigen B27; MRI SI, Magnetic Resonance Imaging of sacroiliac joint; mSASSS, modified Stoke Ankylosing Spondylitis Spine Score; NR, not reported; nr-axSpA, non-radiographic axial spondyloarthritis; PBO, placebo; SD, standard deviation; SPARCC MRI SIJ, The Spondyloarthritis Research Consortium of Canada (SPARCC) Magnetic Resonance Imaging (MRI) Sacroiliac Joint (SIJ).
† Mean (SD); ‡ N (%)

Etanercept

Three conference abstracts reported on one double-blind RCT of patients with nr-axSpA administered etanercept 50 mg once weekly on a background of NSAIDS versus placebo (126, 128 , 185). One publication reporting on this RCT (68) was published after the electronic database search update was conducted so was therefore obtained directly from Pfizer's internal database. This publication was not considered further for the NMA due to the timing of publication, and since full details of this trial were already extracted using data from the clinical study reports as described in section 4.2.3.

Adalimumab

Two publications reported on two double-blind RCTs of adalimumab 40 mg administered every other week versus placebo (105, 112). An FDA report provided additional information on the Sieper et al, 2013, ABILITY-1 RCT (139).

Quality assessment

A quality assessment for risk of bias was carried out for all RCTs in the AS and nr-axSpA indications for which there were full-text publications available and which were eligible for data extraction (Appendix D).

Assessment of objective signs of inflammation in nr-axSpA trial as indicated by MRI positivity and CRP elevation

The market authorisation for anti-TNFs licensed for nr-axSpA specify that patients should display objective signs of inflammation as indicated by elevated C-reactive protein (CRP) and/or magnetic resonance imaging (MRI) evidence. This section summarises the extent to which the ITT population of the trials relate to the licensed population, providing an analysis of the proportion of patients in each trial who adhere to the licensed population.

-

Table 74 provides an analysis of patients who displayed elevated or normal CRP levels and/or MRI positive indication for sacroiliitis in the etanercept double-blind RCT (126, 128 , 185) and the adalimumab double-blind RCT (105, 112). Of note is that 12% of patients enrolled in the etanercept RCT and 32% enrolled in the adalimumab RCT had both a normal CRP and were MRI negative for sacroiliitis at baseline, indicating that they are not included in the licensed population. These patients may have a reduced response to anti-TNF alpha treatment in comparison to patients who had either a positive MRI for sacroiliitis (81% in etanercept RCT; 48% in adalimumab RCT) or elevated CRP (43% in etanercept RCT; 39% in adalimumab RCT) at baseline screening. It should be noted in reference to the 1031 study:

- Analyses on MRI positivity and CRP level were prespecified, but for purposes of providing evidence on comparative clinical efficacy, the results of ITT analyses were used for the study population, as this maintains randomisation, and for this reason a small number of patients who did not meet the licenced population were included in the study results.
- The MRI negative group is approximately 20% of all patients. Therefore, conclusions from analyses on this population are not powered to detect differences between etanercept and placebo.
- In addition, the definition of elevated CRP varies between trials. In the case of study 1031, this was >3 mg/L. However, for the adalimumab ABILITY-1 study (112), the

definition of elevated CRP was a CRP concentration above upper normal limit, with the latter value undefined in the publication.

Table 74: Patients with evidence of nr-axSpA symptoms at baseline in etanercept study 1031 and adalimumab ABILITY-1 trials

Study		Normal CRP (<3 mg/L)	Elevated hsCRP (>3 mg/L)
Study1031, (CSR-205791) (141) Sep 2013 Primary analysis of primary outcome	**MRI negative for sacroiliitis at screening**	ETN: [redacted] PBO: [redacted]	ETN: [redacted] PBO: [redacted]
Study1031, (CSR-205791) (141) Sep 2013 Primary analysis of primary outcome	**MRI positive for sacroiliitis at screening**	ETN: [redacted] PBO: [redacted]	ETN: [redacted] PBO: [redacted]
Sieper, 2013 ABILITY-1 (112)	**MRI negative for sacroiliitis at screening**	ADA: 30/91 (32.97%) PBO: 29/94 (30.85%)	ADA: 15/91 (16.48%) PBO: 19/94 (20.21%)
Sieper, 2013 ABILITY-1 (112)	**MRI positive for sacroiliitis at screening**	ADA: 25/91 (27.47%) PBO: 28/94 (29.78%)	ADA: 21/91 (23.07%) PBO: 15/94 (15.96%)

Abbreviations: ADA, adalimumab; CRP, C-reactive protein; CZP, certolizumab pegol; ETN, etanercept; MRI, magnetic resonance imaging; PBO, placebo.

While a similar breakdown of the above nr-axSpA signs of inflammation criteria is not reported for the certolizumab pegol RCT, Table 75 indicates that both the etanercept 1031 and adalimumab ABILITY -1 trials included relatively high numbers of patients (range: 57%-62%) with normal baseline CRP levels, whereas the proportion was much smaller in the certolizumab pegol trial (12%). Patients without elevated CRP, i.e. normal CRP, are likely to have a lower response to anti-TNF therapy, so inclusion of these patients in the etanercept and adalimumab trials is likely to reduce the observed efficacy of these treatments in these specific trial populations. It should be noted that although only a small proportion of patients in the certolizumab pegol study had sacroiliitis detected by MRI, a large proportion (certolizumab 200 58.6%, certolizumab 400 52.3%, and placebo 53.3%) also had sacroiliitis detected by x-ray (140).

In the case of the certolizumab pegol RCT, all patients had to have CRP levels > upper limit of normal (ULN=7.9 mg/L) and/or sacroiliitis on MRI according to the ASAS/OMERACT definition (90). Therefore there was no population of patients who had both normal CRP and a negative MRI. As these are the patients broadly considered to demonstrate the smallest benefit from anti-TNF treatment, it would be expected that the RAPID-axSpA trial population would have a higher average response to anti-TNF treatment.

It should be noted that these and other baseline differences between the nr-axSpA trials discussed in more detail in section 4.2.6.2 indicate the presence of heterogeneity between the key nr-axSpA RCTs included in the NMA.

Table 75: Comparison of MRI sacroiliitis and elevated CRP levels between the three key nr-axSpA trials

Publication & trial name	MRI Sacroiliitis positive, N (%)	MRI Sacroiliitis negative, N (%)	Elevated CRP yes, N (%	Elevated CRP no, N (%)
Study 1031, (CSR-205791) (141) Sep 2013 Primary analysis of primary outcome	ETN: 87 (82.08) PBO: 87 (79.82)	ETN:19 (17.92) PBO: 22 (20.18)	ETN: 48 (45.28) PBO: 44 (40.37)	ETN: 58 (54.72) PBO: 64 (58.72)
Sieper, 2013 ABILITY-1 (112)	ADA: 46 (50.5) PBO: 43 (45.7)	ADA: 45 (49.5) PBO: : 48 (51.1)	ADA: 36 (39.6) PBO: 36 (38.3)	ADA: 55 (60.4) PBO: 58 (61.7)
FDA 2013 RAPID-axSpA (140)	CZP 200: 32 (28.8) CZP 400: 35 (32.7) PBO: 31 (29.0)	CZP 200: NR CZP 400: NR PBO: NR	CZP 200: 99 (89.2) CZP 400: 94 (87.9) PBO: 94 (87.9)	CZP 200: 12 (10.8) CZP 400: 13 (12.1) PBO: 13 (12.1)

Abbreviations: ADA, adalimumab; CRP, C-reactive protein; CZP, certolizumab pegol; ETN, etanercept; PB, placebo.

4.2.3 Additional data sources for the NMA

In addition to the publication and conference abstract data identified from the literature search, additional sources of RCT data were identified which were relevant to the NMA. First, the manufacturer clinical study reports (CSRs) were identified for five of the etanercept AS RCTs (10-12, 14, 93) and for the single etanercept nr-axSpA RCT (Table 72) to supplement the conference abstracts which reported data on this study (126, 128, 185). Where required, the CSRs were used to supplement study design, patient characteristics and outcomes data required to assess the feasibility of conducting an NMA, and for the conduct of the NMA itself.

For the certolizumab pegol RCT (90), additional outcomes data (e.g. standard error, standard deviation or mean arm change from baseline) were identified from the Cimzia EMA Assessment report (186). For two adalimumab RCTs, (91, 107), and the golimumab RCT (69) additional outcomes data were sourced from the NICE TA233 ERG report (187).

Studies considered eligible for inclusion in the systematic review but which on further scrutiny were not eligible for inclusion in the NMA due to lack of relevant outcomes data or other reasons are listed in Appendix E with reasons for their exclusion.

4.2.4 NMA feasibility assessment

In order to gauge the appropriateness of proceeding with an NMA it was first necessary to perform an assessment of: 1) the RCT evidence for the interventions of interest to ascertain whether they do form one evidence network for each population and outcome of interest; and 2) the distribution of study and patient characteristics that may affect treatment effects across direct comparisons of the evidence networks.

For the current NMA there was no information on which patient characteristics are likely to be effect modifiers. Previous analyses have indicated that some patient characteristics may be prognostic factors, e.g. age, duration of disease, functional status at initiation of therapy

and elevated levels of CRP at baseline (see section 1.5.1). In addition, there is some evidence from the etanercept 1031 trial that CRP and sacroiliac joint SPARCC score may be treatment effect modifiers (see section 3.3.1.3). However, in general there is a paucity of published evidence on the extent to which prognostic factors act as treatment effect modifiers.

Feedback from clinical experts consulted on the matter indicates there is the potential that some of the prognostic factors could act as effect modifiers. In particular, they identified CRP, Spinal/sacroiliac joint MRI score, HLA-B27, symptom duration, and age as potential effect modifiers. On this basis, it is conservative to assume that all of the baseline characteristics identified in section 4.2.1.1 may have an impact on efficacy. Analysis for the economic evaluation based on study 1031 and 314-EU suggested that CRP, HLA-B27, gender (nr-axSpA only), and age (AS only) were significant predictors of response for BASDAI and BASFI. These findings are presented in section 5.4.2.1.

4.2.4.1 Etanercept studies

Of the ten AS RCT publications eligible for inclusion in the NMA based on availability of outcomes data (10-14, 93, 98, 101, 104, 110), three were excluded at the feasibility assessment stage for the following reasons:

- RCT compared etanercept 50 mg once weekly to an unlicensed dose of etanercept 50 mg every other week (101)
- RCT compared etanercept 50 mg once weekly to an unlicensed dose of etanercept 50 mg twice weekly (110)
- RCT compared etanercept 50 mg once weekly to sulfasalazine (the latter is not a relevant comparator for the decision problem) (14)

In addition, a further publication was excluded (98) since it reported the same results from an RCT as the initial main publication (12).

The 2011 SPINE RCT (13) reported on a population with advanced and severe spinal ankylosis. The population from the SPINE trial differed from the other etanercept RCTs in several ways. The advanced stages of disease of these patients meant they were generally older and had disease symptoms for a longer time than patients from other trials; disease duration in the SPINE trial was double the disease duration reported in other RCTs (20 vs 10 years, respectively). For this reason, this study was excluded from the base case analysis, and was included as a sensitivity analysis for transparency and to assess its impact on the results.

Note that a sixth study which compared etanercept 50 mg once weekly to infliximab 5 mg was excluded from the base case NMA, but included as a sensitivity analysis as it was an open-label trial rather than a double-blind RCT (104). All studies included in the base-case NMA were double-blind RCTs.

For purposes of the NMA, studies using etanercept 25 mg twice weekly and 50 mg once weekly were combined into one group since these doses were shown to be as effective as one another, with a similar safety profile, in a large Phase III RCT (12). This may reduce the uncertainty in the NMA for indirect comparison estimates including etanercept.

In summary, four publications reporting on four etanercept RCTs were included in the base-case NMA analyses (10-12, 93). These were placebo-controlled comparisons to etanercept

25 mg twice weekly, and/or 50 mg once weekly. Therefore, an indirect comparison was required to compare etanercept to 1) adalimumab, 2) golimumab, and 3) certolizumab pegol.

Table 76-Table 78 shows the mean arm change from baseline results for these RCTs for the outcomes of CRP, BASFI and BASDAI at 12 weeks and which were used as the raw data for the NMA. Note that the impact of excluded etanercept studies on the NMA results for AS are discussed in Appendix E.

For the nr-axSpA indication, data was provided from one RCT of etanercept 50 mg once weekly versus placebo, and was obtained where required from the CSRs (141, 161) where there was insufficient data in the publicly available conference abstracts (126, 128 , 185). The 2013 CSR reports adjusted means for all relevant outcomes at 12 weeks, while the 2014 CSR (includes results from a 48-week open-label extension) reports unadjusted means for the double-blind 12 week data to compare with the 48 week unadjusted data.

Both adjusted and unadjusted means (arm change from baseline) are shown in Table 80 and Table 81 for the outcomes of ASDAS-CRP, BASFI and BASDAI at 12 weeks. However, only the unadjusted mean data was used in the NMA for consistency with the data reported in the other comparator treatment RCTs.

4.2.4.2 Infliximab studies

Of the three AS publications reporting on two double-blind RCTs of infliximab versus placebo eligible for inclusion in the NMA, based on availability of outcomes data (95, 106, 108), the two publications of an RCT comparing low-dose infliximab 3 mg to placebo (106, 108) were excluded from consideration at the feasibility assessment stage since this dose is not licensed in the UK.

Table 76-Table 78 show the mean arm change from baseline results for the single infliximab 5 mg versus placebo RCT (95) for the outcomes of CRP, BASFI and BASDAI at 12 weeks, and which were used as the raw data for the NMA.

For the nr-axSpA indication, there were no RCTs of infliximab identified by the systematic review; infliximab is not licensed in nr-axSpA so is not a relevant comparator for this indication.

4.2.4.3 Adalimumab studies

Of the three AS publications reporting on two double-blind adalimumab RCTs versus placebo, (70, 91, 107) one publication was excluded (70) since it duplicated the data reported in the earlier main publication (91).

The mean arm change from baseline results for the two adalimumab 40 mg versus placebo RCTs (91, 107) for the outcomes of CRP, BASFI and BASDAI at 12 weeks, which were used as the raw data for the NMA are shown in Table 76-Table 78.

For the nr-axSpA indication, two publications reporting on two double-blind RCTs of adalimumab 40 mg versus placebo contributed outcomes data to the NMA (105, 112).

Table 79-Table 81 shows the mean arm change from baseline for the two placebo-controlled adalimumab RCTs for the outcomes of ASDAS-CRP, BASFI and BASDAI at 12 weeks which were used as the raw data for the NMA.

4.2.4.4 Certolizumab pegol studies

One double-blind RCT investigating certolizumab pegol 200 mg administered every two weeks versus certolizumab pegol 400 mg every four weeks versus placebo in patients with AxSpA, included outcomes data relevant to the NMA for patient sub-groups with AS and nr-axSpA (90). Since the response rates in the AS and nr-axSpA populations were similar for the certolizumab pegol 200 mg and 400 mg dose arms in this RCT, these arms were combined in the NMA for both indications.

Table 76-Table 78 shows the mean arm change from baseline results for the certolizumab pegol RCT AS sub-group for the outcomes of CRP, BASFI and BASDAI at 12 weeks. Table 79-Table 81 shows the mean arm change from baseline results for the certolizumab pegol RCT nr-axSpA sub-group for the outcomes of ASDAS-CRP, BASFI and BASDAI at 12 weeks which were used as the raw data for the NMA.

4.2.4.5 Golimumab studies

One study publication reporting on a double-blind RCT comparing golimumab 50 mg versus 100 mg administered every 4 weeks versus placebo in an AS population included outcome data relevant to the NMA (69). Table 76-Table 78 shows the mean arm change from baseline results for the placebo-controlled golimumab trial for the outcomes of CRP, BASFI and BASDAI at 12 weeks which were used as the raw data for the NMA.

Table 76: AS studies - CRP outcome

Ankylosing Spondylitis: Efficacy CRP						
Study details					**CRP arm level change from baseline (mg/dL)**	
Author, year	**Treatment arms**	**No of patients randomised**	**No of patients analysed**	**Follow-up period, (weeks)**	**Mean**	**SE**
Davis, 2003 (10)	ETN	138	138	12	-1.3†	0.17†
	Placebo	139	139	12	0.00†	0.20†
Calin, 2004 (11)	ETN	45	45	12	-1.34†	0.31†
	Placebo	39	39	12	0.00†	0.41†
van der Heidje, 2006 (12)	ETN 50 mg once weekly	155	155	12	-1.40†‡	0.13†‡
	ETN 25 mg twice weekly	150	150	12	-1.40†‡	0.13†‡
	Placebo	51	51	12	-0.47†	0.33†
Dougados, 2011 (13)	ETN	39	39	12	-16	0.44†
	Placebo	43	43	12	-1	0.29†
van der Heidje 2006 (91)	ADA	208	208	12	-1.3	0.10†
	Placebo	107	107	12	-0.1	0.10†
Inman, 2008 (69)	GOL 50 mg	138	138	14	-0.70†	0.13†
	GOL 100 mg	140	140	14	-0.50†	0.14†
	Placebo	78	78	14	0.00†	0.09†

Abbreviations: ADA, adalimumab; CRP, C-reactive protein; CZP, Certolizumab Pegol; ETN, etanercept; GOL, golimumab; NR, not reported; PBO, placebo; SE, standard error.
†calculated; ‡ pooled ETN values for 50mg and 25 mg doses

Table 77: AS studies - BASFI outcome

Ankylosing Spondylitis: Efficacy BASFI						
Study details					**BASFI arm level change from baseline**	
Author, year	**Treatment arms**	**No of patients randomised**	**No of patients analysed**	**Follow-up period, (weeks)**	**Mean**	**SE**
Davis, 2003 (10)	ETN	138	138	12	-1.67†	0.20†

	Placebo	139	139	12	-0.33†	0.21†
Gorman, 2002 (93)	ETN	20	20	12	-2.30†	0.46†
	Placebo	20	20	12	-0.10†	0.59†
Calin, 2004 (11)	ETN	45	45	12	-2.06†	0.33†
	Placebo	39	39	12	-0.33†	0.31†
van der Heidje, 2006 (12)	ETN 50 mg once weekly	155	155	12	-2.68†¶	0.13†¶
	ETN 25 mg twice weekly	150	150	12	-2.68†¶	0.13†¶
	Placebo	51	51	12	-1.07†	0.31†
Dougados, 2011 (13)	ETN	39	39	12	-2.2	0.28†
	Placebo	43	43	12	-1	0.27†
Giardina, 2010 (104)	ETN	25	25	12	-1.5†	0.45†
	INF	25	25	12	-2.6†	0.45†
Braun, 2002 (95)	INF	35	34	12	-2†	0.33†
	Placebo	35	35	12	-0.1†	0.39†
Landewe, 2012 (90)	CZP 200 mg	65	65	12	-1.72‡¶	0.23¶
	CZP 400 mg	56	56	12	-1.72‡¶	0.23‡¶
	Placebo	51	NR	12	-0.58‡	0.31†
Lambert 2007 (107)	ADA	38	NR	12	-2§	0.31†
	Placebo	44	NR	12	-0.1§	0.37†
van der Heidje 2006 (91)	ADA	208	208	12	-1.86§	0.15†
	Placebo	107	107	12	-0.45§	0.21†
Inman, 2008 (69)	GOL 50 mg	138	138	14	-1.62†	0.19†

	GOL 100 mg	140	140	14	-1.73†	0.182†
	Placebo	78	78	14	0.28†	0.19†

Abbreviations: ADA, adalimumab; BASFI, Bath Ankylosing Spondylitis Functional Index; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; NR, not reported; PBO, Placebo; SE, standard error. †Calculated; ‡Source Cimzia EMA assessment report (186); §TA233 ERG assessment report (187); ¶ pooled ETN values for 50mg and 25 mg doses

Table 78: AS studies - BASDAI outcome

Ankylosing Spondylitis: Efficacy BASDAI						
Study details					**BASDAI arm level change from baseline**	
Author, year	**Treatment arms**	**No of patients randomised**	**No of patients analysed**	**Follow-up period, (weeks)**	**Mean**	**SE**
Davis 2003 (10)	ETN	138	138	12	-2.36†	0.19†
	Placebo	139	139	12	-0.45†	0.18†
Calin 2004 (11)	ETN	45	45	12	-2.72†	0.34†
	Placebo	39	39	12	-0.85†	0.35†
van der Heidje 2006 (12)	ETN 50 mg once weekly	155	155	12	-3.25†‡‡	0.13†‡‡
	ETN 25 mg twice weekly	150	150	12	-3.25†‡‡	0.13†‡‡
	Placebo	51	51	12	-1.51†	0.32†
Dougados 2011 (13)	ETN	39	39	12	-2.6	0.32†
	Placebo	43	43	12	-1.4	0.30†
Giardina 2010 (104)	ETN	25	25	12	-1†	0.49†
	INF	25	25	12	-3†	0.48†
Braun 2002 (95)	INF	35	34	12	-3.2†	0.30†
	Placebo	35	35	12	-0.60†	0.24†
Landewe 2012 (90)	CZP 200 mg	65	65	12	-2.47‡†§§	0.20†
	CZP 400 mg	56	56	12	-2.47‡†§§	0.20†
	Placebo	51	57	12	-1.02‡	0.30†
van der	ADA	208	208	12	-2.6	0.20†

Heidje 2006 (91)	Placebo	107	107	12	-0.8	0.20†
Inman 2008 (69)	GOL 50 mg	138	138	14	-2.94†† †	0.19†
	GOL 100 mg	140	140	14	-2.84†	0.19†
	Placebo	78	78	14	-0.76†† †	0.22†

Abbreviations: : ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; NR, not reported; PBO, placebo; SE, standard error.
†Calculated; ‡source Cimzia EMA assessment report (186); §TA233 ERG assessment report (187); ††From manufacturer (Schering-Plough Ltd) submission to NICE STA;‡‡ pooled ETN values for 50mg and 25 mg doses; §§ pooled CZP values for 200mg and 400mg doses

Table 79: nr-axSpA studies – ASDAS-CRP outcome

nr-axSpA: Efficacy ASDAS-CRP						
Study details					**CRP arm level change from baseline (mg/dL)**	
Author, year	**Treatment arms**	**No of patients randomised**	**No of patients analysed**	**Follow-up period (weeks)**	**Mean**	**SE**
Study 1031, 2013†† (141)	ETN	111	NR	12	-1.1‡	0.12
	Placebo	114	NR	12	-0.49‡	0.11
Study 1031, 2014†† (161)	ETN	111	105	12	[redacted]	[redacted]
	Placebo	114	109	12	[redacted]	[redacted]

Abbreviations: ASDAS, Ankylosing Spondylitis Disease Activity Score; CRP, C-reactive protein; ETN, etanercept; NR, not reported; PBO, placebo; SE, standard error.
†calculated; ‡source CSR 205791 (141); § ASDAS CRP arm level change from baseline (mg/L); ††Study 1031 2013 (141) reports adjusted means and Study 1031 2014 (161) reports unadjusted means.

Table 80: nr-axSpA studies – BASFI outcome

nr-axSpA: Efficacy BASFI						
Study details					**BASFI arm level change from baseline**	
Author, year	**Treatment arms**	**No of patients randomised**	**No of patients analysed**	**Follow-up period (weeks)**	**Mean**	**SE**
Study 1031, 2013[†‡‡] (141)	ETN	111	105	12	-1.41[‡]	0.24[†]
	Placebo	114	108	12	-0.84[‡]	0.23[†]
Study 1031, 2014[††] (161)	ETN	111	105	12	[redacted]	0.21[†]
	Placebo	114	109	12	[redacted]	0.16[†]
Sieper 2013 (112)	ADA	91	90	12	-1.07[†]	0.25[†]
	Placebo	94	94	12	-0.59[†]	0.24[†]
Haibel 2008 (105)	ADA	22	22	12	-2.4[†]	0.39[†]
	Placebo	24	24	12	-0.8[†]	0.38[†]
Landewe 2012 (90)	CZP 200 mg	46	46	12	-2.28[§‡‡†]	0.28[†]
	CZP 400 mg	51	51	12	-2.28[§‡‡†]	0.28[†]
	Placebo	50	50	12	-0.4[§]	0.35[†]

Abbreviations: ADA, adalimumab; BASFI, Bath Ankylosing Spondylitis Functional Index; CZP, certolizumab pegol; ETN, etanercept; NR, not reported; PBO, placebo; SE, standard error.
†calculated; ‡adjusted mean; §From Cimzia EMA assessment report (186); ††Study 1031, 2013 (141) reports adjusted means and Study 1031, 2014 (161) reports unadjusted means. ‡‡ pooled CZP values for 200mg and 400mg doses

Table 81: nr-axSpA studies – BASDAI outcome

nr-axSpA: Efficacy BASDAI						
Study details					**BASDAI arm level change from baseline**	
Author, year	**Treatment arms**	**No of patients randomised**	**No of patients analysed**	**Follow-up period (weeks)**	**Mean**	**SE**
Study 1031, 2013[††] (141)	ETN	111	105	12	-1.96[‡]	0.28[†]
	Placebo	114	108	12	-1.31[‡]	0.27[†]
Study 1031, 2014[††] (161)	ETN	111	105	12	[redacted]	0.22[†]
	Placebo	114	109	12	[redacted]	0.18[†]
Sieper 2013 (112)	ADA	91	90	12	-1.9	0.29[†]
	Placebo	94	94	12	-1	0.29[†]
Haibel 2008 (105)	ADA	22	22	12	-2.7[†]	0.59[†]
	Placebo	24	24	12	-1.2[†]	0.56[†]
Landewe2012 (90)	CZP 200 mg	46	46	12	-3.36[§‡‡†]	0.30[†]
	CZP 400 mg	51	51	12	-3.36[§‡‡†]	0.30[†]
	ETN	111	105	12	-1.52[§]	0.36[†]

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CZP, certolizumab pegol; ETN, etanercept; NR, not reported; PBO, placebo; SE, standard error.
†calculated; ‡adjusted mean; §From Cimzia EMA assessment report (186); ††Study 1031, 2013 (141) reports adjusted means and Study 1031, 2014 (161) reports unadjusted means.

4.2.4.6 Overview of previous UK HTA anti-TNF indirect comparison evidence for AS and nr-axSpA

Previous NICE and SMC appraisals of the anti-TNF agents in AS have included indirect comparisons of etanercept and some of the other comparators of interest. NICE TA143 reported an indirect comparison of etanercept, adalimumab and infliximab in the AS indication (4), and found the following:

- No statistically significant differences in ASAS response rates
- A significantly greater mean reduction for BASDAI score for infliximab versus adalimumab and etanercept treatment at 12 weeks, whereas at 24 weeks there were no statistically significant differences between the three treatments
- For BASFI there was a significant difference between infliximab and adalimumab at 12 weeks and between infliximab and etanercept at 24 weeks.

The NICE Appraisal Committee (AC) noted that it was inappropriate to conclude there were differences in clinical effectiveness between the three treatments based on the results of this indirect comparison.

In NICE TA233 the manufacturer of golimumab conducted a mixed treatment comparison (MTC) that included an indirect comparison of the former with etanercept and adalimumab in the AS indication (5). The NICE AC noted the following with regard to the MTC results:

- Small but non-significant differences between golimumab and the other two TNF-inhibitors were identified for most outcomes

- An analysis conducted by the ERG indicated that golimumab was more efficacious than etanercept and adalimumab on BASDAI improvement, but the 95% CrIs around these estimates were wide and therefore the observed treatment effect was subject to a high degree of uncertainty
- The AC concluded that golimumab was comparable to the other TNF- inhibitors in terms of efficacy, adverse-event profile and risk of treatment discontinuation.

The Scottish Medicines Consortium (SMC) detailed advice for certolizumab pegol in the AxSpA population (including AS and nr-axSpA sub-populations) has recently been published (188). An MTC in the AS sub-population which included an indirect comparison of certolizumab pegol with adalimumab, golimumab and etanercept at Week 24 found no statistically significant differences between certolizumab pegol and the other comparator treatments for any outcomes (ASAS20, ASAS40, BASDAI and BASFI).

For the nr-axSpA sub-population, a fixed effect MTC enabled an indirect comparison of certolizumab pegol versus adalimumab at Week 12 and found:

- No statistically significant differences between treatments for outcomes with the exception of change in BASDAI, BASFI and patient global assessment of disease activity where certolizumab pegol had significant improvements compared to adalimumab
- Heterogeneity in the nr-axSpA patient populations between the studies.

The SMC concluded that certolizumab pegol was broadly comparable with comparators in the MTCs conducted in the AS and nr-axSpA sub-populations (188).

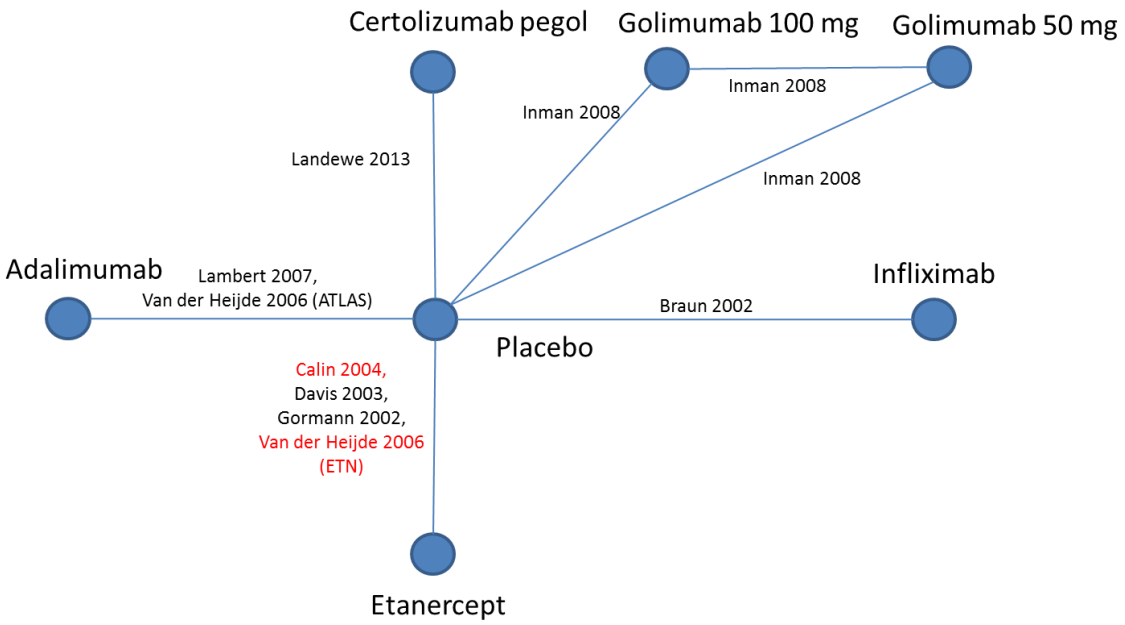
It appears from the indirect comparisons of the anti-TNF agents conducted to date in support of UK HTA submissions that these treatments are all similar in terms of clinical efficacy across both the AS and nr-axSpA populations.

4.2.5 Results of meta-analysis and indirect comparisons

4.2.5.1 Ankylosing Spondylitis

The set of eligible AS studies included 11 RCTs. The complete evidence network for the AS population is shown in Figure 5. Most comparisons were described by only one trial, except for the comparison of etanercept vs placebo (4 trials) and adalimumab vs placebo (2 trials).

Figure 5: Complete network of evidence of randomised clinical trials in the AS population



*Studies in red font indicate that imputation of SE data was required

As mentioned in section 4.2.4.1, the Dougados et al, 2011, SPINE RCT which enrolled a population with advanced disease, differed from the other etanercept RCTs in several ways. The advanced stages of disease of these patients meant they were generally older and had disease symptoms for a longer time than patients from other trials: patients in the spine trial had disease duration twice that of patients in other RCTs (20 vs 10 years, respectively). Therefore it was decided to remove this RCT from the base-case NMA analyses, a decision which was verified by clinical experts in a Pfizer Advisory Board. A sensitivity analysis that includes this study can be found in Appendix F.

For all three outcomes of interest, a connected network was available within the set of eligible studies. Two of the eligible studies (highlighted in red in Figure 5) had no variance measure reported for any outcomes; these studies required imputation of standard errors in order to be included in the analysis. The results of the NMA analyses conducted without imputation are presented in Appendix F. Network diagrams for NMA analyses with and without imputation of SE data are also presented in Appendix F. In addition, results of 1) the imputed analyses that include the Giardina open-label study (104) and 2) non-imputed and imputed analyses that include the SPINE RCT, are also presented in Appendix F as sensitivity analyses.

4.2.5.1 BASFI (with imputation of standard errors) NMA

Nine RCTs (N=1,758) comprising 19 treatment arms were included.

Table 82 presents the results obtained with both fixed and random effects NMA models. The DICs for the fixed effects and random effects models were very similar: 31.58 and 33.35, respectively for fixed and random effects models:

- All active treatments provided greater reductions in BASFI scores than placebo with the exception of certolizumab in the random effects model
- No other comparisons were found to be statistically important.

Table 82: Mean differences in BASFI score among drugs used to treat AS patients (with imputation of SEs)

Mean Differences (Fixed Effects)						
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	Etanercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	Infliximab	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	Certolizumab	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	Adalimumab	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Golimumab 100 mg	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Golimumab 50 mg
Mean Differences (Random Effects)						
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	Etanercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	Infliximab	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	Certolizumab	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	Adalimumab	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Golimumab 100 mg	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Golimumab 50 mg

Note: *Each cell represents the comparison (mean difference and 95% CrI) of the row treatment versus the column treatment; statistically meaningful results, i.e., comparisons where the 95% credible intervals do not overlap 0.00, are bold-faced. Positive mean differences represent superiority of the interventions (column treatment), negative mean differences represent superiority of the comparators (row treatment).*

4.2.5.2 Change in BASDAI score (with imputation of standard errors)

Eight trials (N=1,630) were used for the analysis of BASDAI scores (cm) in the NMA, including studies that required the imputation of SEs.

Table 83 presents the differences in CFB BASDAI (cm) obtained with fixed and random effects NMA models. The DICs for the fixed effects and random effects models were very similar: 25.51 and 27.39, respectively.

In the fixed effects model:

- All active treatments were found to produce significantly greater reductions in BASDAI scores versus placebo
- Infliximab had a significantly greater reduction in BASDAI score than certolizumab
- Other comparisons between active treatments were not found to be statistically significant

In the random effects model:

- Only etanercept was found to provide greater reductions in the BASDAI score than placebo

- Other comparisons between active treatments were not found to be statistically important

Table 83: Mean differences in BASDAI score among drugs used to treat AS patients (with imputation of SEs)

Mean Differences (Fixed Effects)						
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	**Etanercept**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	**Infliximab**	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	**Certolizumab**	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	**Adalimumab**	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**Golimumab 100 mg**	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**Golimumab 50 mg**
Mean Differences (Random Effects)						
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	**Etanercept**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	**Infliximab**	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	**Certolizumab**	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	**Adalimumab**	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**Golimumab 100 mg**	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**Golimumab 50 mg**

Note: *Each cell represents the comparison (mean difference and 95% CrI) of the row treatment versus the column treatment; statistically meaningful results, i.e., comparisons where the 95% credible intervals do not overlap 0.00, are bold-faced. Positive mean differences represent superiority of the interventions (column treatment), negative mean differences represent superiority of the comparators (row treatment).*

4.2.5.3 Change in CRP score (with imputation of standard errors)

Six trials (N=1,458) reporting CRP were included in the NMA. Results obtained with the random effects NMA model were considered too unstable and only results obtained with the fixed effects model are presented in Table 84. The fixed effects model results indicated:

- All active treatments offered significant reductions in CRP levels versus placebo
- Etanercept and adalimumab had significantly greater reductions in CRP levels versus golimumab (100 mg and 50 mg dose)
- No other comparisons were statistically important.

Table 84: Mean differences in CRP score among drugs used to treat AS patients (with imputation of SEs)

Mean Differences (Fixed Effects)					
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	Etanercept	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	Infliximab	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	Adalimumab	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	Golimumab 100 mg	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Golimumab 50 mg

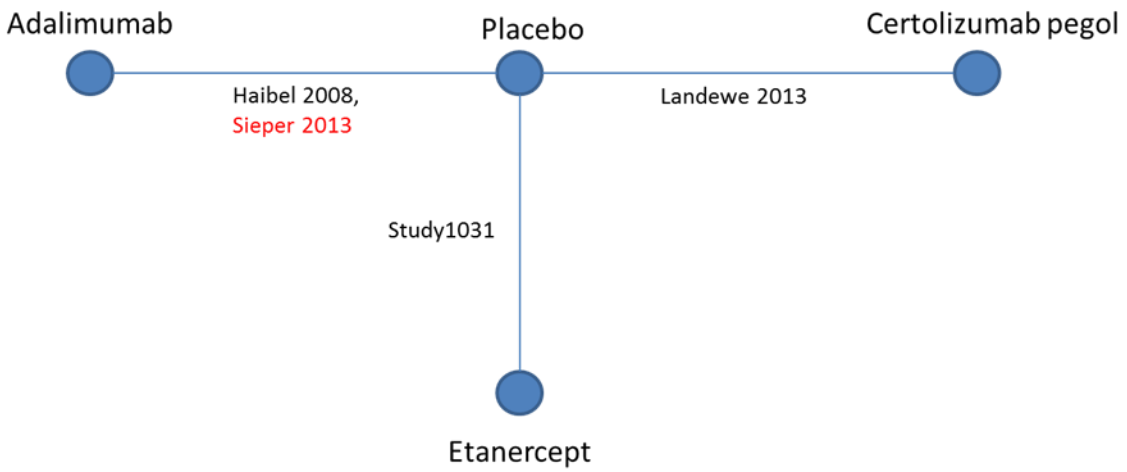
Note: *Each cell represents the comparison (mean difference and 95% CrI) of the row treatment versus the column treatment; statistically meaningful results, i.e., comparisons where the 95% credible intervals do not overlap 0.00, are bold-faced. Positive mean differences represent superiority of the interventions (column treatment), negative mean differences represent superiority of the comparators (row treatment).*

4.2.5.4 Non-radiographic axial spondyloarthritis

The set of eligible nr-axSpA studies included 4 RCTs as presented in Figure 6. The comparison of adalimumab vs placebo was assessed by two trials. One of the eligible studies (highlighted in red in Figure 6) had no variance measure reported for any outcome; this study required imputation of SEs in order to be included in the NMA.

In this population, only one RCT reported change in CRP levels, therefore no analysis was possible for this outcome among the nr-axSpA population. For BASFI and BASDAI, a connected network was available within the set of included studies. The results of the NMA analyses conducted without imputation are presented in Appendix F. Network diagrams for NMA analyses with and without imputation of SE data are also presented in Appendix F.

Figure 6: Complete network of evidence of randomised clinical trials in the nr-axSpA population*



*Studies in red font indicate that imputation of SE data was required.

4.2.5.5 Change in BASFI (with imputation of standard errors)

Four trials (N=604) reported BASFI score (cm) in the nr-axSpA population and were included in the NMA, including one study that required imputation of standard errors. There was not

sufficient data to obtain stable estimates with the random effects model, and results obtained with the fixed effects model are presented (Table 85). The fixed effects model found:

- All three active treatments produced significantly greater reductions than placebo
- Certolizumab had significant improvements compared to both etanercept and adalimumab
- No difference between etanercept and adalimumab.

Table 85: Mean differences in BASFI score among drugs used to treat nr-axSpA patients (with imputation of SEs)

Mean Differences (Fixed Effects)			
Placebo	[redacted]	[redacted]	[redacted]
[redacted]	Etanercept	[redacted]	[redacted]
[redacted]	[redacted]	Adalimumab	[redacted]
[redacted]	[redacted]	[redacted]	Certolizumab

Note: *Each cell represents the comparison (mean difference and 95% CrI) of the row treatment versus the column treatment; statistically meaningful results, i.e., comparisons where the 95% credible intervals do not overlap 0.00, are bold-faced. Positive mean differences represent superiority of the interventions (column treatment), negative mean differences represent superiority of the comparators (row treatment).*

4.2.5.6 Change in BASDAI score (with imputation of standard errors)

Four trials (N=604) describing BASDAI scores in the nr-axSpA population are presented in the NMA, including one trial that required imputation of SEs based on the other three studies. Results obtained with the random effects NMA model were deemed too unstable; only results obtained from the fixed effects model are presented (Table 86). The fixed effects model found:

- Etanercept, adalimumab and certolizumab had significantly greater reductions in BASDAI score versus placebo
- Etanercept had significantly lower improvement in BASDAI compared to certolizumab
- No other comparisons were statistically important

Table 86: Mean differences in BASDAI score among drugs used to treat nr-axSpA patients (with imputation of SEs)

Mean Differences (Fixed Effects)			
Placebo	[redacted]	[redacted]	[redacted]
[redacted]	Etanercept	[redacted]	[redacted]
[redacted]	[redacted]	Adalimumab	[redacted]
[redacted]	[redacted]	[redacted]	Certolizumab

Note: Each cell represents the comparison (mean difference and 95% CrI) of the row treatment versus the column treatment below the diagonal and of the column treatment versus the row treatment above the diagonal; statistically meaningful results, i.e., comparisons where the 95% credible intervals do not overlap 0.00, are bold-faced. Positive mean differences represent superiority of the interventions (column treatment), negative mean differences represent superiority of the comparators (row treatment).

4.2.6 Limitations of the NMA and indirect comparisons

4.2.6.1 Ankylosing Spondylitis

One study comparing etanercept 50 mg once weekly to infliximab 5 mg (104) over two years had an open-label design, and so was excluded from the base-case NMA since the other etanercept and comparator trials had a double-blind design. An NMA including this study is reported in Appendix F as a sensitivity analysis. When interpreting the results from the latter the following should be considered:

- It is possible that the study participant's perceptions of the medication they received in an open-label manner in this trial may have influenced responses on the patient reported outcomes of BASDAI and BASFI
- While the 12-week data from this study indicated some statistically significant differences in favour of infliximab on the BASDAI and BASFI measures, this was not sustained in the long-term in this study, and there were no significant differences between the two treatments at Week 104, suggesting that while infliximab may have a more rapid onset of action, it has no long-term efficacy advantage over etanercept.

The certolizumab RCT primary analysis was performed in an AxSpA population and all other endpoints were considered exploratory, including those on key outcomes conducted in the AS and nr-axSpA subpopulations as defined by local X-ray readings (90). When central X-ray readings were applied to these subpopulations, 21% of the original AS subgroup of patients were rediagnosed as nr-axSpA (37 patients were not centrally read). There were some further limitations to the definition of the two subpopulations which are described in more detail in the nr-axSpA NMA section in section 4.2.6.2.

There were also some differences in baseline characteristics in the four etanercept (10-12, 93) and five comparator RCTs (infliximab 5mg versus placebo (95); adalimumab 40 mg versus placebo RCTs (91, 107); golimumab 50 and 100 mg vs placebo (69); certolizumab pegol 200/400 mg vs placebo (90). In particular, where this information was reported in studies, there were variations in:

- Age: The average age ranged from 39 to 43 years
- Gender: One infliximab vs placebo study (95) reported 65.5% of its population as male, compared to all other studies in the NMA that reported male gender in excess of 70%, reaching 78% in two studies (11, 107)
- Disease duration: Varied from 5.87 years (69) to 15.65 years (95)
- Prior NSAID use: Varied from 81.2% (12) to 100% (91, 104, 107)
- Prior DMARD use: Varied from 37.8% (12) to 100% (91)
- Concomitant DMARD use: Varied from 19.9% (91) to 39.3% (12)
- Concomitant NSAID use: 47.5% (93) in one etanercept RCT with the remaining studies reporting usage between 79% (91) to 90% (69)
- History of uveitis: Varied from 9.8% (12) to 46.5% (95)
- CRP: Lower in the golimumab RCT with a median 1.05 mg/dl (69) compared to the remaining studies where the mean value ranged between 1.95 mg/dl (91) and 2.15 mg/dl (11, 107)

- BASFI: Lower in one etanercept trial at 4.52 (93) compared to the remaining studies where the BASFI baseline total scores were all above 5 points
- Prior anti-TNF use: One study in an AxSpA population, (90) included some patients who had prior anti-TNF treatment, including 20.3% of the AS patient subgroup

The variation in the above baseline characteristics indicates there may be some heterogeneity between the studies in the AS NMA. However, covariate regression analysis could not be carried out since there were too few studies in each separate loop of the network to make this viable.

4.2.6.2 Non-radiographic axial spondyloarthritis

For the nr-axSpA indication, there were some variations in the baseline characteristics of the four RCTs comparing etanercept (161), adalimumab (105, 112) and certolizumab pegol (90) versus placebo included in the NMA. In particular, where this information was reported in studies, there were variations in:

- Age: The etanercept vs placebo RCT (161) reported a lower mean age of 31.5 years whereas in the other three studies the average age was over 36 years (90, 105, 112)
- Gender: The etanercept vs placebo RCT (161) reported a higher percentage of male participants (60.5%) whereas in the other three studies the proportion of male participants was lower than 50% (90, 105, 112)
- Disease duration: Varied from 2.45 years (161) to 7.5 years (90, 105, 112)
- Symptom duration: Varied from 2.5 (15) years to 10.1 years (112)
- Prior NSAID use: Varied from 85.7% (90) to 100% (105, 112, 161)
- Prior DMARD use: Varied from 20% (161) to 38% (90)
- Concomitant NSAID use: Higher in the etanercept RCT (100%) (161) than the comparator studies (range: 79%-83.7%) (90, 112)
- Concomitant DMARD use: Varied from 0% (105) to 28% (90)
- History of uveitis: Varied from 7.9% (161) to 12% (112)
- CRP: Higher in the certolizumab RCT (1.6 mg/dl) compared to the remaining studies where the range was between 0.66 mg/dl (161) and 0.72 mg/dl (112)
- BASFI: Lower in the etanercept RCT at 4.05 (161) compared to the remaining studies where the BASFI baseline total scores ranged from 4.7 (112) to 5.15 (105)
- Prior anti-TNF use: One study in an AxSpA population, (90) included some patients who had prior anti-TNF treatment, including 10.3% of the nr-axSpA patient sub-group

When compared to the assessments of the variations in baseline characteristics in the AS populations (section 4.2.6.1), there was considerably greater heterogeneity observed in the nr-axSpA populations. While covariate regression analysis was not feasible since there were too few studies in each separate loop of the network to make this viable, these findings warranted further analysis using MAIC and STC.

In general the etanercept RCT recruited participants who were younger, had shorter disease duration, and had lower mean BASFI compared to the participants in the other studies. As noted in the etanercept clinical RCT and OLE section in section 3, this was done to reduce the risk of including patients with AS in the nr-axSpA population. This risk was further

minimised by the use of centralised and blinded assessment of the radiograph for diagnosis of nr-axSpA. This approach was taken to avoid the situation in the adalimumab and certolizumab RCTs (90, 112) where considerable reclassification of nr-axSpA has subsequently been reported outside the RCT publications, thereby potentially confounding the results in the current nr-axSpA NMA.

The FDA assessment report of the adalimumab RCT (139) noted the following limitations:

- Of the 185 participants randomised with supposed nr-axSpA (diagnosed via local X-ray), 102 were centrally assessed, and 64/102 were diagnosed with nr-axSpA, and 38 with AS
- The results on key outcomes (ASAS 40, ASAS 20, BASDAI 50, ASAS partial remission, ASAS 5/6) in this nr-axSpA subgroup found no statistically significant differences between adalimumab vs placebo compared to the total clinical trial population and the AS subgroup where statistically significant differences were reported for key efficacy outcomes.

The data from this RCT publication (112) included in the current NMA are therefore confounded by the presence of a significant number of AS patients which is likely to invalidate the indirect comparison of adalimumab vs the other comparator treatments in the current nr-axSpA NMA and produce unreliable results. Since the FDA noted that the centrally-read nr-axSpA post-hoc sub-group may not preserve the baseline comparability between treatment arms achieved by randomisation (139), it was deemed inappropriate to conduct an NMA using this post-hoc analysis. The present NMA is therefore limited by the use of the heterogeneous data from the Sieper et al, 2013, trial publication (112).

Similarly, the FDA assessment report of the certolizumab RCT (90) noted the following limitations:

- Of the original N=147 nr-axSpA subgroup in the RCT publication (based on local X-ray), only 69 patients (49%) retained this diagnosis on central X-ray in a post hoc analysis (note that 37 patients were not assessed)
- Following central radiographic assessment, 21% of patients were reclassified from having AS to nr-axSpA
- While results in the central X-ray nr-axSpA subgroup were still significantly different vs placebo in line with the original publication, for some key outcomes (e.g. BASFI, BASDAI) the magnitude of effect was smaller than observed in the original nr-axSpA subgroup

The data from the nr-axSpA subgroup in the RCT publication (90) included in the current NMA are therefore confounded by the presence of a significant number of AS patients which is likely to invalidate the indirect comparison of certolizumab vs the other comparator treatments. Therefore, the results from the NMA are potentially unreliable. As with the adalimumab centrally-read nr-axSpA post-hoc subgroup, the present certolizumab nr-axSpA post-hoc analysis may also fail to preserve baseline comparability between treatment arms and randomisation is likely to be compromised. Therefore, it was deemed inappropriate to conduct an NMA using this post-hoc analysis. In summary, the present NMA is limited by the use of the heterogeneous data from the Landewe et al, 2012, trial publication (90).

Since only 21% of the original AS subgroup in the RCT of certolizumab pegol were reclassified on central x-ray compared to 49% of the original nr-axSpA subgroup (140) and

37% of the nr-axSpA population in the adalimumab RCT (112), the impact of confounding from subsequent reclassification is likely to be much higher in the smaller nr-axSpA evidence network than in the larger AS network. Furthermore, within the AxSpA classification, the definition of nr-axSpA is privative (i.e. nr-axSpA is defined by the absence of radiographic evidence for sacroiliitis). In such circumstances, it is expected that this group will demonstrate greater heterogeneity than a patient group defined by the presence of a characteristic (e.g. the AS population who must show radiographic evidence for sacroiliitis). For these reasons, additional analyses were undertaken to address these limitations in the nr-axSpA indication.

4.2.7 Additional analyses carried out in the nr-axSpA population

In light of the limitations to the nr-axSpA NMA highlighted above, additional analyses were undertaken to overcome the differences in patient characteristics and potential confounding due to reclassification of AS and nr-axSpA patients.

4.2.7.1 Introduction to matching-adjusted indirect comparison (MAIC) and simulated treatment comparison (STC) methods

While the NMA of the nr-axSpA network is not able to address the differences in baseline characteristics between trials, and is therefore potentially unreliable, the use of patient level data from the Pfizer sponsored 1031 trial affords an opportunity to adjust the indirect comparisons using techniques that can utilise this information. Specifically, in light of the challenges involved in carrying out a NMA to derive indirect comparisons of etanercept, adalimumab and certolizumab pegol in the nr-axSpA network, these comparisons were obtained using matching-adjusted indirect comparison (MAIC) (187) and simulated treatment comparison (STC) methods (188).

These novel techniques are based on a targeted comparison of outcomes observed in specific treatments of interest (unlike NMA, which includes data from all treatments in the therapeutic area). The validity of such analyses relies on the general compatibility of the trials with respect to design, inclusion criteria, and population profiles. Even in otherwise identically designed trials, however, it is possible that the characteristics of the patients differ (even just by chance). Differences in characteristics that are predictive of the outcomes of interest will confound the comparisons of treatments. MAIC and STC are designed specifically to address this issue by adjusting for baseline characteristics. This requires patient-level data on at least one of the trials (which is referred to as the index trial). The methods differ in the way they use the data and how they adjust for potential confounding. MAIC creates weights that, when applied to the patients in the index trial, will yield a profile that matches the comparator trial populations. In STC, the patient-level data are used to create predictive equations for the outcomes, which can then be used to predict outcomes in a population that matches the profile of the comparator trial. When the index and comparator trials used the same treatment (or placebo) in their reference groups (i.e. control), MAIC and STC can be used to assess the presence of any study effects, i.e. distortion in comparisons due to differences between studies beyond population characteristics. This is assessed based on difference in outcomes in the reference arms, after adjustment for patient characteristics.

We applied MAIC and STC to assess the relative effectiveness of etanercept vs adalimumab and certolizumab pegol on changes in BASDAI and BASFI at 12 weeks in patients with nr-axSpA. As the MAIC and STC techniques are targeted comparisons (i.e. a single comparison is made between two individual trials), specific trials were selected for

comparison. These were the etanercept 1031 (126, 161) and 314-EU trials (12), the ABILITY-1 trial for adalimumab (112) and the RAPID-axSpA trials for certolizumab pegol (90). In the case of the comparison between etanercept and adalimumab, the ABILITY-1 trial was selected for use in the analysis over the trial published by Haibel et al, (105) as this was the larger and more recent trial performed for adalimumab in the nr-axSpA population, and utilised the ASAS criteria for including nr-axSpA in the trial. Analyses were carried out with both MAIC and STC to assess the robustness of results.

The following sections describe the trials and their populations and provide additional details on the methods and how these were applied. This is followed by a summary of the results and a discussion of findings.

4.2.7.1 Data Sources and Assessment of Compatibility for Comparison

The compatibility of the data sources, and hence the feasibility and appropriateness of using MAIC and STC, was assessed through a review of the design and population profiles of the studies involved in the analyses. The next sections summarise key information from this review, and highlight discrepancies that present challenges or require specific assumptions in analyses.

General Design Features

The etanercept 1031 trial (136) was a multicentre, 12-week, double-blind, placebo-controlled randomised study of etanercept (50 mg subcutaneously weekly) on a background NSAID in the treatment of adult patients with nr-axSpA with a 92-week open-label extension. Patients were required to have active AxSpA despite therapy with at least two NSAIDs at the optimal dose, with non-radiographic sacroiliitis (by blinded and centralised assessment) defined as those patients who did not meet the modified New York (NY) criteria.

The ABILITY-1 Trial (112) was a multicentre, Phase III, 12-week, double-blind, placebo-controlled randomised study designed to evaluate the efficacy and safety of adalimumab 40 mg administered every other week in adult patients with active AxSpA who are not diagnosed with ankylosing spondylitis (AS), psoriasis, or psoriatic arthritis and who have had an inadequate response or intolerance to one or more NSAIDs or had a contraindication to NSAIDs. ABILITY-1 also had an open label extension of adalimumab treatment for up to an additional 144 weeks. Although ABILITY-1 aimed to exclude patients with AS, based on the arthritis Advisory committee meeting document from the FDA, it was estimated that an average of 37.2% of patients for the entire trial population were reclassified as being AS patients following central assessment of radiographs. (139).

The RAPID-axSpA trial (90) was a multicentre, 24-week, double-blind, placebo-controlled, randomised study designed to evaluate efficacy and safety of certolizumab pegol (400 mg every four weeks or 200 mg every two weeks, after a loading dose period of 400 mg on Weeks 0, 2, and 4) in patients with active AxSpA. The RAPID-axSpA trial had a dose-blind extension to Week 48, and an open-label extension to Week 204. This trial had both nr-axSpA and AS subgroups. As in the ABILITY-1 trial. However, it was noted in the FDA briefing document for certolizumab pegol that of the 147 patients originally diagnosed as nr-axSpA, 72 (49%) were later reclassified as AS patients upon central assessment (6 patients were not centrally assessed) (140).

The trials were similar in terms of their comparator arms and blinding but differed in an important way in the profile of their populations. The 1031 trial included patients with nr-axSpA, while ABILITY-1 and RAPID-axSpA included a mix of nr-axSpA and AS patients.

The publications for these trials either did not report results for the nr-axSpA and AS subgroups separately, or where this was done in the case of the RAPID-axSpA trial (90), there was considerable reclassification within subgroups which was only subsequently detected. Therefore, a comparison of results from 1031 directly with these subsets was not possible. For this reason we sought to balance the mix of nr-axSpA and AS patients involved in the comparison by combining data from the etanercept 314-EU trial (12), which targeted patients with AS, with the 1031 trial to form the basis of comparison (i.e. index trials). Etanercept 314-EU was a randomised, double-blind, placebo-controlled trial evaluating the safety and efficacy of etanercept 50 mg once weekly compared with 25 mg twice weekly in patients with AS.

The pooled 1031 and 314-EU trials were used as the basis for comparison (i.e. index trials) with the ABILITY-1 and RAPID-axSpA trials. This allows adjusting to the mix of nr-axSpA and AS patients in the comparator trials, but assumes that treatment differences in these mixed populations are generalisable to nr-axSpA patients. That is, while outcomes may differ for the two types of diagnoses, the treatment effects are the same (i.e. no interaction between treatment and diagnosis type). Comparisons with adalimumab were also derived based on the 1031 trial alone as both trials were designed to target nr-axSpA subjects only. This assumes that outcomes in ABILITY-1 were not affected by inclusion of patients with AS; that is, the observed outcomes are generalisable to patients with nr-axSpA.

To match the proportion of nr-axSpA patients reclassified as AS in each of the comparisons with adalimumab and certolizumab pegol, the proportion of AS patients included in the pooled 1031 and 314-EU trial analysis was set at 37.2% for the comparison vs adalimumab and 65.8% for the comparison vs certolizumab pegol at the AxSpA level (see Appendix C for details on the calculations used). The proportions are based on the reclassified nr-axSpA data reported in the Humira (139) and Cimzia (140) FDA Arthritis Advisory Committee meeting reports.

Population Profiles

The characteristics of the patients in the etanercept, adalimumab and certolizumab pegol trials are summarised in Table 87. This table lists the variables considered for inclusion in analyses to adjust for confounding. Some variables were not used in the particular targeted comparison (e.g. family history of SpA); these could not be included for adjustment. Table 87 also highlights key differences in characteristics, some of which were highlighted in the discussion of the NMA (section 4.2.6). For example, patients in the 1031 trial were on average younger than those in the other trials (by approximately 8 years), and disease severity scores (BASDAI, BADFI, ASDAS, etc.) tended to be lower. These differences can be taken into account in comparisons with MAIC and STC.

A more important difference lies in the duration of symptoms; the 1031 trial restricted the population to patients experiencing symptoms for less than 5 years. No such restriction was applied in the ABILITY-1 and RAPID-axSpA trials. The mean symptom duration for the 1031 trial is 2.5 years, compared with a mean symptom duration of 10.1 years in the ABILITY-1 trial and a median of 5.5 years in the RAPID-axSpA trial. This discrepancy makes adjustment for symptom duration very difficult; for example, in MAIC, it would be impossible to re-weight patients in 1031 to get a mean value of 10 years (since all are below 5 years). In STC, this would require generating predictions far beyond the range where the data lie. In addition, symptom duration was not available in the 314-EU trial, which instead captured duration of disease (i.e. time since diagnosis of AS).

To confound the comparisons, symptom duration must be predictive of changes in BASDAI and BASFI at 12 weeks. This has been found to be the case in prior studies (56, 107, 112). We assessed this in our context by examining plots of change in outcomes by symptom duration, and repeated this with disease duration in 314-EU. These are presented in Xxxxxxx7 and Xxxxxxx8, which show a wide scatter of points, and weak and statistically non-significant slopes. Based on these, it was deemed appropriate to omit symptom and disease duration from the MAIC and STC analyses. Any effect due to the weak associations was indirectly adjusted by accounting for study effects in comparative analyses.

Table 87: Summary Baseline Characteristics

	ETN				ADA		CZP	
	nr-axSpA (1031)		axSpA (1031 + 314)		nr-axSpA		axSpA	
	ETN	PBO	ETN	PBO	ADA	PBO	CZP	PBO
N	**106**	**109**	**411**	**160**	**91**	**94**	**218**	**107**
Demographics								
Male, %	64.2	56.9	[redacted]	[redacted]	48	43	62	61
White, %	74.53	72.48	[redacted]	[redacted]	100	97	91	89
Age, yrs	31.9	32	[redacted]	[redacted]	37.6	38.4	39.90	39.44
Disease characteristics								
Symptom duration, yrs	2.42	2.45	N/A	N/A	10.1	10.1	9.99	11.30
Symptom duration, yrs, median	2.3	2.4	N/A	N/A	N/A	N/A	7.39	7.70
Concomitant DMARD use, %	19.8	19.3	[redacted]	[redacted]	19	17	28	36
Concomitant NSAID use, %	100	100	[redacted]	[redacted]	79	79	88	86
History of SpA features								
HLA-B27 positive, %	67.0	76.1	[redacted]	[redacted]	82	74	77	81
Arthritis (past or present), %	47.2	52.3	N/A	N/A	35	52	53	57
Good prior response to NSAIDs, %	[redacted]	[redacted]	N/A	N/A	70	74	N/A	N/A
Heel enthesitis (past or present), %	43.4	40.4	N/A	N/A	40	40	N/A	N/A
Family history of SpA, %	26.4	21.1	N/A	N/A	31	25	N/A	N/A
Anterior uveitis (past or present), %	7.5	8.3	N/A	N/A	13	11	N/A	N/A
Dactylitis (past or present), %	4.7	10.1	N/A	N/A	11	11	N/A	N/A
Inflammatory bowel disease (past or present), %	0.9	0.9	N/A	N/A	4	6	N/A	N/A
BASDAI (0–10)	5.96	5.97	[redacted]	[redacted]	6.4	6.5	6.45	6.40
BASFI (0–10)	4.18	3.87	[redacted]	[redacted]	4.5	4.9	5.35	5.50
ASDAS	2.98	2.97	[redacted]	[redacted]	3.2	3.4	3.85	4.00
Patient global assessment of dis. activity (0–10)	5.77	5.78	[redacted]	[redacted]	6.8	6.8	N/A	N/A
Total back pain (0–10)	5.51	5.48	[redacted]	[redacted]	6.9	7.0	N/A	N/A
Inflammation/morning stiffness (0–10)	[redacted]	[redacted]	[redacted]	[redacted]	6.5	6.7	N/A	N/A
CRP, mg/l	6.76	6.43	[redacted]	[redacted]	6.8	7.6	17.27	22.35
CRP, mg/l, median	[redacted]	[redacted]	N/A	N/A	N/A	N/A	12.50	15.00
MASES (0–13)	3.13	2.94	N/A	N/A	3.5	3.9	N/A	N/A
SF-36 PCS (0–100)	[redacted]	[redacted]	N/A	N/A	33.9	33.1	N/A	N/A
SPARCC MRI sacroiliac joint score (0–72)	8.04	7.73	N/A	N/A	5.1	4.7	N/A	N/A
SPARCC MRI spinal score (0–108)	4.66	3.47	N/A	N/A	4.1	4.6	N/A	N/A

Abbreviations: ASDAS, Ankylosing Spondylitis Disease Activity Score; AxSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C-reactive protein; CZP, certolizumab pegol; DMARD, disease modifying anti-rheumatic drug; ETN, etanercept; HLA-B27, human leucocyte antigen B27; MASES, Maastricht Ankylosing Spondylitis Entheses Score; MRI, magnetic resonance imaging; NA, not applicable; nr-axSpA, non-radiographic axial spondyloarthritis; NSAID, non-steroidal anti-inflammatory drug; SF-36 short form 36; SpA, spondyloarthritis; SPARCC, Spondyloarthritis Research Consortium of Canada.

Green = data available for all studies; Grey = data not available for all studies; AxSpA = pooled nr-AxSpA and AS population

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Summary of Compatibility of Studies

The etanercept 1031/314-EU, ABILITY-1 and RAPID-axSpA trials were generally similar in design and captured and reported BASDAI and BASFI changes at Week 12. The populations of these trials were similar, with sufficient overlap in baseline characteristics to allow adjustment for potential confounding in MAIC and STC analyses. Exceptions to this are variables that were not captured in all sources and symptom and disease duration. These were omitted from analyses since they did not appear to be significantly associated with the outcomes. We assessed any residual impact from these differences, and adjusted for potential differences in variables that are missing in one or both studies in the comparison by measuring the difference in placebo outcomes in the trials after adjustment for baseline characteristics. We refer to this as a study effect as it also captures the impact of differences in studies beyond population profiles. The next section describes how these

are derived and applied. Analyses adjusting for this study effect produce comparisons between treatments when the placebo arms are made to be identical.

Another challenge was the mixture of different diagnoses nr-axSpA, AxSpA and AS in the ABILITY-1 and RAPID-axSpA trials. We incorporated the 314-EU trial in some analyses to be able to match the proportion of AS patients in the comparator trials. This introduced new assumptions, in particular, that the type of diagnosis may impact the magnitude of change in outcomes, but does not modify the effect of treatment.

4.2.7.2 Methods

Analysis Strategy

The following sections describe the analytic process involved in MAIC and STC analyses. Both methods were applied to compare changes in BASDAI and BASFI over 12 weeks of treatment with etanercept vs adalimumab and etanercept vs certolizumab pegol. In addition, the same methods were applied to the placebo arms of the trials being compared to quantify and adjust for any study effects that may distort comparisons beyond differences in the study populations. With both methods, results are reported in terms of the difference in change from baseline in BASDAI and BASFI along with 95% confidence intervals, adjusted for population differences and then further adjusted for study effects.

The combination of analyses, populations used, and variables included are summarised in Table 88. To minimise confounding, analyses aimed to include all available variables characterising the patients. This requires the variables being available in both studies involved in the comparison, which was not always the case. Symptom duration, history of SpA features variables, MASES, SF-36, and SPARCC MRI score variables were not available for the etanercept 314-EU trial, therefore they all had to be excluded from all analyses that included this trial. Patient global assessments of disease activity, total back pain, inflammation/morning stiffness were not available for the RAPID-axSpA trial, and, were therefore, excluded from the comparisons with CZP.

Another constraint to including variables for adjustment was the extent of overlap in their distributions in the studies being compared. Thus, while symptom duration was available for both the 1031 index and ABILITY-1 comparator trials, it was not feasible to robustly match the populations on symptom duration using MAIC since symptom duration could not exceed 5 years in 1031, while the mean duration in ABILITY-1 was around 10 years. Neither was it feasible to match on the ratio of concomitant NSAID use (since 1031 had 100% of patients on NSAIDs) or the proportion of patients with AS (since 1031 excluded these patients). Thus, symptom duration, NSAID use, and AS diagnosis were excluded from the analyses including 1031 as index trial only. We also omitted from analyses, based on the 1031 trial, history of SpA features variables, MASES, SF-36, and SPARCC MRI scores because their inclusion in matching resulted in unbalanced MAIC weights and very small effective sample sizes.

To ensure consistency and comparability of results between MAIC and STC analyses, equations created for STCs included the same variables used in matching of populations in MAICs. The only exception to this was the inclusion of BASDAI and BASFI scores, which were strongly correlated. Thus, the equations only include baseline values corresponding to the outcome being modelled. This is not expected to cause distortion, as baseline BASFI was not statistically significant in the presence of baseline BASDAI in the equation for change in BASDAI, and likewise for the BASFI equation. Otherwise, the equations retained

all predictors considered regardless of statistical significance to ensure comparability with MAIC results.

Table 88 summarises the analyses that were carried out, as well as the variables that were included in adjustment, and those that were omitted. These analyses were carried out for changes in BASDAI and BASFI at Week 12.

Table 88: Summary of analyses

Treatment Comparison/ Study Effect	Populations	Analysis method	Comparison is Adjusted for	Included variables	Excluded variables
ETN vs. ADA **PBO_E vs. PBO_A**	1031 vs. ABILITY-1	MAIC	Baseline characteristics	1i	1e
			Baseline characteristics and Study effect		
		STC	Baseline characteristics		
			Baseline characteristics and Study effect		
	1031+314 vs. ABILITY-1	MAIC	Baseline characteristics	2i	2e
			Baseline characteristics and Study effect		
		STC	Baseline characteristics		
			Baseline characteristics and Study effect		
ETN vs. CZP **PBO_E vs. PBO_C**	1031+314 vs. RAPID (AxSpA)	MAIC	Baseline characteristics	3i	3e
			Baseline characteristics and Study effect		
		STC	Baseline characteristics		
			Baseline characteristics and Study effect		

1i: Age, White, Male, Concomitant DMARDS use, baseline BASDAI, baseline BASFI, CRP, HLA-B27 positive, ASDAS, Patient global assessment of disease activity, Total back pain, Inflammation/Morning stiffness. Note that STC equations for baseline BASDAI and BASFI only included the corresponding baseline value.

1e: Symptom duration, Concomitant NSAID use, Arthritis, Good prior response to NSAIDs, Heel enthesitis, Family history of SpA, Anterior uveitis, Dactylitis, Inflammatory bowel disease, MASES, SF-36 PCS, SPARCC MRI sacroiliac joint score, SPARCC MRI spinal score, AS vs. nr-AxSpA diagnosis

2i: Age, White, Male, Concomitant DMARDS use, Concomitant NSAIDS use, baseline BASDAI, baseline BASFI, CRP, HLA-B27 positive, ASDAS, Patient global assessment of disease activity, Total back pain, Inflammation/Morning stiffness, AS vs. nr-AxSpA diagnosis. Note that STC equations for baseline BASDAI and BASFI only included the corresponding baseline value.

2e: Symptom duration, Arthritis, Good prior response to NSAIDs, Heel enthesitis, Family history of SpA, Anterior uveitis, Dactylitis, Inflammatory bowel disease, MASES, SF-36 PCS, SPARCC MRI sacroiliac joint score, SPARCC MRI spinal score

3i: Age, White, Male, Concomitant DMARDS use, Concomitant NSAIDS use, baseline BASDAI, baseline BASFI, CRP, Median of CRP, HLA-B27 positive, ASDAS, AS vs. nr-AxSpA diagnosis. Note that STC equations for baseline BASDAI and BASFI only included the corresponding baseline value.

3e: Symptom duration, Arthritis, Good prior response to NSAIDs, Heel enthesitis, Family history of SpA, Anterior uveitis, Dactylitis, Inflammatory bowel disease, Patient global assessment of disease activity, Total back pain, Inflammation/Morning stiffness, MASES, SF-36 PCS, SPARCC MRI sacroiliac joint score, SPARCC MRI spinal score

General Principles of MAIC and STC Analyses

Adjusting for Population Differences

The aim of MAIC and STC analyses are to calculate an adjusted estimate of the outcome expected with the index treatment in a population that matches the profile of the comparator treatment's characteristics. Suppose, we denote by $Y_E(X_E)$ the outcome (change from baseline in BASDAI or BASFI) observed with etanercept in the 1031+/-314-EU trials where the population profile is given by the vector of baseline characteristicsX_E. We denote by $\hat{Y}_E(X_*))$ the predicted outcome with etanercept in now a different population described by X_*. Thus, for a comparison with adalimumab, we aim to calculate $\hat{Y}_E(X_A)$, where X_A corresponds to the profile reported in the ABILITY-1 trial. The indirect comparison can then be calculated as:

$$\Delta_{EA} = \hat{Y}_E(\bar{X}_A) - Y_A(\bar{X}_A),$$

Where $\bar{X}_A$is the vector of mean values of baseline characteristics in ABILITY-1.

MAIC and STC differ in the way they derive $\hat{Y}_E(X_*)$. In MAIC, this is obtained by applying weights (w_i) to individual observations in the etanercept trials:

$$\hat{Y}_E(X_A) = \sum w_i \times Y_{Ei}(X_{Ei}),$$

where the weights are calculated such that $\bar{X}_A = \frac{\sum w_i \times X_{Ei}}{\sum w_i}$, *i* indexes outcomes (Y) and characteristics (X) of individuals in the etanercept trials.

In STC, the adjusted outcomes are predicted from regression equations for each outcome of interest. In this case, equations are created for changes in BASDAI and BASFI from patient-level data from the etanercept trials. The adjusted outcomes with etanercept are then predicted by setting predictors to the mean values observed in the comparator trials. For example, for the comparison with adalimumab, the predicted value is given by:

$$\hat{Y}_E(X_A) = \hat{\beta}'\bar{X}_A,$$

where $\hat{\beta}$ represents the vector of coefficients in the predictive equation for BASDAI or BASFI. The indirect comparison can then be calculated as above.

Study Effect Assessment

The comparisons described in the previous section reflect estimates of the relative effectiveness of treatments after adjusting for differences in patient characteristics. These estimates may be subject to residual confounding when some variables are not available in both studies, or when there are differences in study design or background factors. This is referred to as a study effect, and can be assessed from outcomes observed in the reference arms of the studies being compared when these were assigned the same treatment – placebo, in our analyses – after adjusting for population differences with MAIC and STC. For example, the study effect in the comparison of etanercept and adalimumab is obtained from:

$$\Delta_{P_E P_A} = \hat{Y}_{P_E}(X_{P_A}) - Y_{P_A}(X_{P_A}),$$

where P_E and P_A designate the placebo arms of the etanercept and adalimumab trials, respectively.

A large study effect (e.g. in the same order of magnitude as the adjusted effect estimate between treatments when the treatment effect is significantly large) may be an indication of lack of compatibility between studies, and would trigger a closer examination of the studies and assumptions. Otherwise, the effect estimates can be further adjusted to remove any residual bias. This is given by:

$$\Delta_{EA}^{*} = \Delta_{EA} - \Delta_{P_E P_A}.$$

Uncertainty of Effect Estimates

Standard errors and confidence intervals can be readily derived for the effect estimates. For instance, the variance of Δ_{EA} is given by $var\left(\hat{Y}_E(\bar{X}_A)\right) + var\left(Y_A(\bar{X}_A)\right)$. The second term in the sum is the variance of the observed change from baseline in outcomes for adalimumab. Since these were not reported in the publications of the ABILITY-1 or RAPID-axSpA trials, the variance of the observed change from baseline in the etanercept trials were used to calculate standard errors of the means for adalimumab and certolizumab pegol. The variance of the adjusted prediction, i.e. $var\left(\hat{Y}_E(\bar{X}_A)\right)$, is produced from the analyses. In MAIC, this is given from the calculation of the weighted mean of outcomes; in STC, it is the variance of the prediction, which can be obtained directly from statistical software. Similarly,

$$var(\Delta_{EA}^{*}) = var(\Delta_{EA}) + var(\Delta_{P_E P_A}).$$

Application of MAIC

Analyses followed the steps as described by Signorovitch et al, 2010 (189). The key step in these analyses is the derivation of weights to balance the populations. As noted above, the weights must be such that, when applied to the characteristics of the patients in the etanercept trials, the weighted mean values of characteristics match the mean values of the comparator trials.

Weights are derived from a propensity-score type equation that predicts whether a given type of patient originates from the primary treatment trial or the competing drug's trial as a function of baseline characteristics. A logistic regression model is fitted including all available predictors (as described in Table 88):

$$\log(\varphi_i) = log\left(\frac{p_i}{1-p_i}\right) = \gamma' X_i,$$

where φ_i represents the odds of being in the etanercept vs comparator trial, γ' is a vector of coefficients and X_i correspond to the predictors available for inclusion. Given that only aggregate data are available for the comparator trial, the β coefficients of the variables in the equation must be estimated by method of moments (instead of likelihood maximisation, as is the case usually). The estimated equation is then applied to patients in the etanercept trials to obtain predicted odds, which are used as the basis to calculate weights by normalising these as follows:

$$w_i = \frac{\hat{\varphi}_i}{\sum \hat{\varphi}_j} \times N_E,$$

where N_E is the sample size of the etanercept group in the index trials. This process must be repeated for each pair of arms being compared, including for the comparison of outcomes in placebo arms to assess study effects.

Prior to applying these to derive comparisons, the distribution of the weights is examined to ascertain the spread and the effective sample size contributing to analyses. If the index and comparator trial distribution ranges have major differences in variables included in the matching, it is possible that an important fraction of patients are given weights very close to 0 and a small fraction receive very large weights to obtain a match. The effective sample size is given by $(\sum w_i)^2/(\sum w_i^2)$ and can be thought of as the extent of information from the etanercept population contributing to the prediction of adjusted outcomes in comparative analyses.

The key assumptions involved in a MAIC are that:

- The studies are generally compatible for comparison: that is, differences in outcomes can be attributed to differences in the population and the treatments received. Differences in design or background factors have little or negligible impact on outcomes, or can be taken into account in the analyses.
- Matching the mean values of the characteristics is sufficient.
    - This implies, for instance, that the correlation between characteristics is the same in the index and comparator trials.
- All potential confounders are available and have been included in the matching.
    - In our analyses, some variables were not available in all sources. We dealt with this by assessing and adjusting for residual study effects.
- The effects derived from the indirect comparison apply to the entire population; that is, there are no subgroup effects or effect modification.
    - In our analyses this implies that, in comparisons involving a mix of nr-axSpA and AS patients, the effects calculated can be generalised to patients with nr-axSpA.

MAIC analyses were performed using R v3.1.0. Comparisons summarised in Table 88 were produced, including results adjusted for population characteristics and further adjusted for study effects, along with 95% confidence intervals.

Application of STC

Analyses followed the analytic process described by Caro & Ishak, 2010 (190). The key step in an STC is the derivation of the equations for change from baseline in BASDAI and BASFI at Week 12. Multiple linear regression models were created for these, including the same predictors that were included in MAIC analyses (as described in Table 88). For treatment comparisons, equations were derived from data from the etaercept arm of the 1031 and combined 1031+314-EU trials. To assess study effects, equations were also derived from the placebo arm of the 1031 and combined 1031+314-EU trials. All potential predictors were retained in these equations, regardless of statistical significance of the coefficient estimates. This was done to align with the MAIC process which forces all predictors in the matching; thus, following the same approach would ensure fair assessment of the consistency of results. The only exception was that the equation for BASDAI included baseline BASDAI scores but omitted baseline BASFI, as these were strongly correlated. Similarly, baseline BASDAI was left out of the equation for change in BASFI. This was done to reduce co-linearity between predictors. This was not a concern in the MAIC since the matching of populations is not outcome-specific, but rather population/treatment specific. The R2 statistic was checked for each equation as a general assessment of fit.

Equations derived from the etanercept arm were then applied to predict outcomes in populations matching the profiles of the ABILITY-1 and RAPID-axSpA trials by setting predictors to the mean values of corresponding variables. This represents changes in BASDAI or BASFI with etanercept in a population that matches the comparator populations (i.e., $\hat{Y}_E(X_A)$ and $\hat{Y}_E(X_C)$). Similarly, predictions were generated from the placebo equations of the etanercept trials to predict placebo outcomes in the comparator populations. Differences between the predicted values and those observed in the placebo arms of the comparators were considered to reflect residual confounding and study effects. These were incorporated to derive fully adjusted indirect comparisons, as described in the previous sections.

The key assumptions involved in STC are that:

- The studies are generally compatible for comparison: that is, differences in outcomes can be attributed to differences in the population and treatments received. Differences in design or background factors have little or negligible impact on outcomes, or can be taken into account in the analyses.
- All potential confounders are available and have been included in the predictive equations.
    - In our analyses, some variables were not available in all sources. We dealt with this by assessing and adjusting for residual study effects.
- The assumptions invoked by the regression models are met and relationships between the outcomes and predictors are captured accurately.
- The equations derived from the index trial are applicable to the population of the comparator populations. That is, the coefficients of predictors in the equations would be the same if derived from one of the comparator trials.
    - This is analogous to the assumption of no effect modification noted above for MAIC.

STC analysis was performed using SAS v9.2. Comparisons summarised in Table 88 were produced, including results adjusted for population characteristics and further adjusted for study effects, along with 95% confidence intervals.

4.2.7.3 Results

Profile of Study Populations

The profile of the populations of the etanercept 1031 and 314, ABILITY-1 and RAPID-axSpA trials is summarised in Table 87. Key areas of difference were mentioned and discussed in section 4.2.7.1).

Matching of Patients in the etanercept and Comparator Trials

The distribution of weights derived from matching analyses in the MAIC is summarised in Xxxxxx89, which also shows the original sample sizes of the groups in the etanercept trials as well as the effective sample sizes after application of the weights. For example, in the comparison of etanercept vs adalimumab based on the 1031 trial, more than 25% of patients received a weight of less than 0.05 and less than 25% received weights exceeding 1.0. This leads to a considerably lower effective sample size of 25 compared with the original population of 102. There was a similar loss of information when matching the placebo arm and groups from the pooled 1031+314-EU-based analyses. This is attributable to large differences in characteristics included in the matching, such as age and disease severity measures (BASDAI, BASFI, ASDAS, and CRP). Weights derived for comparisons with

certolizumab pegol had a more uniform distribution, but still led to a nearly 40% loss of effective sample size for etanercept and over 65% for placebo.

[redacted]

			Sample Size (N)		Distribution of Weights					
Comparison	Index	Arm	Actual	Effective	Min	Q1	Med	Mean	Q3	Max
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

* Weights of 0 in the table are rounded. These represent values below 0.05.

Equations used in STC Analyses

The coefficients of the equations derived in STC analyses for BASDAI and BASFI are provided in Table 90 and Table 91. As noted above, all predictors considered in each equation were retained, irrespective of statistical significance. The tables show that few predictors had a significant effect on outcomes, which led to relatively weak R^2 values in all cases (< 0.40). This may be due to a few possible factors: the outcomes may be susceptible to high variability across patients with the same characteristics; the outcomes have few known/measurable determinants; and, some known/measured determinants are missing from the equations. The first two factors are not necessarily problematic in this application, but the third would lead to residual confounding. It is assumed that adjusting for study effects will alleviate this issue, and minimise the impact on results.

Table 90: Final models for BASDAI, restricted to variables adjusted in MAIC

	Enbrel 1031+314				Enbrel 1031 only			
	Placebo		ETN		Placebo		ETN	
Parameter	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value
Intercept								
BASDAI (per 1 point increase)								
Age (per 1 year increase)								
White vs. Other								
Male vs. Female								
Concomitant DMARD use (yes vs. no)								
CRP (per 1 mg/l increase)								
HLA-B27 (positive vs. negative)								
ASDAS (per 1 point increase)								
Patient global assessment of dis. activity (per 1 point increase)								
Total back pain (per 1 point increase)								
Inflammation/morning stiffness (per 1 point increase)								
Ankylosing Spondylitis (yes vs. no)								

Abbreviations: ETN, etanercept; BASFI, Bath Ankylosing Spondylitis Functional Index; DMARD, disease modifying anti-rheumatic drug; CRP, C-reactive protein; HLA-B27, human leucocyte antigen B27; ASDAS, Ankylosing Spondylitis Disease Activity Score; AS, ankylosing spondylitis; SE, standard error.

Table 91: Final models for BASFI, restricted to variables adjusted in MAIC

	Enbrel 1031+314				Enbrel 1031 only			
	Placebo		ETN		Placebo		ETN	
Parameter	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value
Intercept								
BASFI (per 1 point increase)								
Age (per 1 year increase)								
White vs. Other								
Male vs. Female								
Concomitant DMARD use (yes vs. no)								
CRP (per 1 mg/l increase)								
HLA-B27 (positive vs. negative)								
ASDAS (per 1 point increase)								
Patient global assessment of dis. activity (per 1 point increase)								
Total back pain (per 1 point increase)								
Inflammation/morning stiffness (per 1 point increase)								
Ankylosing Spondylitis (yes vs. no)								

Abbreviations: ETN, etanercept; BASFI, Bath Ankylosing Spondylitis Functional Index; DMARD, disease modifying anti-rheumatic drug; CRP, C-reactive protein; HLA-B27, human leucocyte antigen B27; ASDAS, Ankylosing Spondylitis Disease Activity Score; AS, ankylosing spondylitis; SE, standard error.

Comparison of BASDAI Change from Baseline[a]

Etanercept vs adalimumab

Table 92 and Table 93 summarise results obtained from MAIC and STC comparing BASDAI changes at 12 weeks with etanercept and adalimumab. Table 87 shows the derivation of the study effect. The placebo arm in the ABILITY-1 trial showed a reduction of one point over 12 weeks. Adjusted estimates from MAIC and STC trial predicted a reduction of [redacted] and [redacted] points, respectively based on the placebo arm in the 1031 trial. This yields a study effect estimate of [redacted]) from the MAIC analysis and [redacted]) from STC, both of which are not statistically significant and of relatively small magnitude. Results based on pooled analyses of the 1031 and 314 trials yielded a larger study effect estimate from both methods [redacted]). The estimates were still not statistically significant, however.

Table 92: Comparison of BASDAI for placebo arms of etanercept and ABILITY-1 Trials

Index Trial	Method	Observed Change from Baseline		Predicted Change for PBO in ETN Trial	Study Effect
		PBO in ETN Trial	PBO in ABILITY-1		
1031	**MAIC**			[redacted]	[redacted]
	STC	-1.66	-1.00	[redacted]	[redacted]
1031 + 314	**MAIC**			[redacted]	[redacted]
	STC	[redacted]	-1.00	[redacted]	[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; PBO, Placebo; ETN, etanercept; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Table 93 reports the results of indirect comparisons adjusting for patient characteristics and study effects. ABILITY-1 showed a 1.9 point improvement in BASDAI over 12 weeks with adalimumab. The predicted response with etanercept after adjustment for patient characteristics based on the 1031 trial were [redacted] with MAIC and [redacted] with STC, which yields an adjusted incremental change of [redacted] with MAIC and [redacted] with STC. After adjusting for study effects, these drop to [redacted] and [redacted] with MAIC and STC, respectively.

[a] Please refer to Table 88 for list of variables included in the adjustment in each of these analyses. The same set of variables were used in MAIC and STC for a given comparison (except for BASFI omitted from the BASDAI STC equations, and vice versa).

Table 93: Comparison of BASDAI for etanercept vs adalimumab based on etanercept and ABILITY-1 trials

Index Trial	Method	Observed Change from Baseline		Predicted Change with ETN Adjusted For		Treatment Effect	
		ETN	ADA	Population Characteristics	Population Characteristics + Study Effect	Population Characteristics	Population Characteristics + Study Effect
1031	**MAIC**	-2.35	-1.90	[redacted]	[redacted]	[redacted]	[redacted]
	STC			[redacted]	[redacted]	[redacted]	[redacted]
1031 + 314	**MAIC**	[redacted]	-1.90	[redacted]	[redacted]	[redacted]	[redacted]
	STC			[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; ETN, etanercept; ADA, Adalimumab; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Results based on the pooled etanercept analyses yielded similar results, with fully adjusted incremental change estimates of [redacted] with MAIC and [redacted] with STC. While the 95% confidence interval of the MAIC result nearly excluded 0, the magnitude of the effect remained consistent with other analyses.

Etanercept vs certolizumab pegol

Table 94 and Table 95 summarise results obtained from MAIC and STC comparing BASDAI changes at 12 weeks with etanercept and certolizumab pegol based on the 1031+314-EU trials. The comparison of placebo arms revealed a residual study effect of [redacted] points with MAIC and [redacted] with STC. While these estimates are not statistically significant, the magnitude of the effect remains substantial, particularly in comparison with the size of the relative effect of etanercept vs certolizumab pegol (Table 95). Adjusting for differences in patient characteristics of the etanercept and certolizumab pegol populations produced estimates of incremental change in BASDAI of [redacted] points from the MAIC and [redacted] points from the STC, which were smaller than the differences in placebo response in the studies. Thus, accounting for study effects effectively nullified these small differences, and produced fully adjusted effect estimates of no difference with MAIC and a difference of [redacted] points with STC with confidence intervals for both covering the null value.

Table 94: Comparison of BASDAI for placebo arms of etanercept and RAPID-axSpA Trials

Index Trial	Method	Observed Change from Baseline		Predicted Change for PBO in ETN Trial	Study Effect
		PBO in ETN Trial	PBO in RAPID		
1031 + 314	**MAIC**	[redacted]	-1.20	[redacted]	[redacted]
	STC			[redacted]	[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; PBO, Placebo; ETN, etanercept; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Table 95: Comparison of BASDAI for etanercept vs certolizumab pegol based on etanercept 1031 and RAPID-axSpA trials

Index Trial	Method	Observed Change from Baseline		Predicted Change with ETN Adjusted For		Treatment Effect	
		ETN	CZP	Population Characteristics	Population Characteristics + Study Effect	Population Characteristics	Population Characteristics + Study Effect
1031 + 314							
	MAIC	[redacted]	-2.8	[redacted]	[redacted]	[redacted]	[redacted]
	STC			[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; ETN, etanercept; CZP, Certolizumab; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Comparison of BASFI Change from Baseline

Etanercept vs adalimumab

Table 96and Table 97 summarise results obtained from MAIC and STC comparing BASFI changes at 12 weeks with etanercept and adalimumab. Differences in outcomes observed in the placebo arms of the ABILITY and etanercept 1031 + 314-EU trials appears to be mostly accounted for by differences in patient characteristics. Adjusting for these yielded relatively small study effect estimates which did not reach statistical significance in either MAIC or STC analyses (Table 96).

Table 96: Comparison of BASFI for placebo arms of etanercept and ABILITY-1 Trials

Index Trial	Method	Observed Change from Baseline		Predicted Change for PBO in ETN Trial	Study Effect
		PBO in ETN Trial	PBO in ABILITY-1		
1031	**MAIC**	-0.89	-0.60	[redacted]	[redacted]
	STC			[redacted]	[redacted]
1031 + 314	**MAIC**	[redacted]	-0.60	[redacted]	[redacted]
	STC			[redacted]	[redacted]

Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index; PBO, Placebo; ETN, etanercept; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

The study effect in the MAIC based on the 1031 trial was of greater magnitude than the change in outcomes observed after adjusting for patient characteristics. That is, the observed change in BASFI with etanercept decreased [redacted] points when adjusted to the profile of the adalimumab population, whereas the study effect was estimated to be [redacted] points. Thus, the final estimate was affected more by the adjustment of study effect, rather than for patient characteristics. Both suggested a weak non-significant effect with an estimated incremental change of [redacted] after full adjustment. The STC analyses yielded a stronger difference [redacted] but still failed to reach statistical significance (Table 97).

Comparisons based on the 1031+314-EU trials produced similar results and conclusions. Study effects were slightly stronger in these analyses (Table 96), but the final comparison measures were of similar magnitude [redacted]) from MAIC [redacted] from STC) and were not statistically significant.

Table 97: Comparison of BASFI for etanercept vs adalimumab based on etanercept and ABILITY-1 trials

Index Trial	Method	Observed Change from Baseline		Predicted Change with ETN Adjusted For		Treatment Effect	
		ETN	ADA	Population Characteristics	Population Characteristics + Study Effect	Population Characteristics	Population Characteristics + Study Effect
1031	**MAIC**	-1.59	-1.10	[redacted]	[redacted]	[redacted]	[redacted]
	STC			[redacted]	[redacted]	[redacted]	[redacted]
1031 + 314	**MAIC**	[redacted]	-1.10	[redacted]	[redacted]	[redacted]	[redacted]
	STC			[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index; ETN, etanercept; ADA, Adalimumab; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Etanercept vs certolizumab pegol

As in the BASDAI analyses, comparisons of BASFI with certolizumab pegol revealed larger study effects than those observed in analyses of etanercept vs adalimumab. Both MAIC and STC suggested study effects near 0.5 points, but in both cases, the estimates were not statistically significant (Table 98). Nevertheless, the magnitude of these study effects exceeded the treatment effects observed after adjustment of population characteristics (Table 98). For example, in the MAIC of etanercept vs certolizumab pegol, after adjusting for population profiles, etanercept was associated with 0.30 points incremental reduction in BASFI at 12 weeks. The study effect derived from this comparison suggested the placebo arm in the etanercept trials had an incremental reduction of 0.49 points after adjusting for populations differences. Thus, when this is taken into account, the relative effectiveness of etanercept vs certolizumab pegol is estimated to be 0.19 (-0.56, 0.94). Similarly, the population-adjusted treatment effect from STC was -0.27 (-1.18, 0.65), which is reversed after adjustment for study effects (0.21 [-1.07, 1.49]). Thus, in these comparisons, differences in outcomes between etanercept and certolizumab pegol are only partly accounted for populations differences and a larger share of this stems from predictors missing from the adjustment or other study-level differences.

Table 98: Comparison of BASFI for placebo arms of etanercept and RAPID-axSpA Trials

Index Trial	Method	Observed Change from Baseline		Predicted Change for PBO in ETN Trial	Study Effect
		PBO in ETN Trial	PBO in RAPID		
1031 + 314	**MAIC**	[redacted]	-0.50	[redacted]	[redacted]
	STC			[redacted]	[redacted]

Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index; PBO, Placebo; ETN, etanercept; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Table 99: Comparison of BASFI for etanercept vs certolizumab pegol based on etanercept and RAPID-axSpA trials.

Index Trial	Method	Observed Change from Baseline		Predicted Change with ETN Adjusted For		Treatment Effect	
		ETN	CZP	Population Characteristics	Population Characteristics + Study Effect	Population Characteristics	Population Characteristics + Study Effect
1031 + 314	**MAIC**	[redacted]	-2.0	[redacted]	[redacted]	[redacted]	[redacted]
	STC			[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index; ETN, etanercept; CZP, certolizumab; MAIC, Matching-adjusted indirect comparison; STC, simulated treatment comparison.

Discussion of results

Indirect comparisons were carried out with MAIC and STC methods, which adjust for differences between the populations being compared as well as any residual confounding or study effects. Findings were generally consistent with both methods, and did not reveal any statistically significant differences between etanercept and comparators in analyses of changes in BASDAI and BASFI from baseline at 12 weeks. Comparisons of etanercept and adalimumab were carried out based on the 1031 (nr-axSpA) and 1031+314-EU (axSpA) populations and yielded similar findings. Though not statistically significant, the magnitudes of the adjusted treatment effect estimates were larger in comparisons of etanercept vs adalimumab for BASDAI and BASFI; comparisons with certolizumab pegol were notably smaller.

MAIC and STC are designed to overcome challenges related to heterogeneity arising in NMA by focusing comparisons between compatible trials and adjusting for population and study differences. The latter is a particularly important feature in this application, since some variables were not available in all studies included in the comparison. Furthermore, studies differed in some aspects, like the duration of symptoms and the mix of diagnoses of patients in the study. More specifically, ABILITY-1 and RAPID-axSpA included patients with AS; to compensate for that and allow adjustment for diagnosis type in analyses, patients from the etanercept 314-EU trial were added to the index population for some analyses. Despite this, comparisons with the RAPID-axSpA trial produced study effect estimates that exceeded the adjusted treatment effect estimates. Nevertheless, the magnitudes of the study effects [redacted] suggesting that the studies were reasonably compatible for comparison.

Both MAIC and STC rely on important assumptions, key among these is that studies being compared are compatible. This was deemed to be the case in these analyses, although some key issues were noted (e.g. symptom duration and SPARCC MRI scores). In addition, outcomes and potential confounders should be measured and available in all studies. This was not the case in our analyses. To deal with these, we incorporated adjustments for study effects in each comparison. This effectively adjusts for any differences between the placebo arms beyond population differences. STC analyses also rely on the accuracy of the predictive equations. Our analyses showed that much of the variability in BASDAI and BASFI changes were unexplained by the predictors included in the equations. It is unclear whether this is due to missing predictors from the equations or whether these outcomes are subject to large random variation. Adjustment for study effects helps to deal with the first issue. We also note that matching the etanercept trial populations to comparators led to a substantial loss of effective sample size, which increases the variance of treatment effect estimates. The low R^2 of the STC equation has a similar impact. Finally, both analyses assumed that the treatment effect derived from the indirect comparisons is not modified by patient characteristics; that is, the relative effectiveness of the treatments is the same in all subgroups of patients. In particular, analyses based on the pooled 1031+314-EU trials assumed the type of diagnosis did not modify the effects and that the patients from the 314-EU trial are similar to the nr-axSpA patients reclassified as AS in the ABILITY-1 trial and the AS patients (including the nr-axSpA patients reclassified as AS patients) in the RAPID-axSpA trial. Since patient-level data is only available from the index trial, it is impossible to examine this assumption in detail. MAIC and STC of etanercept vs adalimumab did show consistent results with both the 1031 and 1031+314-EU data used as the index. This is not a

formal test of effect modification since both analyses included nr-axSpA, but it does suggest that adjusting for population and study effects accounted for the differences in diagnosis mix.

In summary, these analyses suggest etanercept, adalimumab and certolizumab pegol have a similar impact on BASDAI and BASFI after 12 weeks of treatment after adjustment for population differences and study effects. Adjusting for study effects was a key feature of the analyses to account for missing variables, differences in diagnoses of patients, and other, possibly unknown, differences between the studies. The conclusion from this analysis is consistent with the other overall conclusions of other HTA bodies that have examined interventions in the nr-axSpA population. For example, in their appraisal of certolizumab pegol vs adalimumab, the SMC concluded that certolizumab pegol was considered broadly comparable to other comparators in the AS and nr-axSpA subpopulations.

5. Pharmaco-economic Evaluation

5.1 Summary of cost-effectiveness

An economic evaluation was conducted to determine the cost-effectiveness of etanercept versus conventional care and other anti-TNF agents used in clinical practice in England and Wales for the treatment of ankylosing spondylitis (AS) and non-radiographic axial spondyloarthritis (nr-axSpA).

Methodology

- A patient-level simulation model (specifically, a discrete event simulation [DES]) was developed in Microsoft® Excel with Visual Basic for Applications (VBA).
- The model primarily considered two populations, reflecting those detailed in the scope for this multiple technology appraisal (MTA) and an additonal analysis population relevant to the decision problem:
  - **AS:** people with severe active ankylosing spondylitis whose disease has responded inadequately to, or who are intolerant to NSAIDs. The economic model provides comparisons between ETN, CZP, ADA, GOL, INF and CON for this population.
  - **Nr-axSpA:** people with severe axial spondyloarthritis without radiographic evidence of ankylosing spondylitis but with objective signs of inflammation, whose disease has responded inadequately to, or who are intolerant to non-steroidal anti-inflammatory drugs (NSAIDs). The economic model provides comparisons between ETN, CZP and ADA and CON for this population.
  - **AxSpA:** In addition, the full axial spondyloarthritis population was considered in a separate comparison of ETN against CZP as it was not possible to consider all comparators in each population due to differences in trial populations and availability of clinical data.
- The analysis was conducted from the perspective of the UK National Health Service (NHS) and Personal Social Services (PSS). The high burden of ankylosing spondylitis on productivity, unemployment and absenteeism has been demonstrated previously (191); therefore the omission of such costs and health effects may underestimate the economic benefit of anti-TNF agents.
- Data to populate the model were derived from key clinical trials for etanercept and results of a clinical systematic review, network meta-analysis (NMA), and in a separate analysis, a match adjusted indirect comparison (MAIC). The model structure was developed in accordance with current OMERACT (Outcome Measures in Rheumatology) guidance (192). The model was therefore constructed around the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and Bath Ankylosing Spondylitis Functional Index (BASFI). This is also in line with previous economic evaluations in AS (193-203).
- Many aspects of the economic evaluation were informed by de-novo analyses of the etanercept Phase III clinical data. These included:
  - Week 12 response
  - Week 48 response
  - Treatment discontinuation
  - Health-related quality of life (HRQoL) estimates
- Parameter uncertainty was tested through comprehensive sensitivity analyses which

included probabilistic sensitivity analysis (PSA) and one-way deterministic sensitivity analysis.

Results

- In the base-case, the model reported that the incremental cost-effectiveness ratio (ICER) for ETN vs conventional care was £20,938 and £23,195 in the AS and nr-axSpA populations, respectively.
- In both AS and nr-axSpA populations, the ICERs of other anti-TNFs versus conventional care were similar to ETN, with the exception of the INF strategy in AS which had an ICER of £37,741 per QALY gained. The probabilistic and deterministic sensitivity analysis demonstrated that the results were robust to parameter uncertainty with ICERs consistently less than £30,000 per QALY gained.
- In the AS population, when ETN was compared incrementally with other anti-TNFs there were no differences observed between ICERs, demonstrating that, in this population, anti-TNFs have similar cost effectiveness estimates. However, in the nr-axSpA population differences in the ICERs between ETN, CZP and ADA were observed based on the traditional NMA which did not account for the heterogeneity between the clinical trial populations. When the adjustment is made for this heterogeneity using the MAIC technique, the differences in cost and QALYs are reduced, thus supporting the clinical conclusions of no difference between anti-TNFs and that they are all cost effective options for the treatment of nr-axSpA. This is further supported by the results of the probabilistic sensitivity analysis which showed that when variation in assumptions and clinical data are considered, the ranges of ICERs are consistently overlapping for ETN, CZP and ADA.
- Overall from the results of the economic analysis, we can conclude that the continued investment in ETN for the treatment of AS and future use of ETN in nr-axSpA represents good value for money for the NHS.

5.2 Design of the economic evaluation

An economic evaluation was conducted to determine the cost-effectiveness of etanercept versus relevant comparators used in clinical practice in England and Wales for the treatment of ankylosing spondylitis (AS) and non-radiographic axial spondyloarthritis (nr-axSpA).

5.2.1 Patient group considered

The following patient groups were considered in the model as outlined in the NICE scope for this multiple technology appraisal (MTA).

5.2.1.1 nr-axSpA population

The nr-axSpA population was defined based on the scope issued by NICE and was defined as people with severe axial spondyloarthritis without radiographic evidence of ankylosing spondylitis but with objective signs of inflammation, whose disease has responded inadequately to, or who are intolerant to, NSAIDs.

Analyses adjusting for differences in study baseline characteristics

The clinical systematic review identified that the baseline characteristics of nr-axSpA patients within the randomised controlled trials of certolizumab pegol and adalimumab were heterogeneous, and potentially differed in characteristics that could act as treatment effect modifiers. Furthermore, the populations of these trials also included sizable proportions of

AS patients who were originally classified as nr-axSpA on the basis of a difference between centralised and localised readings of x-rays (section 4.2.6.2).

To address the differences in the proportions of AS patents in the trials due to reclassification upon central assessment, analyses were conducted using match adjusted indirect comparison (MAIC) and simulated treatment comparison (STC) techniques (detailed in 4.2.7.1) that incorporated AS patients from the etanercept 314-EU trial. These analyses are referred to collectively as “analyses adjusting for differences in study baseline characteristics”. A comparison of the results from the MAIC and STC approaches show that while the results of the two analyses are similar, when considering comparisons between etanercept and both adalimumab and certolizumab, the MAIC analysis provides a lower overall comparative estimate of the benefit of etanercept, and is therefore considered overall to be the more conservative of the two approaches. To maintain consistency in the analysis utilised in the economic section, the MAIC was used throughout as the adjusted comparative efficacy measure between etanercept versus adalimumab and etanercept versus certolizumab (section 4.2.7.1). For the analysis comparing etanercept against certolizumab pegol, it was possible to address the issue of patient reclassification and differences in baseline characteristics by utilising the RAPID-axSpA trial results that were also available at the level of AxSpA patients, an approach not possible in the comparison of etanercept versus adalimumab. We note that although not explicitly detailed within the scope, the AxSpA population encompasses both nr-axSpA and AS patients, thus making it a relevant comparison to the decision problem outlined in the scope.

5.2.1.2 AS population

The AS population was defined based on current NICE guidance in TA143 and TA233 (4, 5):

- The patient’s disease satisfies the modified New York criteria for diagnosis of AS
- There is confirmation of sustained active spinal disease, demonstrated by:
  - a score of at least 4 units on the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), and
  - at least 4 cm on the 0 to 10 cm spinal pain visual analogue scale (VAS)
- These should both be demonstrated on two occasions at least 12 weeks apart without any change of treatment
- Conventional treatment with two or more non-steroidal anti-inflammatory drugs taken sequentially at maximum tolerated or recommended dose for 4 weeks has failed to control symptoms.

5.2.2 Interventions considered

Table 100 presents the interventions considered within the economic evaluation (as determined by the NICE scope and reflecting licensed populations for each intervention). Comparators include anti-TNF agents and conventional care, which consists of NSAIDs and physiotherapy as defined within the NICE scope.

Table 100: Comparators considered within the primary analysis

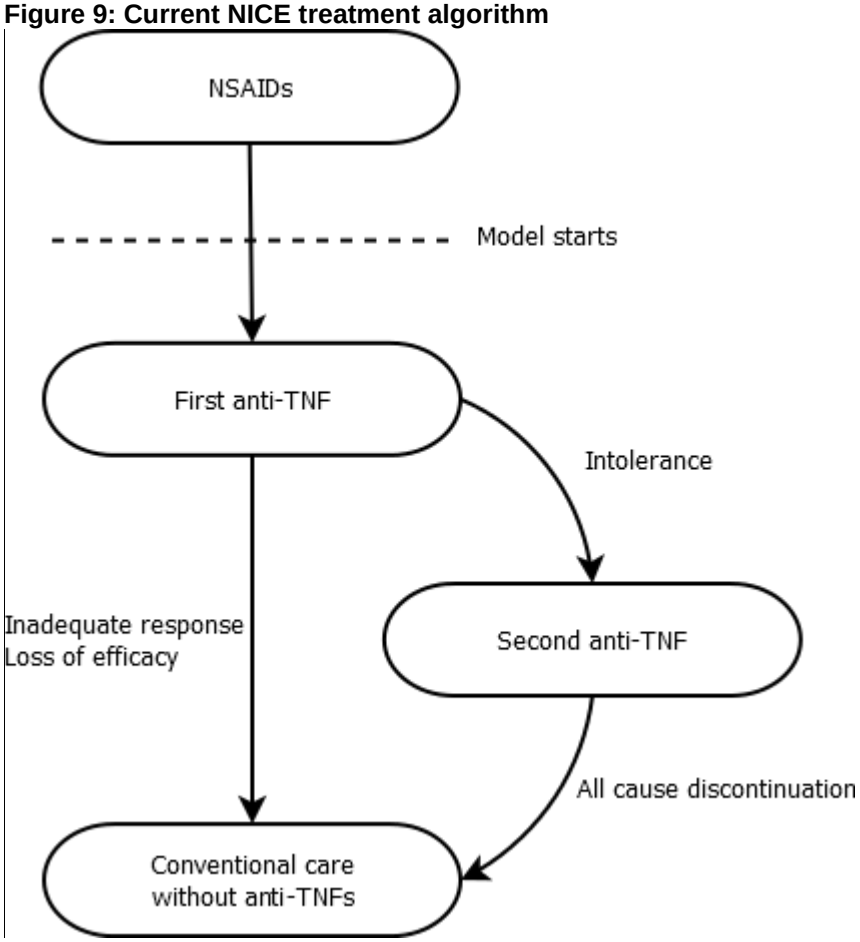
Therapy	AS	nr-axSpA
CON†	✓	✓
ETN	✓	✓

Therapy	AS	nr-axSpA
ADA	✓	✓
CZP	✓	✓
GOL	✓	
INF	✓	

Abbreviations: ADA, adalimumab; AS, ankylosing spondylitis; CON, conventional care; CZP, certolizumab pegol ; ETN, etanercept; GOL, golimumab; INF, infliximab; nr-axSpA, non-radiographic axial spondyloarthritis.
† Established clinical management without anti-TNF agents includes treatment with non-steroidal anti-inflammatory drugs (NSAIDs) and physiotherapy.

For each anti-TNF agent, a treatment strategy was defined based on current NICE guidance, TA143 and TA233 (4, 5), which permits use of a second anti-TNF agent only in patients who have been shown to be intolerant of a first anti-TNF agent before the end of the 12-week initial assessment period (Figure 9). The use of a second anti-TNF agent following all-cause discontinuation was also considered in a sensitivity analysis. Such sequential use has been documented in real-world observational studies previously (204), however the available data is limited and necessitated the adjustment of first-line efficacy data using observational data sources. Details of this scenario are provided in Appendix K. For conventional care, it was assumed that there were no alternative treatments, i.e. a conventional care patient was assumed to remain on conventional care until death.

Figure 9: Current NICE treatment algorithm



Abbreviations: NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor.

Treatment strategies compared alternative first line anti-TNF agents based on the treatment pathway presented in Figure 9 for the modelled populations (nr-axSpA and AS). See Table 100 for comparators by patient population. These treatment strategies represent a modification of the current NICE AS treatment algorithm (4, 5), which recommends etanercept, adalimumab and golimumab, generalised to permit comparison between all therapies detailed in the NICE scope. In the absence of published treatment pathways, the same approach has been assumed for the nr-axSpA population as for the AS population, with the exclusion of golimumab and infliximab, which are not currently licenced for use in nr-axSpA.

Comparisons of all treatment strategies are made against the continued use of conventional care and incremental analysis comparing all treatment pathways is also performed. In the analyses adjusting for differences in study baseline characteristics (section 5.2.1), etanercept is compared against adalimumab and certolizumab pegol strategies separately, as the MAIC permits only pairwise comparisons.

Note that as a modelling simplification, a specific second line anti-TNF agent is not defined. Instead, at this point an anti-TNF agent is allocated randomly to the individual, with the caveat that the individual cannot be allocated to the same therapy twice. We treat this approach as a proxy for estimating the cost-effectiveness of permitting sequential use of anti-TNF agents in general, rather than explicitly comparing alternative strategies of therapy beyond the first anti-TNF agent. This was considered appropriate given the paucity of data in this population; observational data describing the use of anti-TNF agents in second and subsequent lines of therapy have been published previously (196, 204). However, no randomised controlled trial (RCT) evidence was identified for this population during the systematic review (SR) (section 2). As a result, estimates of efficacy in a first line population are adjusted to reflect potentially lower efficacy at this line of therapy and this represents a limitation of the analysis.

5.2.3 Type of economic analysis performed and main outcome measure

A cost utility analysis was conducted, in line with current NICE guidance (205). The primary outcome was the incremental cost-effectiveness ratio (ICER), expressed as a cost per quality adjusted life year (QALY).

5.2.4 Viewpoint of the analysis

The analysis was conducted from the perspective of the UK National Health Service (NHS) and Personal Social Services (PSS). The analysis excluded patients' out-of-pocket expenses, carers' costs and lost productivity.

5.2.5 Time horizon and discounting

The model considers a lifetime time-horizon to reflect the natural history of nr-axSpA and AS (i.e. a chronic condition). A systematic review of economic evaluations of biologic treatments for nr-axSpA and AS conducted for this submission (section 5.3.2.3) found that the majority of papers identified used a time-horizon of over 20 years (4, 193-195, 198-200, 202, 203). Both Kobelt et al, 2007 (199) and the Schering-Plough submission for TA143 (4) implemented a lifetime and 70 year time horizon, respectively. Therefore, the long-term time horizon approach used in the current analysis is in line with previous economic evaluations.

Costs and benefits occurring in the future were discounted at a rate of 3.5% per annum in line with current NICE and Treasury guidance (205, 206). The reference year for the discounting of costs was 2014.

5.2.6 Outline of the Excel model

The economic model (developed in Microsoft Excel® with Visual Basic for Applications) uses a discrete event simulation (DES), a type of patient level simulation (PLS) approach to economic modelling. Patient-level approaches have been undertaken extensively in AS (193, 194, 203, 207, 208) and other rheumatic diseases (209, 210). Recent guidance from the NICE Decision Support Unit recommends the use of patient-level approaches in the following relevant situations (211):

- Model non-linearity with respect to heterogeneous patient characteristics.
    - Considerable between-patient heterogeneity exists within the considered populations (212, 213). SpA is associated with a heterogeneous clinical presentation; no single shared distinguishing feature exists for the conditions comprising SpA (213).
- Patient flow is determined by time since last event or history of previous events (the requirement for memory in the model).
    - The time-dependency requirement becomes more burdensome when modelling the second anti-TNF agent in a sequence (if an individual has failed first line therapy due to intolerance).
    - In TA233, the manufacturer of golimumab was unable to adequately incorporate (as assessed by the Evidence Review Group) sequential use of anti-TNF agents within the submitted Markov framework due to these complexities.

Two disease-specific health measures are followed over time and relate to patients' quality of life and medical resource use: the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and the Bath Ankylosing Spondylitis Functional Index (BASFI). As noted by Bansback et al, 2007 (192), in preliminary proposals for defining a reference case for economic evaluations in AS, the BASDAI and BASFI are most frequently used outcome measures in economic evaluations. They are also widely used in clinical practice within England and Wales, and this was confirmed at the etanercept advisory board (section 5.3.3). Choice of clinical outcomes is further discussed in section 5.4.1.1. The model was based principally on BASDAI and BASFI, including important aspects such as predicting response to treatment, disease progression, resource use costs and utilities.

BASDAI and BASFI are continuous measures and it was considered advantageous to be able to avoid dichotomising these into crude health states such as response and non-response. During TA143, the NICE DSU reported that internal inconsistency was created by the crude classification into binary response and non-response (1). While more nuanced approaches are clearly possible within a cohort-based framework, this may lead to increasing complexity.

The use of these outcome measures precludes the explicit inclusion of progression from nr-axSpA to AS. This was however unavoidable on the basis of a lack of available data on:

- the type of patients that progress;
- the rate of progression;

- the effectiveness of therapies in patients refractory to first line treatment who have progressed;
- the clinical impact of radiographic progression (i.e. the presence or absence of radiographic sacroiliitis does not necessarily alter clinical outcomes such as BASDAI and BASFI).

The model structure is summarised in Figure 10 and is applicable to both nr-axSpA and AS populations.

DES requires the generation of a hypothetical patient cohort, in which each patient is assigned a set of clinical characteristics at baseline. These included baseline age, gender, BASDAI, BASFI, weight (kg), disease duration, VAS fatigue, CRP and presence of the human leucocyte antigen-B27 (HLA-B27) gene. These characteristics were selected based on findings of previously published studies of response prediction (192). Although there is the option in the model to sample from a simulated hypothetical cohort[b], the base case analysis samples patients (with replacement) directly from baseline data from studies 1031 and 314-EU in the nr-axSpA and AS populations, respectively (see section 5.4.1). Sampling with replacement means that each sampled patient is replaced back into the data and may be selected again. This approach allows correlation between patient characteristics to be maintained without recourse to distributional assumptions.

The flow of nr-axSpA and AS patients through the model is the same but each population is subject to either AS- or nr-axSpA-specific parameters where possible. The model begins after the failure of two NSAIDs to control symptoms. A patient enters the model on initiation of the first treatment in the treatment sequence which will differ depending on the strategy being evaluated (see Table 100 for comparators by population)[c]. The same patient then moves through the treatment strategy until death. The last treatment possible for all sequences is conventional care. Each patient is exposed to each treatment strategy, before a new patient is sampled.

Although it may also be desirable to model other aspects of the disease such as inflammation or radiological progression, the lack of long-term data and inability to link such outcomes to costs and other economic endpoints meant that such a structure was not feasible.

Response to treatment (in terms of change in BASDAI and BASFI at Week 12) was allocated stochastically (i.e. subject to random variation) for each patient (section 5.4.2.1). Change in BASDAI and BASFI were estimated using a multivariate regression including baseline characteristics. Stochastic variation was introduced by randomly sampling from the

---

[b] The hypothetical cohort, if used, is generated using the method of Tannenbaum et al (214), which assumes multivariate normality to preserve correlation between variables

[c] Note that the model begins at the point of divergence in the treatment sequences, which occurs after the failure of two NSAIDs.

regression residuals. However, to preserve correlation between changes in the two outcome measures, the residuals for both BASDAI and BASFI models were sampled jointly assuming a multivariate normal distribution.

Relative effects from the NMA (or MAIC in the analyses adjusting for differences in study baseline characteristics) were applied to the etanercept initial response estimates to obtain the change in BASDAI and BASFI for other therapies.

Patients were assumed to discontinue therapy if response was not achieved. Response is defined based on the ASAS handbook (215) as:

- Reduction of the BASDAI score to 50% of the pre-treatment value (BASDAI 50) at week 12

A further improvement from Week 12–48 is assumed for patients who continue to receive anti-TNF agents (i.e. anti-TNF responders) in order to prevent underestimating the QALY gain of these patients. Improvements in the BASDAI and BASFI have been observed to continue for approximately 2 years following initiation of anti-TNF agents (9, 122). This is estimated as the change from baseline in BASDAI and BASFI from baseline at Week 48 for etanercept. This estimated change is dependent on the patient's Week 12 BASDAI and BASFI scores and other baseline characteristics. The change from baseline at Week 48 for all anti-TNF agents is assumed to be the same as etanercept. This was based on evidence from the only long-term study identified during the clinical SR to compare two anti-TNF agents (etanercept and infliximab) which suggested that differences observed at Week 12 were not present at Week 48, and was confirmed at the etanercept advisory board (section 5.3.3) (104). See section 5.4.3.2 for details.

Patients in the conventional care arm did not receive continued improvement after Week 12. Instead, BASDAI and BASFI were assumed to revert to baseline immediately at Week 12. The assumption of 0% response in the conventional care arm after Week 12 has been made previously in economic evaluations in both nr-axSpA (216) and AS (1). Although response in BASDAI and BASFI has been observed in the control arm of clinical trials, objective measure of disease severity such as CRP remain unchanged in these subjects (see studies in section 3). It is therefore considered unlikely that the sustained improvements in measures of disease activity and functioning would be observed in the long-term (see section 3.5.14). This assumption was confirmed at the etanercept advisory board (section 5.3.3). However, in order to ensure that this assumption is fully tested four alternative assumptions were explored in sensitivity analysis:

- No BASDAI or BASFI improvement on CON at week 12
- BASDAI and BASFI scores return gradually to baseline between Weeks 12 and 48
- 20% of patients on CON retain their 12 Week response beyond Week 48
- 50% of patients on CON retain their 12 Week response beyond Week 48

Following determination of treatment response at Week 12, patients are exposed to the risk of the following:

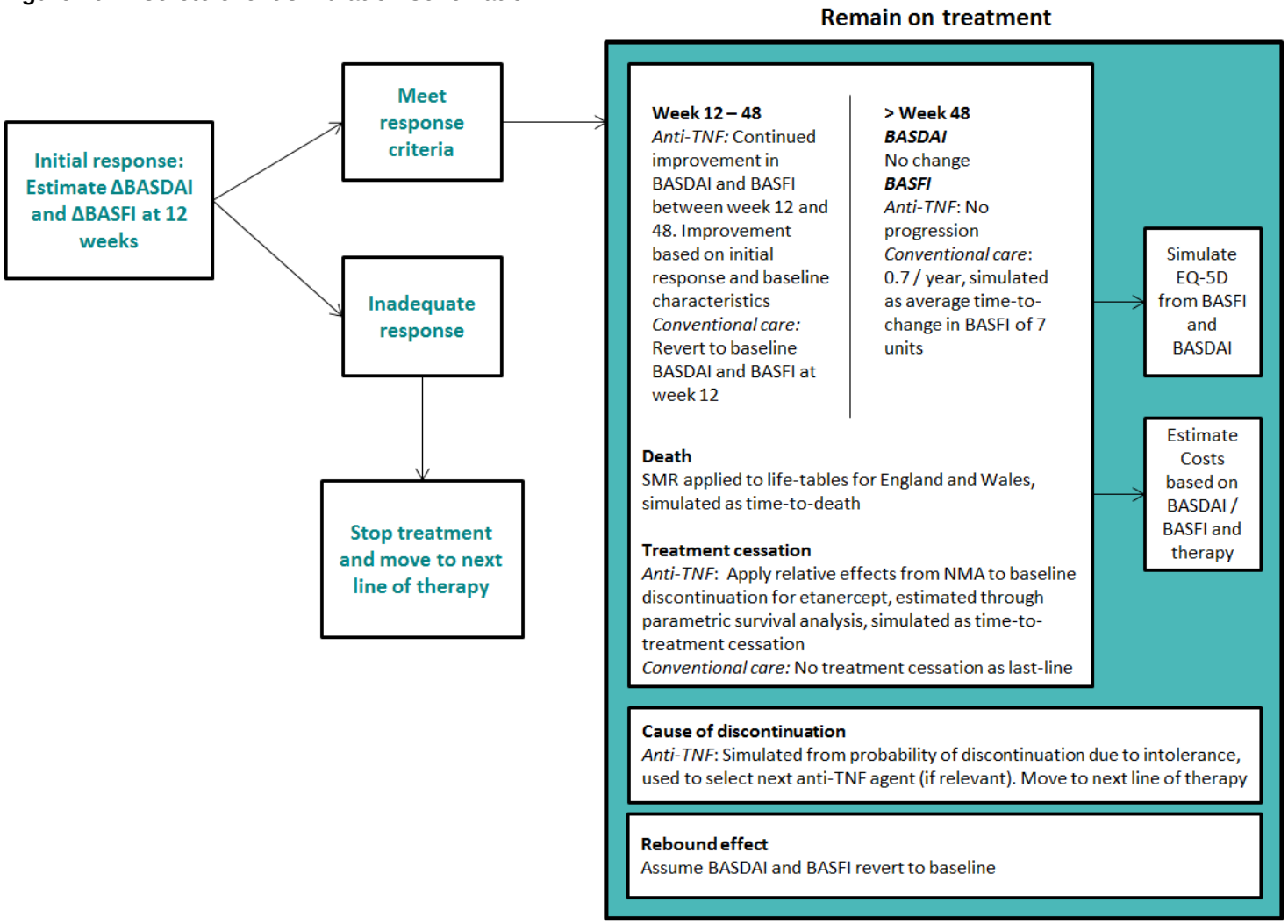
- Treatment discontinuation (section 5.4.5)
- BASFI progression (BASDAI scores were assumed to stay constant; section 5.4.3.3)
- Death (section 5.4.4)

On discontinuation of anti-TNF treatment, BASDAI and BASFI are assumed to return to baseline (i.e. the BASDAI and BASFI values that the individual patient had when they entered the model). This is an assumption which has been made in previous publications (193, 197, 200, 203) and which is supported by recent observations from the Danish Nationwide Registry of Biological Therapies (DANBIO) (204). The time taken to rebound back to baseline is assumed to be 6 months as has been assumed previously in the TA233 submission to NICE (4). The etanercept advisory board suggested an alternative of 3 months, and therefore the time taken to rebound is explored as a scenario analysis. In addition, the cause of discontinuation is determined in order to decide whether use of a further anti-TNF is permitted.

In the base case analysis, costs and outcomes for each treatment strategy are reported separately. Incremental analysis was performed and the ICERs versus conventional care were calculated. Other outcomes that are reported include change in BASDAI and BASFI at Weeks 12, Week 48 and at death, the proportion achieving BASDAI 50 and disaggregated costs and QALYs.

Additional detailed methodological notes are described in Appendix L.

Figure 10: Discrete event simulation schematic



Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; EQ-5D, EuroQoL-5D; NMA, network meta-analysis; SMR, standardised mortality ratio; TNF, tumour necrosis factor.

5.3 Key data sources used in the model

5.3.1 Clinical trial data

The economic model uses analysis of patient-level data from the following studies:

- 314-EU (section 3.5.1)
- 311-EU & 312-EU & 907-EU (sections 3.5.7, 3.5.8, 3.5.9, respectively)
- Study 1031 (section 3.3.1)

Further details of clinical data used to inform the model is provided in section 5.4.

5.3.2 Systematic reviews and data synthesis

A series of systematic reviews (SR) were conducted to identify data on the treatment of nr-axSpA and AS. These reviews are discussed in detail in the following sub-sections.

5.3.2.1 Clinical SR

The clinical SR is discussed in detail in section 2. The results of the NMA are presented in Section 4.2.5. Please note that NMA in nr-axSpA was considered potentially unreliable because of differences in baseline characteristics between etanercept, adalimumab, and certolizumab clinical trials and challenges caused by the reclassification of participants within adalimumab and certolizumab studies. These issues are discussed in more detail in section 4.2.6. As a result, the MAIC and STC analyses were also performed. Section 4.2.7.1 reports the methods and results of these analyses, however the MAIC was used in the analyses adjusting for differences in study baseline characteristics within the economic evaluation as the more conservative (for etanercept) of the two approaches.

5.3.2.2 HRQoL SR

Identification of studies

A systematic review was conducted to identify HRQoL studies from the published literature relevant to the decision problem. Studies reporting health state utility values (HSUVs) for patients with AS and/or nr-axSpA were considered eligible for inclusion.

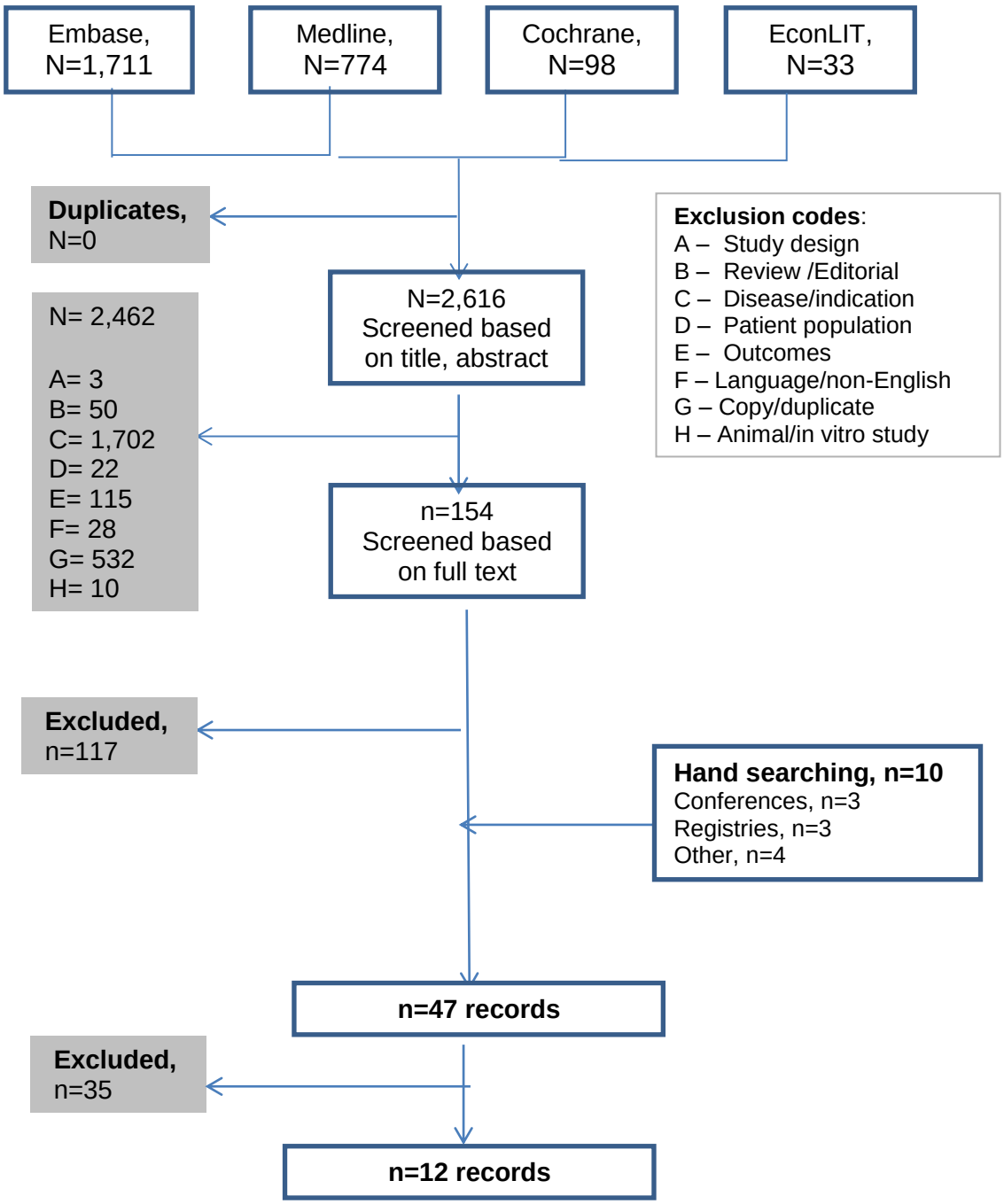
The following databases were searched; Ovid MEDLINE(R)/MEDLINE(R) In-Process & Other Non-Indexed Citations, EMBASE, EconLIT and Cochrane Evidence Based Medicine review (which included the databases NHS Economic Evaluation Database [NHS EED] and Health Technology Assessment [HTA]). Of particular interest to the review were studies that reported EQ-5D HSUVs, in line with the NICE methods guide. Using Boolean operators, the searches combined terms (including Medical Subject Headings [MeSH] as appropriate) for AxSpA and QoL. The search strategy is provided in Appendix G.

Electronic database searches were supplemented by hand searching of relevant HTA publications and their associated ERG reports and manufacturer submissions. In addition, hand searching of bibliographies of included studies, relevant conference proceedings and registries was carried out. Full details of the databases, dates, additional searches, inclusion/exclusion criteria and data abstraction strategy are provided in Appendix G. A summary of the systematic review procedure is shown in Figure 11.

In total, 2,616 citations were retrieved from the electronic database searches. No citations were removed as duplicates, therefore 2,616 titles and abstracts were reviewed. Of these,

154 were obtained for full review. After exclusion of 117 non-relevant citations, 10 extra citations (4 full publications, 1 poster and 5 abstracts from conference proceedings, registries and reference lists of included studies) were included from hand searching. Hand searching of UK HTA agencies yielded no additional eligible publications. Therefore, a total of 47 publications were included.

Figure 11: Schematic of the systematic review of HRQoL evidence



Description of identified studies

The systematic review identified 47 records (36 full, 1 poster and 10 abstracts) that included and presented utilities for patients with AS and/or nr-axSpA (22, 98, 110, 185, 193, 197, 200, 201, 217-255). Full inclusion and exclusion criteria are provided in Appendix G, section 14.1.6

Only three papers reported utilities for patients with nr-axSpA (185, 224, 237). These all used EQ-5D to describe health states but none clearly reported the population by which health states were valued or an elicitation technique. Dougados et al, 2013, was the only publication of the 47 included studies which did not report utilities for AS patients and reported only for nr-axSpA patients (185). Therefore, a total of 46 papers reported utilities for AS patients.

Of the 47 included publications, 41 reported utilities from patients from Organisation for Economic Co-operation and Development (OECD) countries and six of these reported utilities from UK patients (193, 197, 201, 228, 229, 254). Utilities were estimated using a variety of instruments including the EQ-5D (n=39), SF-6D (n=6) and HUI (n=6). Three studies used both EQ-5D and SF-6D (22, 233, 236) and one study reported both HUI and SF-6D (251). The valuation of health states varied according to country-specific tariffs.

Three studies reported a mapping algorithm predicting the relationship between clinical measures of disease severity (BASDAI/BASFI) and HRQoL in patients with AS (193, 200, 201). One study developed an algorithm to predict EQ-5D utilities employing BASDAI, BASFI and age and gender as explanatory variables, one study predicted HUI utilities employing BASDAI, BASFI, gender and race as explanatory variables and one study predicted EQ-5D utilities using only BASDAI and BASFI as explanatory variables.

A full list of excluded papers is available in Appendix G, section 14.1.8 together with rationale for exclusion.

Intervention-specific HSUVs (n=22)

Of the 47 studies that met the inclusion criteria of the review, 22 estimated intervention-specific HSUVs; nine reported data for etanercept (22, 98, 110, 185, 193, 222, 223, 250, 254) (two for etanercept and infliximab) (222, 250) and nine papers reported HSUVs for adalimumab (n=6) (200, 248, 249, 251, 252, 255), infliximab (n=2) (197, 201) and ustekinumab (n=1) (244).

Four of the 22 papers did not specify a particular treatment but stated the type of medication that patients had received (e.g. DMARDs, biological agents, NSAIDs) (225, 233, 236, 239). Follow-up ranged from 12 weeks to 2 years for studies reporting on etanercept, from 12 weeks to 24 weeks for those reporting on adalimumab, and 24 weeks for those reporting on infliximab and ustekinumab. Fifteen of these studies described health states using the EQ-5D, six used HUI and five used SF-6D. Four of the studies were conducted in the UK (193, 197, 201, 254) and two specified the use of a UK tariff (22, 197).

Studies consistent with the NICE reference case (n=12)

Studies considered most appropriate to provide HRQoL evidence for patients with AS or nr-axSpA in the UK included studies that were consistent with the NICE reference case in terms of requirements for HSUV evidence, i.e. health states should be described by patients and valued using UK societal values (256). In addition, studies reporting mapping algorithms to predict utilities from disease-specific instruments for patients with AS were considered as

potentially appropriate. Of the 47 studies that met the inclusion criteria, 12 met these additional criteria (22, 193, 197, 200, 201, 220, 228, 229, 231, 232, 246, 254).

Of the 12 studies, nine reported utilities consistent with the NICE reference case (22, 197, 220, 228, 229, 231, 232, 246, 254), and three studies reported mapping algorithms (193, 200, 201). Details and utility data from the studies meeting the requirements of the NICE reference case are listed in Table 101, and details of the mapping algorithms in Table 102.

Six studies were conducted in the UK, and six in OECD countries. Follow-up time of the studies ranged from 24 weeks to 10 years. Health states were valued based on UK societal preferences; this was clearly stated in one publication (22) and was assumed in nine studies, either from reference to a relevant publication (56, 197, 228, 231, 232, 246), or because the study was conducted in the UK (193, 229, 254). The remaining studies did not report a health state valuation method (200, 201).

In terms of treatment, two studies reported utilities for patients pre-treated with etanercept (22, 254), two studies reported for patients pre-treated with infliximab (197, 201) and one study reported for patients pre-treated with adalimumab (200). The remaining seven studies reported utilities for treatment-naïve patients or did not report details of patients' previous treatment.

The impact on utility could not be assessed for all agents, as EQ-5D data was not reported for all relevant treatments, and HSUVs for patients subsequent to an initial response were not available for all treatments. Therefore, long term changes in utility for all treatments were unable to be incorporated into the model directly, and so some method of extrapolation is required. Mapping algorithms that estimate utility based on disease severity indicators (BASDAI/BASFI) may therefore be considered the most appropriate for this purpose.

Of the three mapping studies, the first by Ara et al, 2007 (193) reports a mapping algorithm derived from a trial by van der Heijde et al, 2006, (257) that converts BASDAI/BASFI into EQ-5D utilities. No additional predictor variables were incorporated into the algorithm. The second study by Marcellusi et al, 2012 is a poster reporting a mapping algorithm converting BASDAI/BASFI to HUI. However, the utility instrument is not clear from the poster (200). The algorithm also incorporates gender and race as predictor variables. The final mapping study by McLeod et al, 2007 (Schering-Plough) reports a mapping algorithm that converts BASDAI/BASFI to EQ-5D and incorporates gender and age as predictor variables (201). Details of the mapping algorithms from the three mapping studies are provided in Table 102 (193, 200, 201). The main limitations of these studies are:

- The relevance to the NICE reference case is unclear (200, 201), or the study does not meet the NICE reference case (193)
- No utilities are reported (200, 201), or those that are reported are based on a different patient population (193)
- Lack of detail in reporting the model, mapping methods and data used.

Thirty-five studies met the inclusion criteria of this SR, but were not considered appropriate to support this submission as they reported HSUVs that were not consistent with the NICE reference case. Details of these studies are provided in Appendix G, Table 164.

Table 101: Utility data from studies meeting the NICE reference case

Note: Boonen et al, 2003, (220) and van Tubergen, 2002, (246) have been included as primary data sources of Boonen et al, 2007, (258) which would have originally met the inclusion criteria but was subsequently excluded on identification of these full primary publications.

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
Ara, 2007 (193) (UK)	Patients from two trials with active AS diagnosed by pain scores and the modified NY criteria (119, 257): (1) van der Heijde: (1a) placebo: age, 39.5 years; disease duration, 8.5 years; BASDAI, 61.1; BASFI, 60; (1b) ETN 1x50 mg: age, 41.7 years; disease duration, 9.0 years; BASDAI, 62.4; BASFI, 60.6; (1c) ETN 2x25 mg: age, 39.7 years; disease duration, 10.0 years; BASDAI, 59.4; BASFI, 57.7; (2) Davis: (2a) placebo: age, 42 years; disease duration, 10.5 years; BASDAI, 59.6; BASFI, 56.3; (2b) ETN 2x25 mg: age, 42 years; disease duration, 10.1 years; BASDAI, 58.1; BASFI, 51.7	• ETN plus NSAIDs - treatment • NSAIDs alone - comparator	NR/ 356 enrolled in van der Heijde (2006)/ NR	Mapping algorithm derived from van der Heijde (2006) to convert BASDAI/ BASFI to EQ-5D (NA)	AS patients, post-treatment, 12 weeks: Treatment non-responder	0.48 (NR) [NR] ((NR))	• This study does not meet the NICE reference case; the study reports a secondary mapping algorithm derived from a trial (12, 257). Although patients described the health states using BASDAI and BASFI, it is unclear from which population the EQ-5D health states valuations are based on. • Mapped utilities may be comparable to an active AS population who have not previously been treated with TNFa inhibitors or DMARDs for at least four weeks. They are based on patients from 11 countries in Europe, with an unknown number from the UK. Therefore, the generalisability of the utilities to the UK population is unknown. Limitations that restrict the usefulness of the mapping algorithm to inform economic evaluation include: ○ published algorithm is reported in an economic evaluation only with no detail reported on the method used to estimate the algorithm (e.g. model description, variables tested, goodness of fit not reported, etc.) ○ response rates, loss to follow up and missing data are not reported, potentially causing bias in the results ○ measures of uncertainty are not reported for utility values.
					AS patients, post-treatment, 12 weeks: Treatment responder	0.79 (NR) [NR] ((NR))	
					AS patients, post-treatment, 12 weeks: Comparator non-responder	0.46 (NR) [NR] ((NR))	
					AS patients, post-treatment, 12 weeks: Comparator responder	0.74 (NR) [NR] ((NR))	
					AS patients, post-treatment, 24 weeks predicted utility from BASDAI and BASFI: Treatment non-responder	0.46 (NR) [NR] ((NR))	
					AS patients, post-treatment, 24 weeks predicted utility from BASDAI and BASFI: Treatment responder	0.80 (NR) [NR] ((NR))	
					AS patients, post-treatment, 24 weeks predicted utility from BASDAI and BASFI: Comparator non-responder	0.42 (NR) [NR] ((NR))	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
					AS patients, post-treatment, 24 weeks predicted utility from BASDAI and BASFI: Comparator responder	0.79 (NR) [NR] ((NR))	
					AS patients, up to 25 years later	NA (available from a graph only)	
Boonen, 2008 (22) Europe and North America	Patients with active AS who met NY modified criteria and who had completed 24 weeks of treatment with ETN or placebo, (1) ETN/ETN group: age, 40.90 years (SD 9.02); disease duration, 10.90 years (SD 7.91); disease severity indicators not reported; (2) PLA/ETN group: age, 41.03 years (SD 10.92); disease duration, 10.75 years (SD 8.83); disease severity indicators not reported	• ETN (25 mg twice weekly) following prior ETN • Placebo treatment following prior ETN	277 potentially eligible/ 257 enrolled/ 220 completed 72 weeks	EQ-5D and SF-6D (UK tariff, TTO)	AS patients, pre-treatment, baseline: ETN/ETN (EQ-5D)	0.69 (0.20) [NR] ((NR))	• This study meets the requirements of the NICE reference case; patients described the health state and health states were valued using UK societal preferences. • Utilities may be comparable to an active AS population who are receiving treatment with ETN. • Generalisability of results to a UK population is unknown. • Limitations that restrict the usefulness of the study to inform economic evaluation include the open-label study design and the absence of HRQOL and utility data in patients who withdrew from treatment due to adverse events or efficacy insufficiency, as these factors may introduce bias into the results. SD is the only measure of uncertainty reported which may also limit usefulness of the utility values. • Extension study, so original paper (10) is needed to fully understand the population used. • Response rates to EQ-5D were good and are unlikely to bias the results.
					AS patients, pre-treatment, baseline: PLA/ETN (EQ-5D)	0.49 (0.30) [NR] ((NR))	
					AS patients, pre-treatment, baseline: ETN/ETN (SF-6D)	0.71 (0.13) [NR] ((NR))	
					AS patients, pre-treatment, baseline: ETN/PLA (SF-6D)	0.65 (0.13) [NR] ((NR))	
Boonen, 2002/ 2003 (220, 221) Europe	Patients with AS diagnosed by modified NY criteria (1) Completing the study: mean age, 43.1 years (range 18-77); mean disease duration, 11.1 years (SD 8.9); mean BASFI, 3.3 (SD 2.6); mean BASDAI, 3.3 (SD 2.1);	NA	NR/ 216 enrolled/ 209 completed follow up	EQ-5D (UK tariff assumed by reference to Dolan 1997 (259), TTO)	AS patients, overall, time averaged over 2 years	0.67 (0.19) [NR] ((NR))	• Relevance unclear - this study does not completely meet the requirements of the NICE reference case; although patients do describe the health states, it is unclear by which population the health states are valued. • Findings may be generalisable to an AS population diagnosed with the modified NY criteria but generalisability of results to a UK population is unknown.
					AS patients: the Netherlands, time averaged over 2 years	0.68 (0.16) [NR] ((NR))	
					AS patients: France,	0.63	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
	(2) Baseline: (2a) the Netherlands; age, 46 years (SD 12); disease duration, 12 years (SD 9); BASDAI, 3.7 (SD 2.1); BASFI, 3.9 (SD 2.4); (2b) France; age, 38 years (SD 12); disease duration, 9 years (SD 7); BASDAI, 2.8 (SD 2.3); BASFI, 2.5 (SD 2.9); (2c) Belgium; age, 42 years (SD 14); disease duration, 11 years (SD 10); BASDAI, 3.1 (SD 1.9); BASFI, 2.6 (SD 2.3).				time averaged over 2 years	(0.23) [NR] ((NR))	• Limitations that restrict the usefulness of the study to inform economic evaluation include: ○ SD is the only measure of uncertainty reported, which may also limit usefulness of the utility values ○ the high percentage of missing data and number of questionnaires missing due to administrative error. • Response rates to questionnaires were good and are unlikely to bias the results.
					AS patients: Belgium, time averaged over 2 years	0.67 (0.14) [NR] ((NR))	
					AS patients: the Netherlands, baseline	0.69 (0.16) [NR] ((NR))	
					AS patients: France, baseline	0.63 (0.29) [NR] ((NR))	
					AS patients: Belgium, baseline	0.67 (0.14) [NR] ((NR))	
Haywood, 2002 (228) UK	Patients with AS diagnosed by modified NY criteria: mean age, 46.1 years (SD 12.6); mean AS symptom duration, 19.8 years (SD 11.8); disease severity not reported	NA	451 patients contacted initially/ 303 patients returned survey at 2 weeks and 289 at 6 months/ 321 patients included in 2-week test-retest analysis; 254 (AS HT) and 248 (general HT) included in 6-month responsive-ness analysis	EQ-5D (UK tariff [assumed from reference to Kind 1998 (260), TTO)	AS patients, test-retest reliability, 2 weeks after initial questionnaire	0.53 (0.35) [NR] ((NR))	• This study meets the requirements of the NICE reference case; patients described the health state and health states were valued using UK societal preferences. • Utilities may be comparable to an adult AS population diagnosed with the modified NY criteria, but not those with learning difficulties or the inability to read written English. • SD is the only measure of uncertainty reported which may limit usefulness of the utility values for economic evaluation. • Findings may be generalisable to a UK population. • There was some loss to follow-up which could bias results, although only 2% of EQ-5D data was missing.
					AS patients: Responsiveness analysis, 6 months after initial questionnaire; AS health transition (better)	0.30 (1.2) [NR] ((NR))	
					AS patients: Responsiveness analysis, 6 months after initial questionnaire; AS health transition (about the same)	-0.25 (1.5) [NR] ((NR))	
					AS patients: Responsiveness analysis, 6 months after initial questionnaire; AS	-0.09 (1.6) [NR] ((NR))	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
					health transition (worse)		
					AS patients: Responsiveness analysis, 6 months after initial questionnaire; general health transition (better)	0.35 (1.3) [NR] ((NR))	
					AS patients: Responsiveness analysis, 6 months after initial questionnaire; general health transition (about the same)	-0.21 (1.4) [NR] ((NR))	
					AS patients: Responsiveness analysis, 6 months after initial questionnaire; general health transition (worse)	-0.15 (1.7) [NR] ((NR))	
Healey, 2013 (229) UK	Patients with confirmed diagnosis of AS using the modified NY criteria: mean age, 46.1 years (SD 12.6); mean AS symptom duration, 19.8 years (SD 11.8)	NA	NR/ 159 enrolled/ 117 invited for follow up; 69 completed follow up	EQ-5D (valuation and elicitation technique unclear but UK tariff assumed as UK study)	AS patients: Assessment 1, 1998	0.64 (0.28) [NR] ((NR))	• Relevance unclear - this study does not completely meet the requirements of the NICE reference case; although patients do describe the health states, it is unclear by which population the health states are valued (UK societal preferences are assumed as the study is conducted in the UK). • Utilities may be comparable to an adult AS population with confirmed diagnosis with the modified NY criteria under the care of a specialist rheumatology centre. • Findings may be generalisable to a UK population. • Limitations that restrict the usefulness of the study to inform economic evaluation include: ○ only approximately 25% of the original cohort of patients invited to participate took part in both study assessments
					AS patients: Assessment 2, 2008	0.61 (0.30) [NR] ((NR))	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
							○ SD is the only measure of uncertainty reported.
Kobelt, 2004 (197) UK	Patients with AS: mean age, 57 years (SD 11.2); disease duration, 30.2 years (SD 11.7); mean BASDAI score, 4.2 (SD 2.3); mean BASFI score, 4.4 (SD 2.8); diagnosis method not reported	NA	2,300 contacted from previous studies/ NR/ 1,413 included in analysis	EQ-5D (UK tariff [assumed from reference to Dolan, 1995 (261), TTO)	AS patients: BASDAI <3	0.80 (NR) [NR] ((NR))	• This study meets the requirements of the NICE reference case; patients described the health state and health states were valued using UK societal preferences. • Utilities may be generalisable to a UK population. • Limitations that restrict the usefulness of the study to inform economic evaluation include: ○ the analysis combines data from a clinical trial, a cohort study and an observational survey with substantially different mean disease scores and patient characteristics ○ no measures of uncertainty are reported for utility values. • As analyses are based on a small clinical trial (Braun, 2002) (95), results can only be considered representative for the type of unremitting disease included in the clinical trial. • A large proportion of patients initially contacted did not return the survey and loss to follow-up data was not reported, potentially causing bias in the results.
					AS patients: BASDAI 3-3.99	0.70 (NR) [NR] ((NR))	
					AS patients: BASDAI 4-4.99	0.64 (NR) [NR] ((NR))	
					AS patients: BASDAI 5-5.99	0.60 (NR) [NR] ((NR))	
					AS patients: BASDAI 6-6.99	0.51 (NR) [NR] ((NR))	
					AS patients: BASDAI >7	0.39 (NR) [NR] ((NR))	
Kobelt, 2006 (231) Canada	AS patients: mean age, 49.6 years (SD 13.3); mean disease duration, 22.3 years (SD 12.6); mean BASDAI, 4.33 (SD 2.19); mean BASFI, 3.56 (SD 2.53);	NA	1,249 contacted initially/ 545 completed questionnaire/	EQ-5D (UK tariff [assumed from reference to Dolan, 1995 (261)], TTO)	AS patients, overall, baseline	0.65 (0.23) [NR] ((NR))	• This study meets the requirements of the NICE reference case; patients described the health state and health states were valued using UK societal preferences. • Utilities may be comparable to a diagnosed AS population, although generalisability of results
					AS patients, BASDAI and BASFI 0-2, baseline	0.865 (NR) [NR]	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
	diagnosis method not reported		545 included in analysis			((NR))	may be restricted to a Canadian population. • Limitations which restrict the usefulness of the study to inform economic evaluation include: ○ some utility data is only presented in graph format so cannot be extracted ○ SD is the only measure of uncertainty reported for utility values. • A large proportion of patients initially contacted did not complete questionnaire, which may bias results but there was no loss to follow-up.
					AS patients, BASDAI and BASFI 2-4, baseline	NR (available on a graph only)	
					AS patients, BASDAI and BASFI 4-6, baseline	NR (available on a graph only)	
					AS patients, BASDAI and BASFI 6-8, baseline	NR (available on a graph only)	
					AS patients, BASDAI and BASFI 8-10, baseline	0.203 (NR) [NR] ((NR))	
Kobelt, 2008 (232) Spain	Patients with AS: mean age, 47.8 years (SD 12.36); mean disease duration, 18.8 years (SD 10.9); mean BASDAI, 4.3 (SD 2.45); mean BASFI, 3.8 (SD 2.85); diagnosis method not reported	NA	700 contacted initially/ 601 completed questionnaire/ 601 included in analysis	EQ-5D (UK tariff [assumed from reference to Dolan, 1995 (261)], TTO)	AS patients, overall, baseline	0.59 (0.30) [NR] ((NR))	• This study meets the requirements of the NICE reference case; patients described the health state and health states were valued using UK societal preferences. • Utilities may be comparable to an AS population, although generalisability of results may be restricted to a Spanish population. • Limitations that restrict the usefulness of the study to inform economic evaluation include: ○ the cross-sectional and uncontrolled nature of the study did not allow for analysis of differences in outcomes between different treatments being used ○ SD is the only measure of uncertainty reported for utility values. • Response rates to questionnaires were good and are unlikely to bias results.
					AS patients: BASFI <3	NR (available on a graph only)	
					AS patients: BASFI 3-3.99		
					AS patients: BASFI 4-4.99		
					AS patients: BASFI 5-5.99		
					AS patients: BASFI 6-6.99		
					AS patients: BASFI >7		
					AS patients: BASDAI <3	NR (available	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
					AS patients: BASDAI 3-3.99	on a graph only)	
					AS patients: BASDAI 4-4.99		
					AS patients: BASDAI 5-5.99		
					AS patients: BASDAI 6-6.99		
					AS patients: BASDAI >7		
Van Tubergen, 2002 (246) The Netherlands	Dutch outpatients with active AS (diagnosis method not reported): Group 1: age, 47 years (SD 10); disease duration, 11 years (SD 6); BASFI, 4.9 (SD 1.8); Group 2: mean age, 48 years (SD 9); mean disease duration, 12 years (SD 5); mean BASFI, 4.2 (SD2.0); Control group: mean age, 48 years (SD 10); mean disease duration, 10 years (SD 6); mean BASFI, 4.2 (SD 2.1)	• Group 1, spa therapy in Austria • Group 2, spa therapy in the Nether-lands • Control group, continued their usual activities and drug treatment	NR/ 120 randomly allocated/ 111 included in analysis	EQ-5D (UK tariff [assumed from reference to Dolan 1996 (262), TTO)	AS patients, pre-intervention, baseline (-2 weeks): Group 1	0.64 (0.22) [NR] ((NR))	• This study meets the requirements of the NICE reference case; patients described the health state and health states were valued using UK societal preferences. • Generalisability of study findings is restricted to Dutch patients with active AS who follow weekly physical therapy and have a disease duration of <20 years. • Limitations that restrict the usefulness of the study to inform economic evaluation include: ○ absence of response rates to questionnaires ○ SD is the only measure of uncertainty reported, which may limit usefulness of the utility values.
					AS patients, pre-intervention, baseline (-2 weeks): Group 2	0.65 (0.22) [NR] ((NR))	
					AS patients, pre-intervention, baseline (-2 weeks): Control group	0.72 (0.10) [NR] ((NR))	
					AS patients, post-intervention, 4 weeks: Group 1	0.10 (0.24) [NR] ((NR))	
					AS patients, post-intervention, 4 weeks: Group 2	0.02 (0.20) [NR] ((NR))	
					AS patients, post-intervention, 4 weeks: Control group	-0.06 (0.18) [NR] ((NR))	
					AS patients, post-intervention, 16 weeks: Group 1	0.12 (0.24) [NR] ((NR))	

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
					AS patients, post-intervention, 16 weeks: Group 2	0.04 (0.21) [NR] ((NR))	
					AS patients, post-intervention, 16 weeks: Control group	-0.04 (0.19) [NR] ((NR))	
					AS patients, post-intervention, 28 weeks: Group 1	0.10 (0.21) [NR] ((NR))	
					AS patients, post-intervention, 28 weeks: Group 2	-0.03 (0.23) [NR] ((NR))	
					AS patients, post-intervention, 28 weeks: Control group	-0.08 (0.28) [NR] ((NR))	
					AS patients, post-intervention, 40 weeks: Group 1	0.03 (0.23) [NR] ((NR))	
					AS patients, post-intervention, 40 weeks: Group 2	-0.01 (0.27) [NR] ((NR))	
					AS patients, post-intervention, 40 weeks: Control group	-0.03 (0.19) [NR] ((NR))	
Wade, 2011 (254) UK	Patients with AS (or RA, PsA, psoriasis) newly prescribed ETN by their specialist between April and November 2009. AS patients had the following baseline characteristics: mean age, 49.2 years (SD 12.3); disease duration and severity not reported;	ETN (dose NR)	~1,000 initially contacted/ 344 enrolled/ 43 AS patients analysed	EQ-5D (valuation and elicitation technique unclear but UK tariff assumed as UK study)	AS patients, pre-treatment, baseline	0.37 (0.37) [NR] ((NR))	• Relevance unclear - this study does not completely meet the requirements of the NICE reference case; although patients do describe the health states, it is unclear by which population the health states are valued (UK societal preferences are assumed as the study is conducted in the UK). • Utilities may be comparable to an AS population who have been newly prescribed ETN and findings may be generalisable to a UK

Study, country	Population	Intervention	No. contacted/ enrolled/ analysed	Utility instrument (valuation, elicitation)	Health states	Mean HSUV, (SD), [SE] or ((CILL, CIUL)) of HSUV	Quality and Relevance Assessment
	diagnosis method not reported.						population. • Limitations that restrict the usefulness of the study to inform economic evaluation include: ○ unknown true number of eligible patients, potentially leading to recruitment bias ○ potential differences between telephone and web-based responses ○ SD is the only measure of uncertainty reported for utility values. • Response rate to EQ-5D for AS patients was good and is unlikely to bias results.

Abbreviations: AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CILL, confidence interval lower limit; CIUL, confidence interval upper limit; DMARD, disease-modifying anti-rheumatic drug; EQ-5D, European Quality of Life 5 Dimensions; ETN, etanercept; HRQoL, health-related quality of life; HSUV, health state utility value; HT, health transition; NA, not available/applicable; NICE, National Institute for Health and Care Excellence; NR, not reported; NSAID, non-steroidal anti-inflammatory drug; NY, New York; PLA, placebo; PsA, psoriatic arthritis; RA, rheumatoid arthritis; SD, standard deviation; SE, standard error; TNF, tumour necrosis factor; TTO, time trade-off.

Table 102: Mapping algorithms reported in the three included economic evaluations

Study	Instrument	Mapping algorithm	Comments	Limitations
Ara, 2007 (Wyeth) (193)	EQ-5D	Utility = 0.923 - 0.004*BASFI - 0.004*BASDAI	• Employs a linear regression model to generate utility estimates based on BASDAI and BASFI scores only as independent predictors	• This study does not meet the NICE reference case; the study reports a secondary mapping algorithm derived from a trial (12, 257). Although patients described the health states using BASDAI and BASFI, it is unclear from which population the EQ-5D health states valuations are based on. • Mapped utilities may be comparable to an active AS population who have not previously been treated with TNFa inhibitors or DMARDs for at least four weeks. They are based on patients from 11 countries in Europe, with an unknown number from the UK. Therefore, the generalisability of the utilities to the UK population is unknown. • Limitations that restrict the usefulness of the mapping algorithm to inform economic evaluation include: ○ published algorithm is reported in an economic evaluation only with no detail reported on the method used to estimate the algorithm (e.g. model description, variables tested, goodness of fit not reported, etc.) ○ response rates, loss to follow up and missing data are not reported, potentially causing bias in the results.
Marcellusi, 2012 (Abbott) (200)	HUI	Utility = 0.948857 - 0.041528*BASDAI - 0.034481*BASFI + 0.047080*Male - 0.063801*White	• Employs a linear regression model to generate utility estimates using HUI data regressed on BASDAI and BASFI as well as gender and race	• Relevance unclear - this study does not meet the NICE reference case; the study reports a mapping algorithm which converts BASDAI/BASFI to HUI so EQ-5D is not used to describe health states and it is unclear from which population health states are valued. • As no utility data is reported, generalisability and comparability of utility results cannot be defined. • Limitations that restrict the usefulness of the mapping algorithm to inform economic evaluation include: ○ lack of detail of the description of the model ○ omission of age from the model formulation (this prevents any consideration of long-term modifying effects of ageing and growing co-morbidities on patient utilities and may cause bias of results) ○ limited information about the mapping algorithm ○ utility data is lacking.
McLeod, 2007 (Schering-Plough) (201)	EQ-5D	Utility = 0.8772129 - 0.0384087*BASDAI - 0.0322519*BASFI - 0.0278913*Male + 0.0016809*Age	• Employs a linear regression model to generate utility estimates developed from a dataset featuring BASDAI, BASFI, age and gender predictor variables	• Relevance unclear - it is unclear if this study meets the requirements of the NICE reference case; the study reports a mapping algorithm to predict utilities based on data from Kobelt, 2004 (197). Although EQ-5D is used to describe health states it is unclear from which population health states are valued. • As no utility data is reported, generalisability and comparability of utility results cannot be defined.

Study	Instrument	Mapping algorithm	Comments	Limitations
				• Limitations that restrict the usefulness of the mapping algorithm to inform economic evaluation include: ○ insufficient detail of the model and mapping algorithm ○ lack of EQ-5D data ○ Also, given the limited range of possible utility values, using a simple linear regression model to estimate a predicting equation for utilities will in general result in biased and inefficient estimates, as values are not restricted to fall within the possible range of utility values. It is unclear if values used in this model have been restricted to prevent this.

Abbreviations: AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; DMARD, disease-modifying anti-rheumatic drug; EQ-5D, European Quality of Life 5 Dimensions; HUI, Health Utilities Index; NICE, National Institute for Health and Care Excellence; TNF, tumour necrosis factor

5.3.2.3 Economic SR

Identification of studies

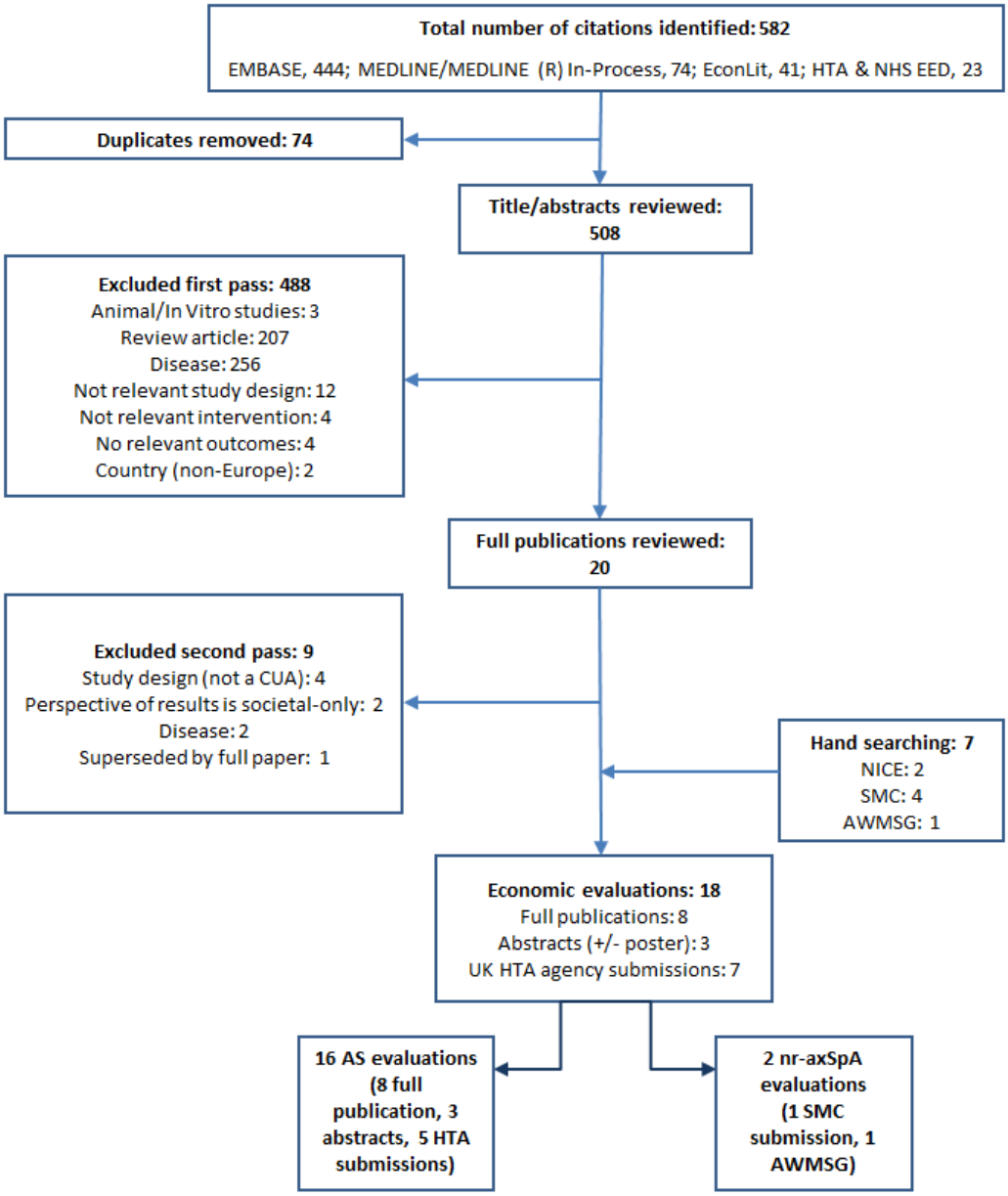
A systematic review was conducted to identify published cost-utility analyses for the treatment of AS and nr-axSpA. Objectives were to identify relevant clinical and cost data inputs, sources and assumptions used in the economic analyses identified.

The following electronic databases were searched; MEDLINE/MEDLINE In-Process, EMBASE, EconLit and Cochrane Evidence Based Medicine review (which included the databases NHS Economic Evaluation Database [NHS EED] and Health Technology Assessment [HTA]). The search strategy is provided in section 15.1.3 in Appendix H. Electronic searches were supplemented by hand searching the following sources; conference proceedings, the Cost-Effectiveness Analysis Registry (CEAR) and HTA submissions for the National Institute for Health and Care Excellence (NICE), Scottish Medicines Consortium (SMC), and All Wales Medicines Strategy Group (AWMSG).

Full details of the databases, dates, additional searches, inclusion/exclusion criteria and data abstraction strategy are provided in sections 15.1.1 to 15.1.6 of Appendix H. A summary of the systematic review procedure is shown in Figure 12.

In total, 582 citations were identified through the electronic searches. Upon removal of duplicates, 508 titles and abstracts were reviewed. After first pass, 20 studies remained, nine of which were excluded during second pass, resulting in eleven citations for final inclusion (eight full publications and three conference abstracts). Hand searching of UK HTA agencies also identified seven HTA submissions for inclusion. In total, 18 economic evaluations met the applied criteria.

Figure 12: Schematic of the systematic review of cost-utility evaluations



Abbreviations: AS, ankylosing spondylitis; AWMSG, All Wales Medicines Strategy Group; CUA, cost-utility analysis; HTA, Health Technology Assessment; NICE, National Institute for Health and Care Excellence; nr-axSpA, non-radiographic axial spondyloarthritis; SMC, Scottish Medicines Consortium.

Description of identified studies

The present SR identified 18 payer-perspective European cost-utility analyses for treatment of AS patients (n=16) and nr-axSpA patients (n=2) with TNF inhibitors or biological DMARDs.

Table 103 provides a summary of the publications identified, including the treatment under review, comparators, model type, time horizon and incremental cost-utility ratio (ICUR). A summary of the clinical, utility and cost inputs of each analysis is presented in Appendix I, Table 166.

Of the 11 full papers and abstracts, agents evaluated were etanercept (193, 203), adalimumab (194, 200), infliximab (196-199) and golimumab (195, 263). A further full paper

was the Assessment Group Report conducted by the Liverpool Reviews and Implementation Group (LRiG) which informed the MTA TA143 (201). This paper evaluated etanercept, adalimumab and infliximab and considered the three manufacturer de novo submissions for each of these drugs. Among the UK HTA submissions (for which there was adequate information on the associated economic evaluation reported), agents evaluated were etanercept (208), adalimumab (207) and golimumab (264). The two nr-axSpA evaluations included were both manufacturer submissions for adalimumab to the SMC (265) and AWMSG (216).

Also provided in Table 103 is a summary of the disease severity at baseline which was used in the models, as measured using BASDAI scores and BASFI scores. BASDAI and BASFI are key inputs for economic models in AS, reflecting intervention efficacy and natural disease progression.

Table 103: Summary of included evaluations

Reference, Country	Disease severity at baseline†	Treatment(s) evaluated	Model used	Time horizon for base case (years)	ICURs for base case analysis
AS					
Ara 2007 (193), UK	Inclusion criteria: BASDAI ≥40 (scale 0–100) Mean scores: NR	Etanercept + NSAIDs vs. NSAIDs	Patient simulation	2, 5, 15, 25	£27,594 (2 yrs), £23,649 (5 yrs), £22,580 (15 yrs), £22,704 (25 yrs)
Botteman 2007 (194), UK	Inclusion criteria: Patients who had BASDAI ≥4cm (scale 0–10cm) and VAS score for total back pain ≥4cm (scale 0–10cm) on two separate occasions at least 4 weeks apart, as per BSR guidelines Mean scores: VAS: 75.15mm, BASDAI: 6.94cm and BASFI: 6.04	Adalimumab vs. conventional therapy	Patient simulation	48 weeks, 5 years and 30 years	£47,083 (48 wks), £26,332 (5 yrs), £23,097 (30 yrs)
Fautrel 2010 (196), France	Inclusion criteria: all trial patients who agreed to participate in evaluation Mean scores: BASDAI: 6.2 (1.4), BASFI: 5.6 (1.9), Utility (GH-RS): 0.4 (0.2)	Infliximab (every 6 weeks) vs. infliximab (on demand)	Markov	1	€50,760 (95% CIs, €17,963 to €216,452)
Kobelt 2004 (197), UK	Inclusion criteria: all trial patients Mean scores: BASDAI: 6.3 (1.4) in placebo arm and 6.5 (1.1) in the infliximab arm BASFI: 5.1 (2.2) in placebo arm and 5.5 (1.8) in the infliximab arm	Infliximab vs. placebo	Decision analytic model using retrospective observational study	2	£73,300
Kobelt 2007 (199), UK	Inclusion criteria: NR Mean scores: Braun trial: BASDAI, placebo: 6.3, Infliximab: 6.5. BASFI, placebo: 5.1, Infliximab: 5.5 ASSERT trial BASDAI, placebo: 6.5, Infliximab= 6.6. BASFI placebo: 6.0, Infliximab: 5.7	Infliximab vs. placebo	Markov	Lifetime	With no progression while on treatment, £28,332 (using Braun et al. trial) and £26,751 (using ASSERT trial) With 50% progression while on treatment: £35,332 (using Braun et al. trial) and £34,067 (using ASSERT trial) With same progression in both groups: £49,417 (using Braun et al. trial) and £46,167 (using

Reference, Country	Disease severity at baseline†	Treatment(s) evaluated	Model used	Time horizon for base case (years)	ICURs for base case analysis
					ASSERT trial)
Kobelt 2008 (198), Spain	Inclusion criteria: BASDAI ≥4 Mean scores: BASDAI (SD): 4.3 (2.45); one-third of patients had a BASDAI <3 BASFI (SD): 3.4 (2.85); 46% of patients at a BASFI <3	Infliximab vs. placebo	Decision tree (clinical trial period) and Markov (open extension period)	40	With no progression whilst on treatment, €15,152 (using Braun et al. trial) and €5,307 (using Spanish trial) With 50% progression while on treatment: €22,520 (using Braun et al. trial) and €8,866 (using Spanish trial) With same progression in both groups: €31,721 (using Braun et al. trial) and €13,659 (using Spanish trial)
McLeod 2007 (201), UK (Assessment Group for TA143 model)	Inclusion criteria: BASDAI ≥4 and VAS pain ≥40mm Mean scores: BASDAI: various scores of 4.0–9.0 BASFI: various scores 4.0–9.0	Adalimumab, etanercept, infliximab vs. conventional management	Cohort with Markov-like transitions	Short term model: 1 Long-term model extension: beginning of year 2 to end of year 20	Short-term (1 year) model, vs. conventional treatment: adalimumab, £57,258; etanercept, £57,261; infliximab, £120,109 Long term (years 2-20) model, vs. conventional treatment: adalimumab/etanercept, £98,910; infliximab, £175,000
Neilson 2010 (203), Germany	Inclusion criteria: BASDAI ≥40 Mean scores: NR	Etanercept + usual care (including NSAIDs) vs. usual care alone (including NSAIDs)	Patient simulation	25	€54,815
Conference abstracts					
Farrell 2011 (195), UK	NR	Golimumab vs. adalimumab, etanercept, standard care (NSAIDs, DMARDs, Cox-2 inhibitors and physiotherapy)	Markov	20	vs. conventional therapy: golimumab, £30,043; adalimumab, £30,187; etanercept, £30,810
Marcellusi 2012 (200), Italy	NR	Adalimumab vs. golimumab, placebo	Markov	40	vs. placebo: adalimumab, €33,704; golimumab €39,149 For adalimumab vs. golimumab: adalimumab was dominant
Muszbek 2012 (263), UK	NR	Golimumab vs. other TNF-α inhibitors, conventional treatment	Markov model with an initial decision tree	20	vs. conventional treatment: £15,353

Reference, Country	Disease severity at baseline†	Treatment(s) evaluated	Model used	Time horizon for base case (years)	ICURs for base case analysis
UK HTA agency submissions					
NICE TA143 (4)	Inclusion criteria: BASDAI ≥4.0/10 Mean scores: BASDAI, 6.1/10; BASFI, 5.9/10	Etanercept + NSAIDs vs. NSAIDs (Wyeth submission)	Patient simulation	25	~£13,200
	Inclusion criteria: VAS ≥4/10, BASDAI ≥4/10 Mean scores: Overall: BASDAI, 6.8; BASFI, 5.8 Anti-TNF-agent: BASDAI, 6.67/10; BASFI, 56.4/100 Placebo: BASDAI, 6.94/10; BASFI, 60.4/100	Adalimumab + NSAIDs vs. NSAIDs (Abbott submission)	Patient-based state transition model	30	£23,097
	Inclusion criteria: BASDAI ≥4.0/10 Mean scores: BASDAI, 6.41/10; BASFI, 5.75/10	Infliximab vs. placebo (Schering-Plough submission)	Combined decision tree and Markov chain structure	70	Corrected results†: £50,380 (using Braun et al trial) and £40,889 (using ASSERT trial)
NICE TA143 (4)	BASDAI, various scores of 4.0–9.0 BASFI, various scores 4.0–9.0	Assessment group (LRiG) Adalimumab, etanercept, infliximab	Short-term model, vs. conventional treatment and long term model, vs. conventional treatment	20	Short-term model, vs. conventional treatment: adalimumab, £57,258; etanercept, £57,261; infliximab, £120,109 Long term model, vs. conventional treatment: adalimumab/etanercept, £98,910; infliximab, £175,000
	BASDAI, various scores of 4.0–9.0 BASFI, various scores 4.0–9.0	NICE appraisal committee decision on most plausible scenario following DSU reconciliation of the 3 manufacturer models and Assessment group (LRiG) model Adalimumab, etanercept, infliximab	Assessment of manufacturer and LRiG models	20	Assuming no disease progression after 20 weeks for TNF-α inhibitor treatment responders in the Assessment group's (LRiG) model: adalimumab/etanercept, £22,000; infliximab, £49,000
NICE TA233 (5)	Inclusion criteria: GO-RAISE trial patients	Golimumab (50 mg, once a month) vs. adalimumab (40	Decision tree (short term) and Markov	20‡	Conventional treatment: adalimumab, £26,747; etanercept, £26,600; golimumab, £26,597

Reference, Country	Disease severity at baseline†	Treatment(s) evaluated	Model used	Time horizon for base case (years)	ICURs for base case analysis
	Mean (SD) scores: BASDAI, 6.54 (1.55); BASFI, 5.04 (2.33)	mg, every 2 weeks), etanercept (25 mg twice weekly or 50 mg once weekly), conventional therapy (MSD submission)	(long term)		From an ERG clarification letter, requesting the model adaptation to consider the effect of a 2nd anti-TNF (sequential treatment), manufacturer submitted the following (by halving the discontinuation rate to 7.5%): vs. conventional treatment: adalimumab, £26,547; etanercept, £26,462; golimumab, £26,460 Further requested analysis of golimumab +/- methotrexate, vs. conventional treatment: golimumab + methotrexate, £38,968; golimumab without methotrexate, £26,352
	BASDAI, 6.54 BASFI, 5.04	Golimumab vs. conventional therapy	ERG corrected model: golimumab vs. conventional therapy	60.1‡	£26,954‡
SMC 2005 (208)	NR	Etanercept vs. NSAIDs	Patient simulation	15	£11,700
SMC 2006 (207)	NR	Adalimumab (40 mg by subcutaneous injection every two weeks) vs. conventional therapy (NSAIDs, non-drug therapy)	Patient simulation	30	£23,000 (30 yrs horizon) and £26,000 (5 yrs horizon), £47,000 (48 wks horizon)
SMC 2011 (264)	NR	Golimumab vs. conventional therapy (combination of NSAIDs and conventional DMARDs)	Decision tree/Markov	20	For golimumab 100 mg, vs. conventional DMARDs: £32,546, (was less cost-effective than adalimumab and etanercept) For golimumab 50 mg, ICER was similar to that of adalimumab and etanercept (values not reported)
nr-axSpA (UK HTA agency submissions)					
AWMSG 2013 (216)	BASDAI ≥4	Adalimumab (40 mg every two weeks) vs. conventional care (defined as a combination of physiotherapy and/or NSAIDs, or palliative care in patients with NSAID intolerance)	Markov	40	£16,154

Reference, Country	Disease severity at baseline†	Treatment(s) evaluated	Model used	Time horizon for base case (years)	ICURs for base case analysis
SMC 2013 (265)	NR	Adalimumab (+/- conventional therapy†) vs. continuing conventional therapy only (combination of physiotherapy and NSAIDs, or palliative care for those unable to tolerate NSAIDs)	NR	40	£16,154 (base case) For patients receiving ≥2 prior NSAIDs, in line with ASAS guidelines: £17,000 From additional analysis provided by the manufacturer, which assumed conventional therapy responders retained a level of clinical benefit, produced ICURs of between £25,000-£28,000

Abbreviations: ASAS, Assessment In Ankylosing Spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BSR, British Society of Rheumatology; DMARDs, Disease-modifying anti-rheumatic drugs; DSU, Decision Support Unit; ERG, Evidence Review Group; GH-RS, General health rating scale; ICERs, incremental cost-effectiveness ratios; ICURs, incremental cost-utility ratios; LRiG, Liverpool Reviews and Implementation Group; NR, not reported; NSAIDs, non-steroidal anti-inflammatory drugs; VAS, visual analogue scale.
†An increasing BASDAI/BASFI/VAS score indicates increasing severity of disease. Scores required for inclusion of patients in the model (Inclusion criteria) and mean scores of patients actually included in the model (Mean scores) are provided, where reported.
‡In the base case of the manufacturer submission, the time horizon is 20 years; however golimumab is a lifetime treatment. Therefore, the evidence review group (ERG) conducted an analysis with 60.1 year time horizon as a close approximation of a lifetime time horizon.

ICERs reported in published economic evaluations for anti-TNF treatment

Base-case ICERs reported for adalimumab vs conventional care in nr-axSpA were reported as £16,154 (216, 265). Including the manufacturer's patient access scheme (PAS), certolizumab was reported to be the dominant strategy vs adalimumab in both nr-axSpA and AS populations[d] (188). This study was identified.

Across all AS evaluations, the ranges of reported base case ICERs versus relevant comparators were: etanercept, £22,000 (4) to €54,815 (203); adalimumab, £22,000 (4) to £30,187 (195); infliximab, €5,307 to £49,000 (4) (excluding a 2 year horizon model, reporting £73,300 (197)); golimumab £15,353 (202) to £32,546 (264). The base case ICERs reported are listed in Table 103.

Table 104 summarises the key drivers of cost-effectiveness identified in sensitivity analyses in previously published economic evaluations of AS. The important determinants of cost-effectiveness were annual withdrawal rate/discontinuation rate of treatment, disease state costs and time horizon, baseline BASDAI/BASFI, and treatment costs. A recent HTA for adalimumab in nr-axSpA found that the key drivers of cost-effectiveness were treatment benefits over time, non-treatment costs, utilities and time-dependent discontinuation rate (216).

Table 104: Key drivers of cost-effectiveness from sensitivity analyses reported in published economic evaluations

Parameter	Citation(s)
Patient characteristics	
Baseline BASDAI/BASFI	(5, 195, 201)
Age	(200, 201)
Effectiveness	
Percentage of patients with a long-term 'response'	(201)
Annual withdrawal rate/discontinuation rate of treatment	(193, 198, 201, 203)
Long-term spontaneous recovery assumption for placebo treated patients	(201)
Benefits associated with QoL	(193, 203)
BASDAI/BASFI, BASFI progression rate	(201, 208)
Costs	
Disease (state) costs (medical)	(193, 201, 203, 208)
Treatment costs	(199-201)
Patient weight (requiring a greater dose)	(200)
Administration costs (for infliximab only)	(201)
Other	
Time horizon	(5, 195, 198, 199)

[d] This document was identified following searches of the SMC and AWMSG websites after the initial search dates of the SR.

Parameter	Citation(s)
Discounting practices	(198)
Non-response imputation method	(200)

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; QoL, Quality of Life.

None of the studies attempted to model sequential treatment with anti-TNF agents following failure or intolerability of a treatment. One discrete event simulation (DES), which modelled sequential treatment in AS, has been conducted (266). However, it was excluded from the present systematic review as the results were reported from the societal perspective. In one evaluation, the ERG for NICE TA233 (golimumab in AS) requested additional analyses from the manufacturer concerning sequential use of treatments. The manufacturer provided an estimated effect of sequential treatments by halving the discontinuation rate (which the ERG found to be a poor approximation for sequential treatment); they did not build flexibility into their model to evaluate sequential treatment from the outset.

When comparing results for etanercept by type of model used (patient simulation vs Markov), there does not appear to be a substantial difference between results. Details of only two evaluations for nr-axSpA (both for adalimumab) were identified (216, 265). Given the lack of modelling studies in nr-axSpA, this provides the rationale for development of a new etanercept model for the NICE MTA submission.

5.3.2.4 Quality assessment of economic evaluations

Of the evaluations identified in the SR, eight were cost-utility, full publications relevant for Quality Assessment. The Quality Assessment was performed using Appendix 11 of the NICE Single Technology Appraisal specification for manufacturer submission of evidence, as adapted from the Drummond Checklist (267). The full results of Quality Assessment are provided in Appendix J, Table 168.

5.3.3 Expert opinion

A proposed model methodology was critiqued at an advisory board meeting held on 10th July 2014 (described hereafter as the etanercept advisory board). The advisory board was not used to derive data for the model. However, it was used to confirm and critique model structure and assumptions. The assembled panel was comprised of eight rheumatologists.

5.4 Data used in the model

5.4.1 Patient baseline characteristics

Characteristics of individual patients in each population were sampled (with replacement) directly from the baseline of the relevant studies identified in the SR (see section 5.3.1). This method has the advantage of maintaining correlation between variables without reliance on strong distributional assumptions.

Table 105 presents a summary of the population characteristics assumed within the model for both populations (discussion of data sources is provided in section 5.3.1). The key differences between the nr-axSpA and AS populations were that nr-axSpA patients were likely to be younger, have shorter disease duration and lower BASFI. Nr-axSpA patients also tended to have lower CRP and a lower proportion were male.

Table 105: Baseline characteristics of patients sampled in model

	AS population (study 314-EU)		nr-axSpA population (study 1031)	
	Mean	**SD**	**Mean**	**SD**
Age, years	41	11	32	8
Male, %	74%		60%	
BASDAI	60.93	16.78	59.61	18.15
BASFI	59.24	20.21	40.27	24.78
CRP, mg/L	20.94	22.75	6.59	10.53
Disease duration, years	9.34	8.59	2.45	1.84
Weight, kg	76.42	17.48	[redacted]	[redacted]
HLA-B27 positive, %	[redacted]		[redacted]	

Abbreviations: AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, c-reactive protein; nr-axSpA, non-radiographic axial spondyloarthritis; HLA-B27, human leukocyte antigen B27; SD, standard deviation.

5.4.1.1 Generalisability

Several sources are available to determine the generalizability of the etanercept data used in clinical practice within England and Wales. The most relevant was considered to be the British Society for Rheumatology Biologics Registers (BSRBR) as reported in Lord et al, 2010 (192), however data are only available for 261 patients with AS. Compared to these patients, the 314-EU cohort has lower BASDAI, BASFI, CRP and disease duration. Therefore, these characteristics were considered in a scenario analysis (section 5.7.2) in order to explore the impact of these differences. Larger cohorts are available from DANBIO and the German Spondyloarthritis Inception Cohort (GESPIC), although these are less relevant in the UK setting. Compared to DANBIO, the Study 314-EU cohort had higher BASFI, longer disease duration and higher CRP and compared to GESPIC patients had lower BASDAI, were older and had longer disease duration. The patient characteristics in Study 314-EU and the three registries are shown in Table 106.

Table 106: AS patient characteristics in registry data and study 314-EU

Characteristic	Study 314-EU (12)	DANBIO (Glintborg 2010 et al†) (268)	BSRBR (Lord 2010 et al) (192)	GESPIC (Sieper et al) (269)
Age	41	41	43	35.6
Gender (% male)	74	74.6	82	64
Disease duration (years)	9.32	4.4	13	2.8
BASDAI	60.93	57.8	76	40
BASFI	59.24	49.1	79	-
CRP	20.9	13	23	-
HLA-B27+ %	[redacted]	67.7	-	88.2

Abbreviations: AS, Ankylosing Spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BSRBR, British Society for Rheumatology Biologics Registers; CRP, C-reactive protein; DANBIO, Danish Nationwide Registry of Biological Therapies; HLA-B27, human leukocyte antigen B27; GESPIC, German Spondyloarthritis Inception Cohort
†Weighted average of switchers and non-switchers

5.4.2 Week 12 response

At 12 weeks, the following are estimated in the model:

- Change from baseline in BASDAI at Week 12
- Change from baseline in BASFI at Week 12

These measures were preferred because:

- OMERACT (Outcome Measures in Rheumatology) preliminary proposals recommends use of BASDAI *and* BASFI in economic modelling within AS (192)
- BASDAI and BASFI are measured in almost all clinical studies, and have been demonstrated to be closely correlated to health utilities, mortality and costs
- As noted within the preliminary OMERACT guidance, absolute (as opposed to relative) measures are much preferred for economic modelling as they avoid arbitrary thresholds
- BASDAI and BASFI are regularly collected in clinical practice to measure treatment effects and disease severity

The BASDAI score is used to determine whether an individual is considered a responder at Week 12, and changes in both the BASDAI and BASFI at Week 12 are used in subsequent predictions of Week 48 response (section 5.4.3). Response at Week 12 was defined as a reduction of the BASDAI score to 50% of the pre-treatment value (BASDAI 50), based on the current NICE definition of treatment response (4).

It is noted that the response criteria for short-term treatment in clinical studies are typically defined by the Assessment of Spondyloarthritis international Society (ASAS) response criteria which is a relative response criteria in which the patient either responds or does not (i.e. a binary response). Although such outcomes are the primary endpoints in most contemporary clinical studies, they were not used for this economic model for the following reasons:

- Clinical expert opinion suggested that ASAS response criteria are not widely used in clinical practice

- The ASAS handbook recommends the use of the BASDAI to define response to anti-TNF agents (215)
- ASAS is a (comparatively) new outcome measure, and therefore the use of ASAS could exclude earlier clinical evidence
- In the context of AS, absolute outcome measures are much preferred to relative response criteria in economic evaluations, as noted by Bansback et al, 2007 (192). Using the absolute measures of BASDAI and BASFI allows direct estimation of resource use costs and health utility values in the model, as well as being used to predict response to treatment.

5.4.2.1 Week 12 response for etanercept

Etanercept RCT data were used to predict a baseline initial response for etanercept for both nr-axSpA and AS populations; relative effects from the NMA (or MAIC in the analyses adjusting for differences in study baseline characteristics) were applied to this to estimate initial response to comparator therapies. Results of the NMA are presented in section 4.2.5 and summarised in section 5.4.2.2. In order to account for the correlation between BASDAI and BASFI, a multivariate multiple regression was estimated.

Regression approach (multivariate multiple regression)

The regression analysis considered all subjects within the modified intention-to-treat (mITT) populations and missing data were imputed using a last-observation carried forward approach as used within the primary analysis of the etanercept clinical trial programme.

Multivariate regression is a technique that estimates a single regression model with more than one outcome variable. When there is more than one predictor variable in the model, it is known as a multivariate, multiple regression model. This approach allows the prediction of correlated samples of change in BASDAI and change in BASFI to be drawn. Specifically, correlated changes in BASDAI and BASFI at Week 12 are drawn based on the residuals from the multivariate regression, which are assumed to be multivariate normally distributed.

Regression model selection

The following variables were included in analyses predicting treatment response in economic evaluations identified in the economic SR including NICE TA233 (5).

- Baseline BASDAI
- Baseline BASFI
- Baseline age
- Gender

Additional non-systematic internet searching identified the following additional baseline predictors for achieving clinical response (270):

- Baseline acute phase reactants (CRP)
- HLA-B27

Initially, all variables were included in the regression models and their statistical significance and direction of effect were evaluated.

In the regression for nr-axSpA, the effect of baseline age was not statistically significant. Moreover the direction of effect was different in the BASDAI model compared to the BASFI

model (for nr-axSpA, study 1031). Backwards stepwise selection (using an arbitrary restriction of p<0.2) excluded this variable from both the BASDAI and BASFI models in the nr-axSpA population (study 1031) and this variable was therefore removed. Gender, although not statistically significant, had consistent direction between models and was therefore retained as an *a priori* identified covariate.

Conversely, in AS, the effect of baseline age was highly statistically significant, and the effect of gender, while non-significant, also had different effect direction in the BASDAI and BASFI models and was therefore excluded. CRP was statistically significant in BASDAI, but not BASFI but had consistent direction and so was retained as an *a priori* identified covariate. This gave rise to slightly different model specification between nr-axSpA and AS models.

The results of these analyses confirm the hypothesis that the differences in baseline characteristics between the certolizumab, adalimumab, and etanercept nr-axSpA studies may lead to differences in effect size and hence support the requirement to perform analyses adjusting for differences in baseline characteristics.

The R-squared values are reported below; these indicate the proportion of total variation of outcomes explained by the model, and range between 0 and 1. A value of 0 indicates that the model explains none of the variability in the data. A value of 1 indicates that the model explains all the variability of the data. In general, therefore, higher R-squared values indicate better model fit to the data in question. Note that although these models have only modest explanatory power, they are combined with the resulting residuals (and their variance-covariance matrix) to ensure that model predictions adequately reflect the clinical trial data.

Analysis results

Nr-axSpA population

The results of multivariate multiple regressions for the nr-axSpA population are presented in Table 107. The models of BASDAI and BASFI had R-squared values of [redacted], respectively. To interpret the regression coefficients, consider the coefficient of CRP for change in BASDAI ([redacted]). For a one unit change in CRP, the resulting change in the change in BASDAI is [redacted] (0-100 scale). For example, consider an nr-axSpA patient with mean baseline characteristics (see Table 105), the expected change in BASDAI at Week 12 is given as:

[redacted]

For an individual with a CRP of 5 units higher than the mean (i.e. [redacted]), the expected change in BASDAI, all else being equal, would become [redacted] (because [redacted] $\cdot\, 5 \approx -2$). Please note that for prediction within the model, a stochastic approach is used to determining change at Week 12 (see Appendix L), not simply the linear predictor as described in this illustrative example.

Table 107: Change in BASDAI and BASFI from baseline at Week 12 in a nr-axSpA population (n=103)

	Coefficient	SE	p-value	Lower 95% CI	Upper 95% CI
BASDAI					
Male	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	Coefficient	SE	p-value	Lower 95% CI	Upper 95% CI
Baseline BASDAI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BASFI					
Male	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASFI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; CRP, C-reactive protein; HLA-B27, human leukocyte antigen B27; nr-axSpA, non-radiographic axial spondyloarthritis; SE, standard error.

AS population

The results of multivariate multiple regression for the AS population are presented in Table 108. The models of BASDAI and BASFI had R-squared values of 0.2840 and 0.2353, respectively.

Table 108: Change in BASDAI and BASFI from baseline at week 12 in AS population (n=294)

	Coefficient	SE	p-value	Lower 95% CI	Upper 95% CI
BASDAI					
Age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASDAI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27 positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BASFI					
Age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASFI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27 positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; CRP, C-reactive protein;HLA-B27, human leukocyte antigen B27; SE, standard error.

5.4.2.2 Week 12 response for other therapies (primary analyses)

Relative effects from the NMA results (or MAIC in the analyses adjusting for differences in study baseline characteristics) are applied to the etanercept estimates in order to predict initial response for the other anti-TNF agents and conventional care. In probabilistic sensitivity analysis, treatment effects were drawn from the WinBUGS output (CODA) to

preserve correlation. Full details and discussion are provided in section 4 but are summarised here for convenience.

Table 109 and Table 110 present treatment effects from the NMA for nr-axSpA and AS populations, respectively. Please note that the model uses inputs in the form of comparator versus etanercept and hence these results are the reverse of those presented in section 4.2.5. Also note that the model is constructed on the 0-100 (rather than 0-10) scale for BASDAI and BASFI, hence these differences have been scaled up to reflect this difference.

For the NMA in the AS population, results for both the fixed effect and random effects models were presented. These two models did not differ in their goodness of fit using DIC or deviance. For the purposes of providing inputs into the economic model, the results from the random effects analysis were used. Information suggesting that differences in baseline characteristics between different trials can influence outcome means that the assumptions behind the random effects model are considered more plausible.

Table 109: Relative effects from NMA for nr-axSpA, mean change (with imputation): fixed effects

	BASDAI			BASFI		
	Mean difference vs etanercept	**lower 95% CrI**	**upper 95% CrI**	**Mean difference vs etanercept**	**lower 95% CrI**	**upper 95% CrI**
CON	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CZP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations; ADA, adalimumab; AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CrI, credible interval; CZP, certolizumab pegol; ETN, etanercept; N/A, not applicable; nr-axSpA, non-radiographic axial spondyloarthritis.

Table 110: Relative effects from NMA for AS, mean change (with imputation): random effects

	BASDAI			BASFI		
	Mean difference vs etanercept	**lower 95% CrI**	**upper 95% CrI**	**Mean difference vs etanercept**	**lower 95% CrI**	**upper 95% CrI**
CON	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CZP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL†	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
INF	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations; ADA, adalimumab; AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CrI, credible interval; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; N/A, not applicable; nr-axSpA, non-radiographic axial spondyloarthritis.
† the GOL comparative data are based on the 50 mg dose

5.4.2.3 Week 12 response for other therapies in analyses adjusting for differences in study baseline characteristics

As discussed in section 4.2.6.2, for the nr-axSpA indication, there were some variations in the baseline characteristics of the four RCTs comparing etanercept (161), adalimumab (105, 112) and certolizumab pegol (90) versus placebo included in the NMA.

In general the etanercept RCT recruited participants who were younger, had shorter duration of disease symptoms, and had lower mean BASFI compared to the participants in the other studies. As noted in the etanercept clinical RCT and OLE section in section 3, this was done to reduce the risk of including patients with AS in the nr-axSpA population. This risk was further minimised by the use of centralised and blinded assessment of the radiograph for diagnosis of nr-axSpA. This approach was taken to avoid the situation in the adalimumab and certolizumab RCTs (90, 112) where considerable reclassification of nr-axSpA patients has subsequently been found to have taken place, thereby potentially confounding the results in the current nr-axSpA NMA.

The FDA assessment report of the adalimumab RCT (139) noted the following limitations:

- Of the 185 nr-axSpA participants included in the efficacy analysis, 102 were centrally assessed, and 64/102 were diagnosed with nr-axSpA, 38 with AS.
- The results on key outcomes (ASAS 40, ASAS 20, BASDAI 50, ASAS partial remission, ASAS 5/6) in this nr-axSpA subgroup found no statistically significant differences between adalimumab versus placebo compared to the total clinical trial population and the AS subgroup where statistically significant differences were reported for key efficacy outcomes

The data from this RCT publication (112) included in the current NMA are therefore confounded by the presence of a significant number of AS patients which is likely to invalidate the indirect comparison of adalimumab versus the other comparator treatments in the current nr-axSpA NMA. Since the FDA noted that the centrally-read nr-axSpA post-hoc subgroup may not preserve the baseline comparability between treatment arms achieved by randomisation (139), it was deemed inappropriate to conduct an NMA using this post-hoc analysis. The present NMA is therefore limited by the use of the heterogeneous data from the Sieper et al, 2013, trial publication (112).

Similarly, the FDA assessment report of the certolizumab RCT (90) noted the following limitations:

- Of the original N=147 nr-axSpA subgroup in the RCT publication (based on local X-ray), only 98 patients (66.6%) retained this diagnosis on central X-ray in a post hoc analysis
- The original AS subgroup of N=178, increased to N=184 (central reads were not performed for 43 patients as X-rays were unavailable)
- While results in the central X-ray nr-axSpA subgroup were still significantly different versus placebo in line with the original publication, for some key outcomes (e.g. BASFI, BASDAI) the magnitude of effect was smaller than observed in the original nr-axSpA subgroup

The data from the nr-axSpA subgroup in the RCT publication (90) included in the current NMA are therefore confounded by the presence of a significant number of AS patients which is likely to invalidate the indirect comparison of certolizumab versus the other comparator treatments in the current nr-axSpA NMA. As with the adalimumab centrally-read nr-axSpA post-hoc subgroup, the present certolizumab nr-axSpA post-hoc analysis may also fail to preserve baseline comparability between treatment arms and randomisation is likely to be compromised. Therefore, it was deemed inappropriate to conduct an NMA using this post-hoc analysis. In summary, the present NMA is limited by the use of the heterogeneous data from the Landewe et al, 2012, trial publication (90).

As a result, MAIC and STC analyses were performed in attempt address these issues. These methodologies are reported in section 4.2.7. For the purposes of the analyses adjusting for differences in study baseline characteristics in the economic evaluation, the more conservative estimates generated by the MAIC were retained. Results of the MAIC are reported in full in section 4.2.5 but are summarised in Table 111.

Table 111: MAIC data used in analyses adjusting for differences in study baseline

	Certolizumab vs etanercept (AxSpA population)			Adalimumab vs etanercept (nr-AxSpA population)		
Outcome (change from baseline at week 12)	**Mean difference**	**lower 95% CI**	**upper 95% CI**	**Mean difference**	**lower 95% CI**	**upper 95% CI**
BASDAI (0-100)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BASFI (0-100)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AxSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; nr-axSpA, non-radiographic axial spondyloarthritis.

As discussed in section 4.2.7.3, the MAIC leads to smaller differences in effect size than the NMA, particularly in the comparison against certolizumab. The analyses confirmed that differences in baseline characteristics make a relevant difference in the relative effect of treatments compared to standard NMA approaches. The analyses found no evidence of statistically significant difference in the efficacy between the comparator treatments considered. As a consequence, the apparent differences between etanercept and adalimumab, and etanercept and certolizumab in treatment efficacy in the nr-axSpA and AxSpA populations cannot be interpreted at face value, and require adjustment for differences in baseline characteristics between the respective trials.

Please note that for the purposes of the analyses adjusting for differences in study baseline characteristics (see section 5.2.1.1), in which etanercept is compared against certolizumab in an AxSpA population, it was necessary to replicate the mix of nr-axSpA and AS patients seen in clinical practice. The ratio of nr-axSpA to AS patients was set at 50:50. As this was a source of uncertainty, alternative ratios of 70:30 and 30:70 were also considered.

5.4.2.4 Week 12 response following the failure of the first anti-TNF

In the base-case, second line anti-TNF use is only permitted if failure was due to intolerance. In such cases, it was assumed that second line efficacy would not be affected. In the scenario which considered second line use of anti-TNF agents due to failure of the first from any cause, it is assumed that there is a reduced level of response for second line therapies if discontinuation is for causes other than intolerance (i.e. loss of efficacy). An adjustment is therefore applied to the baseline efficacy to reflect this based on a similar approach to that used by Trans-Duy et al, 2011, (266). An adjustment is made to the change in BASDAI and BASFI based on data from Glintborg et al, 2013 (204) to reduce second line efficacy in these patients. The adjustments -30.07% and -32.03% are applied to the change in BASDAI and BASFI respectively at Week 12.

5.4.3 Week 48 response

Following an initial 12 weeks of anti-TNF treatment, longer term follow-up studies, including study 1031, demonstrate that improvements in the BASDAI and BASFI continue to be observed for at least one year of therapy (9, 122, 126, 141, 161, 271). It was therefore considered necessary to model benefits associated with treatment beyond Week 12. In order

to approximate this effect, change in BASDAI and BASFI between Week 12 and Week 48 was modelled.

The modelling of change in BASDAI and BASFI from baseline at Week 48 was conducted using the same approach used for the Week 12 treatment response (see section 5.4.3.2). However, change in BASDAI and BASFI from baseline at Week 12 were included as covariates within the resulting models in order to ensure that an individual's response at Week 48 was dependent on their response at Week 12[e].

In the absence of relative effect estimates at Week 48 for other therapies, it was assumed that patients who remained on anti-TNF treatment beyond week 12 (i.e. anti-TNF responders) would converge at the BASDAI and BASFI levels predicted for etanercept (see section 5.4.3.2 for discussion).

Note that only study subjects randomised to etanercept at baseline were included within the analysis (i.e. subjects that crossed-over from placebo to etanercept at Week 12 were excluded) and again a mITT approach was adopted. In nr-axSpA, study 1031 was again used as the most relevant source of data. In AS, study 314-EU, which was used to estimate Week 12 response, was not of sufficient duration to inform Week 48 response. Therefore, these estimates were taken from study 312-EU, an open-label extension to study 311-EU (section 3.5.8).

5.4.3.1 Week 48 treatment response for etanercept

Coefficients of response at Week 48 regression models can be interpreted in the same way as the Week 12 estimates described in section 5.4.2.1.

nr-axSpA population

The results for the nr-axSpA population are presented in Table 112. The models of BASDAI and BASFI had R-squared values of [redacted], respectively.

[e] In the resulting models, change at Week 12 was the strongest predictor of change at Week 48 for both BASDAI and BASFI

Table 112: Change in BASDAI and BASFI from baseline at Week 48 in nr-axSpA population (n=97)

	Coefficient	SE	p-value	Lower 95% CI	Upper 95% CI
BASDAI					
Male	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27 positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Δ BASDAI week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASDAI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BASFI					
Male	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27 positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Δ BASFI week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASFI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; CRP, C-reactive protein; HLA-B27, human leukocyte antigen B27; nr-axSpA, non-radiographic axial spondyloarthritis; SE, standard error.

AS population

The results for the AS population are presented in Table 113. The models of BASDAI and BASFI had R-squared values of [redacted], respectively.

Table 113: Change in BASDAI and BASFI from baseline at Week 48 in AS population (n=43)

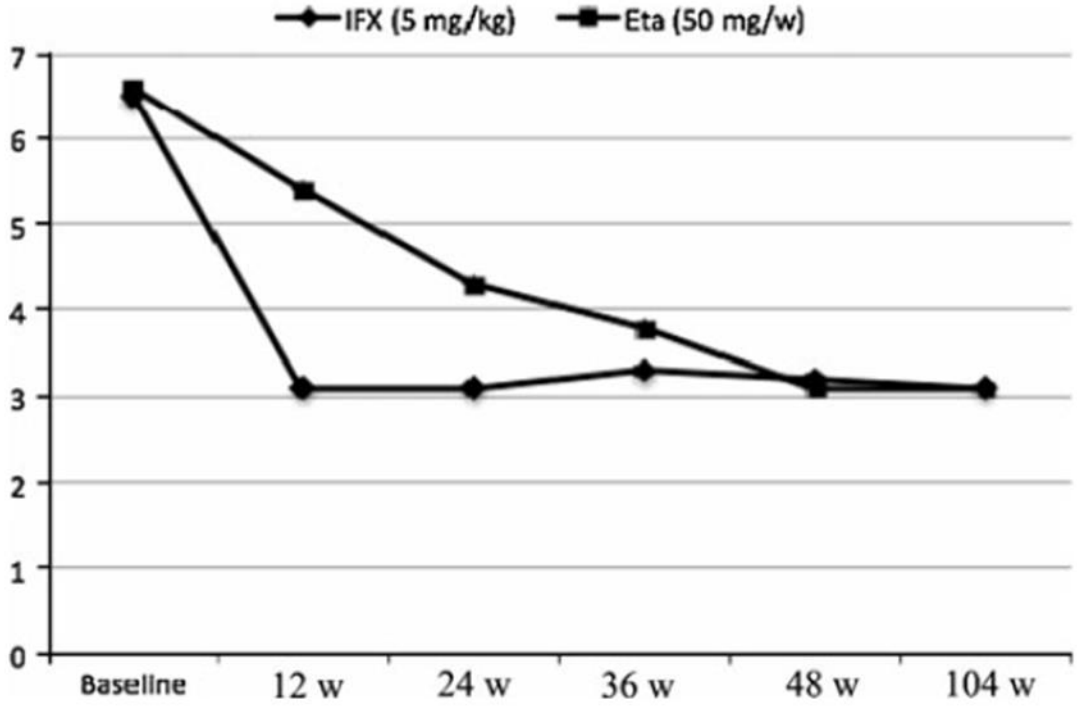
	Coefficient	SE	p-value	Lower 95% CI	Upper 95% CI
BASDAI					
Age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27 positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Δ BASDAI week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASDAI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BASFI					
Age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HLA-B27 positive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Δ BASFI week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline BASFI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AS, ankylosing spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence interval; CRP, C-reactive protein; HLA-B27, human leukocyte antigen B27; SE, standard error.

5.4.3.2 Week 48 response for other therapies

All anti-TNF agents were assumed to achieve the same Week 48 BASDAI and BASFI as etanercept. This assumption is a source of uncertainty and was confirmed at the etanercept advisory board (section 5.3.3). It is difficult to determine the validity of such assumptions as long-term comparative studies of anti-TNF agents are limited; the only long-term study identified during the clinical systematic review which compared anti-TNF agents was Giardina et al, 2010 (see section 3.5.16). The authors observed that the differences observed at Week 12 between infliximab (5 mg/kg) and etanercept (50 mg weekly) are not present at longer treatment durations for both the BASDAI and the BASFI (104) (Figure 13). This suggests that the differences between anti-TNF agents observed at Week 12 may not be observed at longer durations and may therefore be related to onset of action.

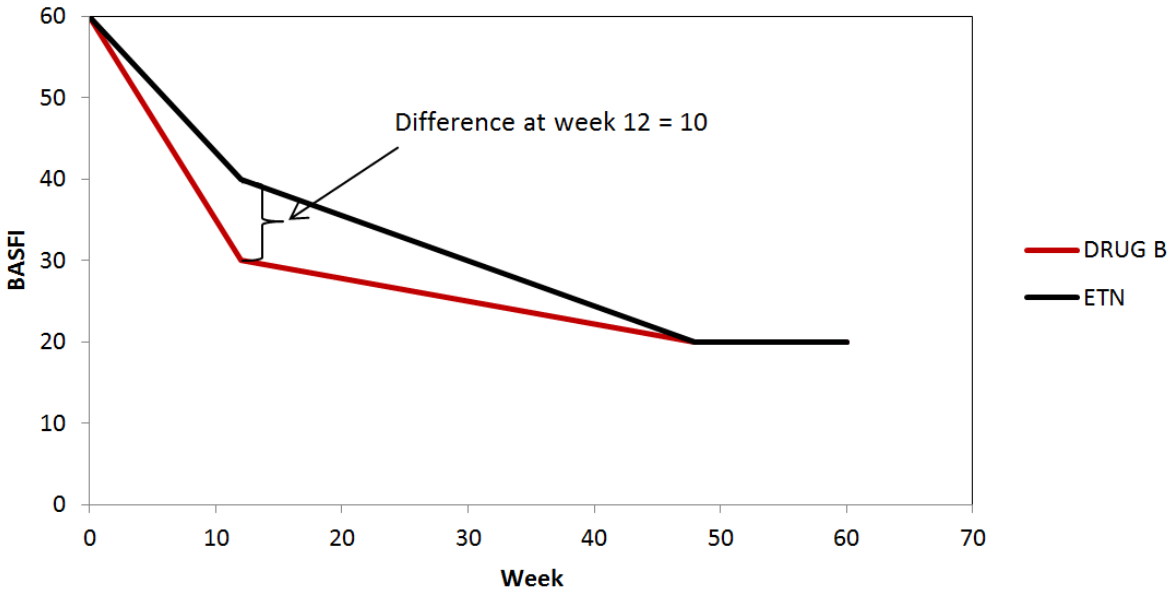
Figure 13: BASDAI levels in patients treated with infliximab or etanercept



Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; Eta, etanercept; IFX, infliximab; w, week.
Reproduced from Giardina et al, 2010 (104)

Figure 14 presents an illustrative example for BASFI of these assumptions (although the same method is used for BASDAI) assuming an illustrative mean difference of 10 at week 12.

Figure 14: Illustrative example of week 48 treatment effect assumption assuming example mean difference of 10 at week 12



Abbreviations: BASFI, Bath Ankylosing Spondylitis Functional Index; ETN, etanercept.

For patients receiving conventional care, BASDAI and BASFI are assumed to return to baseline at Week 12, and hence this assumption is not applicable. The rationale for this assumption is discussed in section 5.2.6.

5.4.3.3 Progression after Week 48

New long-term evidence relating to the effects of treatment with anti-TNF agents has emerged since the publication of TA143. Baraliakos et al, 2013 (9) have recently observed constant BASDAI and BASFI scores in 16 individuals with AS undergoing continuous treatment over 7 years with etanercept (Figure 15; see section 3.5.14). Martin-Mola et al, 2010 also demonstrate relatively constant BASDAI and BASFI scores between Years 2 and 5 in 59 etanercept-treated AS patients (122). In Scottish and Irish individuals, Dean et al, 2014 observe that BASDAI is relatively consistent over time following the introduction of anti-TNF agents (271). Therefore, although random fluctuations exist within the data, many studies now support the assumptions of constant BASDAI and BASFI for anti-TNF agents in the long-term.

The results of extension study 907-EU (section 3.5.9) also demonstrate that the positive efficacy outcomes achieved with etanercept in study 312-EU (section 3.5.8), including improvements in physical function and spinal mobility, were sustained in comparison to baseline throughout the duration of this 156 week extension study.

Figure 15: BASDAI and BASFI over 7 years for etanercept patients

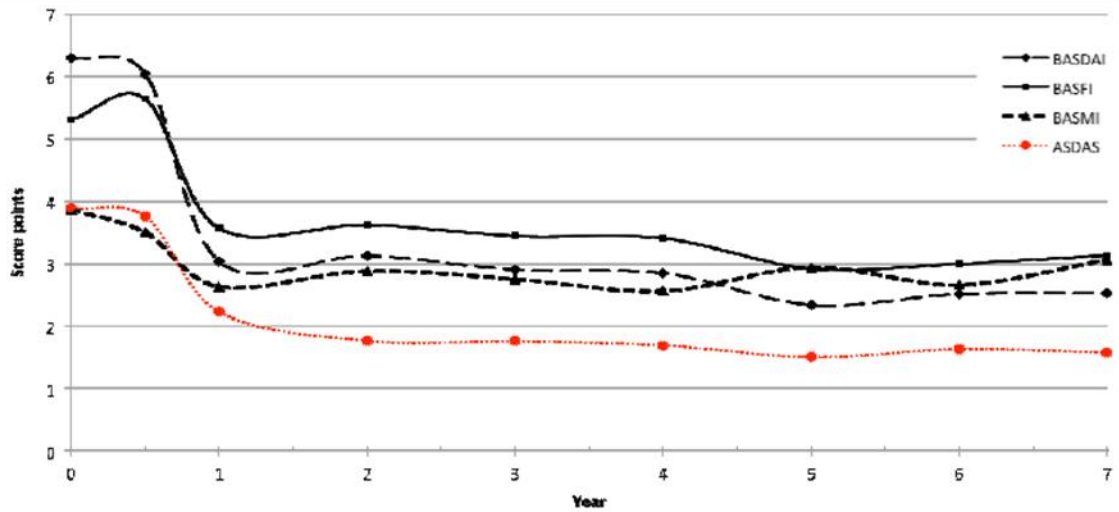


Figure 3 Course of disease related clinical parameters: ASDAS, BASDAI, BASFI, BASMI. Completer analysis. Course of disease related clinical parameters: ASDAS, Ankylosing spondylitis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index, BASFI, Bath Ankylosing Spondylitis Function Index; BASMI, Bath Ankylosing Spondylitis Metrology Index. Completer analysis of the 16 patients who reached Year 7 of the study.

Reproduced from Baraliakos et al, 2013 (9).

Table 114 contains a summary of the assumptions made in previous economic evaluations for disease progression after initial improvement. In summary, the majority of economic evaluations assumed that BASDAI and BASFI remain constant for responders on anti-TNF agents after initial response to treatment.

Table 114: Summary of progression assumptions by study (all converted to 0–100 scale)

Author, year	Progression assumption (BASDAI)		Progression assumption (BASFI)	
	anti-TNF responders	anti-TNF non-responders or conventional care	anti-TNF responders	anti-TNF non-responders or conventional care
Ara, 2007 (193)	Constant at level observed at week 24		Constant at level observed at week 24	0.7 beyond 6 months
Botteman, 2007 (194)	Constant at level observed at week 48		Constant	0.5
Kobelt, 2007 (199)	Constant		3 scenarios: - no progression - 50% reduced progression - no effect of treatment	0.7
Marcellusi, 2012 (200)	Constant at level observed at week 12		Constant at level observed at week 12	0.5
Wyeth (TA143), 2007 (4)	Constant	0.3	Constant	0.3
Abbott (TA143), 2007 (4)	Constant		Constant	0.5
Schering-Plough (TA143), 2007 (4)	Constant		Constant	0.7
McLeod, 2007 (201)	Constant		Constant	0.7
MSD (TA233), 2011 (5)	Constant		Constant on treatment for 2 years	0.7
SMC, 2005 (208)	Constant		Constant	Not reported
SMC, 2006 (207)	Constant		Constant	0.5

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index.

Assumptions have been selected based on those reported by previous evaluations:

- Constant BASFI score for anti-TNF responders at the level observed at Week 48 as assumed in all economic evaluations displayed in Table 114
- Constant increase of 0.7 in BASFI score per year for conventional care after 12 weeks (4, 5, 193, 199, 201) (linear progression of AS has been demonstrated in Brophy et al, 2002 (174)).
- Constant BASDAI score for all patients after 48 weeks (BASDAI progression is not explicitly modelled) as assumed in all economic evaluations displayed in Table 114

- These assumptions are used to model a time to minimally important clinical difference in BASFI, which was defined to be a 7 unit change (272).

5.4.4 Mortality

A standardised mortality ratio (SMR) of 1 for the nr-axSpA population and1.5 for the AS population were applied to general population life-tables. A SMR of 1.5 has been commonly used in previous economic models in AS, including the Abbott and Wyeth model submissions for NICE TA143 (4, 193, 194, 203). A SMR of 1.0 has also (at least implicitly) been used within the literature for nr-axSpA and AS (4, 199, 216, 265, 273). A SMR of 1.5 in the nr-axSpA population and 1 in the AS population were also applied in scenario analyses.

5.4.5 Treatment discontinuation

The most commonly used long-term annual discontinuation rates for anti-TNF agents in previous economic evaluations were 10% (4, 193, 208) and 15% (4, 5, 199, 201, 264). In TA233 the ERG considered that the use of a 15% discontinuation rate across all anti-TNF agents was arbitrary. Although mixed treatment comparisons may be able to inform these estimates the inevitable use of short-term data sources may ignore important differences between therapies which arise as a result of different immunogenicity profiles and other differences (see section 1.5.3). Differences in drug survival have been observed in other rheumatic disease areas. In a systematic review of European national drug registers for rheumatoid arthritis, Arora et al, 2013, show that etanercept was associated with numerically lower discontinuation rates than adalimumab and infliximab was associated with the highest rate (274).

Nell-Duxneuner et al, 2012, reviewed the use of anti-TNF agents in Austria between 2007 and 2009 (275). A total of 694 anti-TNF agent patients with AS were identified. Yearly costs were highest for adalimumab (€14,399 per patient), followed by infliximab (€11,685 per patient) and etanercept (€10,184 per patient). The study concluded that etanercept showed the lowest switching rate and longest 1- and 2-year drug survival (275). Analysis of 310 AS patients from the Czech ATTRA Registry collected from 2003 and 2008 also demonstrated that fewer AS patients discontinue on etanercept. The treatment discontinuation rates were 19, 16 and 8 per 100 patient-years for adalimumab, infliximab and etanercept respectively (164).

Glintborg et al, 2010, observed numerical differences in long-term discontinuation rates anti-TNF treatments, but these differences did not achieve statistical significance (268). This study used the DANBIO registry data to describe drug survival for AS patients receiving their first anti-TNF treatment. Patients received infliximab, adalimumab or etanercept, and hazard ratios were reported for predictors of treatment discontinuation. In the absence of more local long-term comparative evidence, time to discontinuation was estimated based on long-term etanercept data and adjustments are applied for the other anti-TNF agents. It is assumed that patients on conventional treatment will not discontinue as there are no other lines of therapy to follow. A scenario analysis is considered in which the discontinuation rates are set equal to etanercept for all anti-TNF therapies.

5.4.5.1 Treatment discontinuation for etanercept

In order to model time to treatment discontinuation for etanercept, time-to-discontinuation curves (parametric survival curves) were fitted to long-term etanercept open-label study data. This method was implemented in line with NICE decision support unit (DSU) best-

practice guidance on survival analysis (276). Survival curves were fitted for the following distributions:

- Gamma
- Gompertz
- Weibull
- Exponential
- Log normal
- Log-logistic

Survival models were fitted with and without covariates. Covariates for inclusion were identified from Glintborg et al, 2010, who reported gender, baseline VAS fatigue and baseline CRP as predictors of treatment discontinuation for anti-TNF agents in the DANBIO registry (268).

However, when these covariates were included in the survival model, measures of fit, such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were increased versus models without these covariates suggesting the covariates had poor predictive power. This is perhaps unsurprising given the low rates of discontinuation observed. As a result, models were subsequently estimated without covariates meaning that the baseline risk of treatment cessation for etanercept was the same for all individuals in the model.

In order to predict treatment cessation in the population likely to continue treatment after 12 weeks, parametric curves were fitted only to subjects who achieved a BASDAI 50 response at Week 12. Only subjects who were randomised to etanercept at baseline were retained within these survival analyses and subjects who began etanercept during open-label phases of studies were excluded. The survival analysis adopted a mITT approach.

nr-axSpA population

Figure 16 presents cumulative survival conditional on BASDAI 50 status at Week 12 for the nr-axSpA population. There is a clear separation between those subjects who achieved BASDAI 50 response at Week 12 and those who did not (log-rank test for equality; p=0.0266). This suggests that it is not reasonable to assume discontinuation rates based on the overall population in the context of the present decision problem; as patients who achieve a BASDAI 50 response would be expected to have a lower discontinuation rate than the overall population.

Figure 16: Treatment cessation in study 1031 by BASDAI 50 response at Week 12

[redacted]

Abbreviations: BASDAI 50, reduction of the Bath Ankylosing Spondylitis Disease Activity Index score to 50% of the pre-treatment value; LOCF, last observation carried forward.

The distributions that provided the best fit were exponential and log-logistic, based on the minimisation of the AIC and the BIC (Table 115). Unfortunately, Glintborg et al, 2010 (268), do not present Kaplan-Meier estimates of treatment discontinuation and therefore it was necessary to use the reported hazard ratios directly; this required the adoption of a survival model in the proportional hazards metric. Therefore, the log-logistic distribution could not be used without recourse to further assumptions. The exponential models are shown in Table 116 and Table 117. The variance-covariance matrix was used in probabilistic sensitivity analysis (PSA) where appropriate to induce correlation between survival model parameters.

Table 115: Model fit statistics for treatment discontinuation in nr-axSpA

Model	Obs	ll(null)	ll(model)	DF	AIC	BIC
Exponential	46	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	46	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Model	Obs	ll(null)	ll(model)	DF	AIC	BIC
Log-logistic	46	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log normal	46	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	46	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	46	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; DF, degrees of freedom; nr-axSpA, non-radiographic axial spondyloarthritis; ll, log-likelihood.

AS population

Collectively, studies 311-EU, 312-EU and 907-EU (described in sections 3.5.7, 3.5.8, and 3.5.9) represent the longest duration of follow-up available within the etanercept clinical trial programme and therefore these studies were used to inform treatment cessation. Figure 17 presents cumulative survival conditional on BASDAI 50 status at Week 12 for the AS population from study 312-EU. While the number of subjects available is small, the available evidence suggested that the BASDAI 50 response at Week 12 was not a predictor of long-term treatment discontinuation (log-rank test for equality; P=0.8170). It is not immediately clear why the conclusions of these analyses appear different between the nr-axSpA and AS population; this may be a feature of the condition or reflect the increased availability of new therapeutic options between the time periods in which EU-314 and study 1031 were conducted.

Figure 17: Treatment cessation in study 312-EU by BASDAI 50 response at week 12

[redacted]

[redacted]

[redacted]

Abbreviations: BASDAI 50, reduction of the Bath Ankylosing Spondylitis Disease Activity Index score to 50% of the pre-treatment value; LOCF, last observation carried forward.

Table 116: Exponential model for time to discontinuation from anti-TNF in the nr-axSpA population

	Coefficient	SE	z	P > z	Lower 95% CI	Upper 95% CI
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; nr-axSpA, non-radiographic axial spondyloarthritis; SE, standard error.

Table 117: Exponential model for time to discontinuation from anti-TNF in the AS population

	Coefficient	SE	z	P > z	Lower 95% CI	Upper 95% CI
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AS, ankylosing spondylitis; CI, confidence interval; SE, standard error.

The models of discontinuation translate into annual probabilities of discontinuation for etanercept patients who achieve a BASDAI 50 response of 5% and 11% for nr-axSpA and AS populations, respectively.

5.4.5.2 Treatment discontinuation for other therapies

While it is possible to estimate the 'baseline' time on treatment for etanercept using data from etanercept clinical trials, the same level of data for treatment discontinuation for other anti-TNFs is not publically available. In addition, there is an absence of head-to-head comparative studies between different anti-TNF agents; however there is a need to estimate the relative effects between anti-TNF agents within the model. Hence, the time on treatment for etanercept is adjusted for the other anti-TNF agents based on data from the largest available patient population over the longest period of time from the DANBIO Registry reported in Glintborg et al, 2010 (268). Based on this data, it was assumed that other anti-TNF agents have an increased risk of discontinuation compared with etanercept; a hazard ratio of 1.3 is applied for infliximab and 1.12 for adalimumab[f]. In the absence of evidence for golimumab and certolizumab, it is assumed that the relative effect is the same as for adalimumab on the basis that these have common molecular structure and belong to monoclonal antibodies. The available data from DANBIO and ATTRA Registries numerically demonstrated that discontinuation rates for etanercept patients are lower than infliximab and adalimumab (164, 268). With etanercept, there appears to be no correlation between antibody development and clinical response or adverse events (8, 277). Anti-adalimumab or anti-infliximab antibodies have been associated with lower drug levels (87, 88) and treatment

[f] The relative effects for infliximab and adalimumab versus etanercept were calculated as follows:

- Infliximab vs etanercept: 1/0.77
- Adalimumab vs etanercept: (0.86/0.77)

0.77 and 0.86 are the hazard ratios for etanercept vs infliximab and adalimumab vs infliximab respectively from Glintborg et al 2010

discontinuation due to lack of efficacy (7, 84). Sensitivity analysis was performed to explore the uncertainty around discontinuation rates in anti-TNF agents.

5.4.5.3 Determining cause of treatment discontinuation

After discontinuation of the first treatment, the second-line treatment will depend on whether the patient discontinued due to loss of efficacy (also known as inadequate response) or intolerance. According to NICE guidance (4), if the patient discontinues due to intolerance, an alternative anti-TNF agent is recommended as second-line. However, if a patient discontinues due to lack of efficacy, a second anti-TNF is not recommended and therefore the patient would resume on conventional therapy. Study 1031 (nr-axSpA) and studies EU-312 and subsequent follow-up study EU-907 (AS) were used to estimate the probability that discontinuation was a consequence of intolerance. Within the model, allocation of discontinuation due to intolerance or other causes is determined stochastically, by comparing a random number against the probability that discontinuation is caused by intolerance.

In AS, out of the 9 discontinuations across the 22 subjects originally randomised to etanercept achieving BASDAI 50 at Week 12 in studies 312-EU and (subsequently) 907-EU, adverse events were the primary cause in [redacted] of discontinuations.

In nr-axSpA, out of the 17 discontinuations across the 46 subjects originally randomised to etanercept achieving BASDAI 50 at Week 12 in study 1031, adverse events were the primary cause in [redacted] of discontinuations.

5.4.6 Rebound effect

In previous evaluations, discontinuing patients were assumed to either return to baseline BASDAI/BASFI scores or to revert to the scores of the patients in the conventional therapy/ no treatment arm. The time taken to rebound was either assumed to be immediate, or was assumed to take a certain amount of time. The assumptions for the rebound effect in each of the evaluations are summarised in Table 118. Glintborg et al, 2013, provides recent evidence for the return of BASDAI to baseline (204).

Table 118: Assumptions for rebound effects in the previous economic evaluations

Study	Assumptions for rebound effect
Ara 2007 (193), UK	BASDAI and BASFI measurements assumed to revert back to baseline values immediately on withdrawal.
Botteman 2007 (194), UK	When therapy was discontinued, both outcomes were assumed to have ‘caught up’ with conventional therapy by the next visit (i.e.≤12 weeks). In the model, for those patients who discontinued adalimumab therapy, the assumption that their BASFI score ‘catches up’ with conventional therapy is likely to be very conservative.
Fautrel 2010 (196), France	NR
Kobelt 2004 (197), UK	Patients withdrawing from treatment during the trial revert to their baseline scores and are incorporated into the no-treatment group.
Kobelt 2007 (199), UK	Patients revert to the BASDAI and BASFI scores of the no treatment arm at the same time point to account for underlying disease progression.
Kobelt 2008 (198), Spain	NR
McLeod 2007 (201), UK (Assessment Group for TA143 model)	Patients withdrawn from treatment incur an increase in functional score which returns them to the trajectory of non-responders.
Neilson 2010 (203), Germany	For patients who withdraw from treatment, it is conservatively assumed that BASDAI and BASFI measurements revert back to baseline values immediately on withdrawal.

Study	Assumptions for rebound effect
Farrell 2011 (195), (Abstract) UK	NR
Marcellusi 2012 (200), (Abstract) Italy	BASDAI/BASFI scores of non-responders and responders who discontinued returned to average pre-treatment levels after discontinuation.
Muszbek 2012 (202), (Abstract) UK	NR
NICE TA143 (4)	Wyeth model: Two distinct scenarios were modelled for patients who do not respond to treatment; patients either return to their original BASDAI/BASFI scores or they rebound by the same magnitude to their initial improvement.
	Abbott model: Assumed that once patients discontinued treatment, their BASDAI and BASFI scores rebounded to the average score of the conventional therapy (NSAIDS alone) patients.
	Schering-Plough model: NR
NICE TA233 (5)	To model the slow return to the BASDAI/BASFI score of conventional treatment after discontinuation, two 12-week tunnel states ('Just discontinued' and 'Discontinued') were incorporated with a BASFI score halfway between that for conventional treatment and TNF-α inhibitor treatment.
SMC 2005 (208)	NR
SMC 2006 (207)	NR
SMC 2011 (264)	NR
AWMSG 2013 (216)	Upon discontinuation of adalimumab, BASDAI and BASFI scores were assumed to return immediately to the baseline levels.
SMC 2013 (265)	NR

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; NR, not reported; TNF-α, tumour necrosis factor alpha; NSAID, non-steroidal anti-inflammatory drug.

For the base case model, it is assumed that patients rebound back to their baseline BASDAI and BASFI scores and that the rebound takes 6 months based on the TA233 submission to NICE (5). This assumption is varied in separate scenario analyses.

5.4.7 Adverse events

The cost and disutility associated with serious infections were included in the base case. Serious infections were selected for the model as opposed to, for example, serious adverse events (SAEs), since HRQoL consequences associated with infection in alternative populations have been well documented (278). Costs of adverse events have been included previously in economic evaluations in AS (5, 197, 199, 201, 264).

There were no serious infections in study 1031 (section 3.7), therefore the probability for the nr-axSpA population (for all treatments) is 0 in the model. Serious infections are accounted for in the model for the AS population by applying a yearly probability of 0.038 based on the rate in EU-312 (see section 3.5.8). Serious infections were assumed to persist for 28 days and confer a disutility of 0.156 during that time (278). The cost of a serious infection is outlined in section 5.4.9.3.

To account for the differences between rates of serious infections for anti-TNF agents, relative effects were applied in the model. Relative effects versus etanercept (used in the AS population only) were 0.95 for adalimumab, 3.68 for certolizumab, 0.86 for golimumab and 1.09 for infliximab. These were obtained from a network meta-analysis by Singh et al, 2011 (279).

Following a serious infection, the Summary of Product Characteristics (SPC) for all anti-TNF agents stipulates treatment discontinuation, which is not the case for other SAEs. The model considers all-cause treatment discontinuation (section 5.4.5.1) and therefore implicitly includes treatment discontinuation due to serious infections. The treatment of other AEs is unlikely to utilise a significant amount of medical resources or costs to the NHS (280).

5.4.8 Utility data

Mapping between EuroQol 5-Dimension (EQ-5D) and clinical measures such as the BASDAI and BASFI have been performed on numerous occasions previously (4, 5, 193, 200, 201, 203), and the details of the identification of such mapping relationships is described in section 5.3.2.2.

Algorithms identified from previous economic evaluations are detailed in Table 119. All algorithms were estimated using ordinary least squares (OLS).

Table 119: Previously used HRQL algorithms†

Source	Intercept	BASDAI	BASFI	Male	Age
Ara 2007, Neilson 2010 and TA143 (Wyeth) (4, 193, 203)	0.9230	-0.0402	-0.0432	-	-
Marcelluci 2012 (200)	0.9489	-0.0415	-0.0345	-	-
TA143 (Schering-plough), McLeod 2007 and TA223 (5, 201)	0.8772	0.0384	-0.0323	-0.0279	0.0017

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; †In order to estimate EQ-5D using these algorithms, the general format is $EQ5D = Intercept + BASDAI * B_{BASDAI} + BASFI * B_{BASFI} + Male * B_{Male} + Age * B_{Age}$.

No relationships in the nr-axSpA population were identified and therefore estimation of a *de novo* relationship was required. For consistency, *de novo* mapping relationships were estimated in both the nr-axSpA and AS populations (estimated separately). Based on the above identified algorithms, mapping relationships were estimated using OLS. The dependent variables in all estimated models were EQ-5D scores, derived using the UK tariff reported by Dolan (259). The following variables were considered in the analyses, as reported by TA143 (4) and Kobelt et al (199).

- Baseline BASDAI
- Baseline BASFI
- Gender
- Baseline Age

In addition, early exploratory analysis revealed the presence of non-linear relationships between EQ-5D, BASDAI and BASFI. These relationships are presented in xxxx

xXxxxxxx18 and Xxxxxxx19 for nr-axSpA and AS populations, respectively. The utility algorithm models were estimated using OLS, with standard errors clustered around each subject to account for repeated observations.

[redacted]

[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence internal; EQ-5D, EuroQol 5-Dimension.

[redacted]

[redacted]

Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CI, confidence internal; EQ-5D, EuroQol 5-Dimension.

The resulting *de novo* regression models are presented in Table 120 and Table 121 for the nr-axSpA and AS populations, respectively. Model selection was based on AIC. In the nr-axSpA population, model 4 (which included squared terms for BASDAI and BASFI and an interaction between BASDAI and BASFI) had the best fit, while in the AS population, model 3 was the preferred model. Scenario analyses considered using alternative model specifications for mapping BASDAI, BASFI and other characteristics to EQ-5D. Based on visual inspection, the estimated models were very similar between populations and there was a high degree of agreement between the estimated models and those published previously.

Figure 20 and Figure 21 present observed versus predicted EQ-5D in nr-axSpA and AS populations, respectively. Both suffered from over-prediction of EQ-5D at low observed EQ-5D and under-prediction at higher observed EQ-5D, as did all tested models. This is a commonly observed feature within mapping algorithms (281). It is believed that the use of these algorithms is conservative and likely to underestimate the benefits of anti-TNF agents; at low values of BASDAI and BASFI, predicted EQ-5D will be lower than observed values, while at high values of BASDAI and BASFI, predicted EQ-5D will be higher than observed.

Table 120: EQ-5D in nr-axSpA-study 1031 (includes all complete EQ-5D observations)

	Model 1	Model 2	Model 3	Model 4	Model 5
BASDAI					
BASFI					
BASDAI2†					
BASFI2†					
Age					
male					
BASDAI * BASFI‡					
CRP					
BASDAI VAS fatigue					
Constant					
Observations					
Adjusted R^2					
AIC§					

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
† Quadratic term, e.g. BASDAI^2; ‡ interaction terms between BASDAI and BASFI; §Akaike information Criterion; lower scores are preferred.

Table 121: EQ-5D in AS-study 314-EU. (includes all complete EQ-5D observations)

	Model 1	Model 2	Model 3	Model 4	Model 5
BASDAI					
BASFI					
BASDAI2†					
BASFI2†					
Age					
male					
BASDAI * BASFI‡					
CRP					
BASDAI VAS fatigue					
Constant					
Observations					
Adjusted R^2					
AIC§					

Standard errors in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001
† Quadratic term, e.g. BASDAI^2; ‡ interaction terms between BASDAI and BASFI; §Akaike information Criterion; lower scores are preferred.

Figure 20: Predicted vs observed EQ-5D in nr-axSpA population (model 4)

[redacted]

Abbreviations: AS, ankylosing spondylitis; CI, confidence internal; EQ-5D, EuroQol 5-Dimension.

Figure 21: Predicted vs observed EQ-5D in AS population (model 3)

Abbreviatio[redacted]

5.4.9 Costs

The following costs were included: drug costs, administration costs of intravenous anti-TNF agents, pre-treatment monitoring costs and general medical resource use (assumed to be a function of BASDAI and BASFI; see section 5.4.9.2). Monitoring costs were not explicitly included in order to avoid counting such costs twice.

5.4.9.1 Costs: drug acquisition and administration

Drug costs were taken from publicly available sources (279), including patient access schemes (PAS) for certolizumab and golimumab. In accordance with the PAS for certolizumab, patients receive their first 10 vials free in the model (282). The PAS for golimumab states that the 100 mg dose is the same price as the 50 mg dose for patients who meet the conditions of the PAS (283). This is accounted for by only considering the 50 mg price in the model.

For infliximab, for which the dose is based on the individual's weight (costs were calculated for each patient individually) and vial wastage was assumed. Weight is an important determinant of total cost for infliximab. Failure to include the distribution of weight in calculations of therapy costs by, for example, basing costs of IV therapies on an average population weight, could over- or under-estimate the total cost of infliximab.

Administration costs for IV administration of infliximab were based on a regular chemotherapy cost (DCRDN: Deliver subsequent elements of a chemotherapy cycle) from NHS reference costs 2012-13 (284) as assumed in McLeod et al, 2007 (201) , adjusted for inflation. Therapies administered subcutaneously were assumed to be self-administered following instruction. The cost of instruction was based on 1 hour of nurse time (190) as assumed in the manufacturer's submission for etanercept in the treatment of rheumatoid arthritis (TA130) (285). Cost of drug and administration did not differ between indications.

A summary of drug and administration costs with dosing assumptions is provided in Table 122. Conventional care was assumed to consist of treatment with NSAIDs. The costs of NSAIDs was based on diclofenac 150 mg daily, which was the most commonly used NSAID treatment of European (and UK) subjects within study 1031 (section 3.3.1).

Table 122: Drug and administration costs

Tx	Dosing assumptions	Unit cost†	Unit dose (mg)∞	Annual treatment cost	Administration costs	
					First administration	Subsequent administration
NSAIDs	Diclofenac 150 mg daily (279)	£1.10	84 x 50 mg	£14.34	£0.00	£0.00
ETN	50 mg OW (279)	£178.75	1 x 50 mg	£9,327	£49.00‡	£0.00
ADA	40 mg every other wk (279)	£352.14	1 x 40 mg	£9,187	£49.00‡	£0.00
CZP	400 mg 0, 2 and 4 wks then 200 mg every 2 wks (PAS 10 for free) (279)	£357.50	1 x 200 mg	£9,327§	£49.00‡	£0.00
GOL	50 mg monthly (PAS 100 mg dose same cost as 50 mg dose) (279, 280)	£762.97	1 x 50 mg	£9,156	£49.00‡	£0.00
INF	5 mg/kg wk 0, 2 and then every 6 or 8 wks (every 6 weeks assumed) (279)	£419.62	1 x 100 mg	£11,953¶	£302.00§	£302.00§

Abbreviations: ADA, adalimumab; CZP, certolizumab; ETN, etanercept; GOL, golimumab; INF, infliximab; OW, once weekly; NSAIDs, non-steroidal anti-inflammatory drug; PAS, patient-access scheme; Tx, treatment.
†BNF May 2014 (280).
∞Annual dosing is based on the precise number of weeks per year (52.18)
§ Regular chemotherapy cost (DCRDN: Deliver subsequent elements of a chemotherapy cycle) from NHS reference costs 2012-13 as assumed in McLeod et al 2007.
‡ One hour community nurse time from Curtis, 2013 (190).
§ Does not include cost savings due to PAS or additional titration doses (400 mg for first 3 injections).
¶ Based on average weight of 76 kg (study EU-314, section 3.5.1) and does not include costs of titration doses.

5.4.9.2 Costs: resource use

In order to predict resource use costs based on BASDAI and BASFI, previously used algorithms were considered. Algorithms identified in previous HTA submissions in AS are listed below:

1. TA143 Wyeth submission: £351.31 * 1.0665*BASDAI * 1.1707*BASFI (4)
2. TA143 Abbott submission: £708.45 + £750*BASDAI (4)
3. McLeod et al, 2007 and TA233 submission: £1585.30 * exp (0.1832*BASFI) (5, 201)

A more recent publication by Rafia et al, 2012 (191) provided a UK specific and more general algorithm based on BASDAI score only:

- BASDAI <40: Annual cost: £151.96
- 40 ≤ BASDAI <60: Annual cost: £311.08
- BASDAI ≥60: Annual cost: £1039.16

In the base case, data from Rafia et al, 2012 (191) is used as it provides the recent UK specific costs data available. The data presented also permits the separation of cost items (allowing the exclusion of NSAIDs costs).

These estimates appear slightly lower than other recent cost studies; Husain et al, 2014 consider the cost of AS to the UK NHS and observe, using routine data, that the NHS incurs average costs of £183 for GP, £558 for outpatient, £710 for inpatient/day unit and £8 for A&E attendance per AS patient per year (286). A 10 unit increase in BASDAI score was associated with £410 and £263 increase in costs when based on patient-reported data and based on routine data, respectively. In the absence of evidence reporting differences in cost between nr-axSpA and AS populations, it was assumed that these populations would be subject to the same costs. Kiltz et al, 2012 (54) consider the differences between nr-axSpA and AS and conclude that:

- The proportion of male patients in nr-axSpA is lower than AS
- There is a lower burden of inflammation in nr-axSpA compared with AS
- Both the nr-axSpA and AS groups did not differ in terms of health status, disease activity, and physical function

The key determinants of cost identified within the literature to date appear to be disease activity (4, 191) and physical functioning (5, 201), or combinations of the two (4); therefore it was considered reasonable to assume that costs were similar between populations.

5.4.9.3 Costs: adverse events

The cost of a serious infection was assumed to be £1,457 based on a weighted average of relevant NHS costs (275) (Table 123). Relevant HRG codes were identified based on Lekander et al, 2010 (287). HRG costs with the lowest complications and contraindications scores were used.

Table 123: Costs of serious infection (using in scenario analysis only)

Currency Code	Currency Description	Activity	National Average Unit Cost
WA03C	Septicaemia, with CC Score 0-1	44956	£1,792
DZ23G	Bronchopneumonia with CC Score 0-4	5231	£1,252
LA04M	Kidney or Urinary Tract Infections, with Interventions, with CC Score 0-2	2587	£2,289
PA16B	Major Infections with CC Score 0	7859	£1,573
DZ22J	Unspecified Acute Lower Respiratory Infection with CC Score 0-1	21109	£657
DZ21U	Chronic Obstructive Pulmonary Disease or Bronchitis, without NIV, without Intubation, with CC Score 0-3	52421	£1,453
Weighted average cost			**£1,457**

Abbreviations: CC, complications; NIV, Non-invasive ventilation.
Source: NHS reference costs schedules 2012-13 (275).

5.5 Summary of model assumptions

A summary of key model assumptions is provided in Table 124.

Table 124: Key model assumptions

Assumption	AS	nr-axSpA	Reference
Population			
Sampling method	Sampled with replacement from patient level data (section 5.4.1)		Study 314-EU and 1031 (section 3.3.1 and 3.5.1)
Week 12 response			
BASDAI and BASFI response first-line	Multivariate multiple regression from ETN data with relative effects from NMA (or MAIC) applied for other therapies		*De Novo* analysis
Adjustment factor for response to second line anti-TNF	BASFI: -30.07% BASDAI: -32.03% (not applied if failure due to intolerance)		(268)
Response criteria	BASDAI improvement from baseline of at least 50%		(4)
Conventional care	BASDAI and BASFI return to baseline at 12 weeks		(1)
Week 48 response			
BASDAI and BASFI response	Multivariate multiple regression from ETN data assumed for all anti-TNF agents All anti-TNF agents assumed to reach same level at week 48		*De Novo* analysis
Disease progression (BASFI)			
Patients on anti-TNF agents	No progression after week 48		(4, 5)
Patients on conventional care	Progression in BASFI score of 0.7 per year (estimated as time to 7 unit change in BASFI)		(4, 5, 193, 199, 201)
Minimum clinically	7 units		(272)

Assumption	AS	nr-axSpA	Reference
important difference			
AEs (SIs)			
Yearly probability of SIs	0.038	0	Study 312, Study 1031 (section 3.3.1 and 3.5.8
Unit cost of SI	£1,457 (weighted average)		(275)
Disutility of SI	0.156 for 28 days		(278)
Mortality			
SMR	1.5	1.0	(4, 193, 194, 203)
Background mortality	National Life Tables, England & Wales, 2010-12		Office of National Statistics (288)
HRQoL			
Utility algorithm	Derived from Study 1031	Derived from EU-314	*De Novo* analysis
Treatment discontinuation			
Time until discontinuation from ETN	Time-to-discontinuation estimated using parametric survival curves from ETN data		*De Novo* analysis
Time until discontinuation from other treatments	Hazard ratios applied verses ETN (ADA: 1.12, CZP: 1.12, GOL: 1.12, INF: 1.30)		(268)
Proportion of discontinuations due to intolerance	66.7% (95% CI: 29.5 -92.5%)	17.6% (95% CI: 3.8 – 43.4%)	Study 907-EU and Study 1031 (section 3.3.1 and 3.5.9)
Rebound effect between anti-TNF agents	Upon discontinuation from treatment, BASDAI and BASFI return to baseline over a 6 month period in which patients receive no treatment		(188)
Costs			
Administration costs	Therapies administered subcutaneously assumed to be self-administered; the cost of 1 hour of nurse instruction was included. Intravenous administration cost for INF was included.		(190, 285)
Doses and unit costs	Costs for all therapies estimated assuming 100% dose intensity in line with licensed doses		(279)
Monitoring costs	Not included in the base case		
Medical resource use	BASDAI < 40: Annual cost: £151.96; 40 ≤ BASDAI < 60: Annual cost: £311.08; BASDAI ≥ 60: Annual cost: £1039.16		(191)

Abbreviations: AE, adverse event; ADA, adalimumab; AS, ankylosing spondylitis ; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; HRQoL, health-related quality of life; INF, infliximab; NMA, network meta- analysis; nr-axSpA, non-radiographic axial spondyloarthritis; SI, serious infection; SMR, standardised mortality ratio.

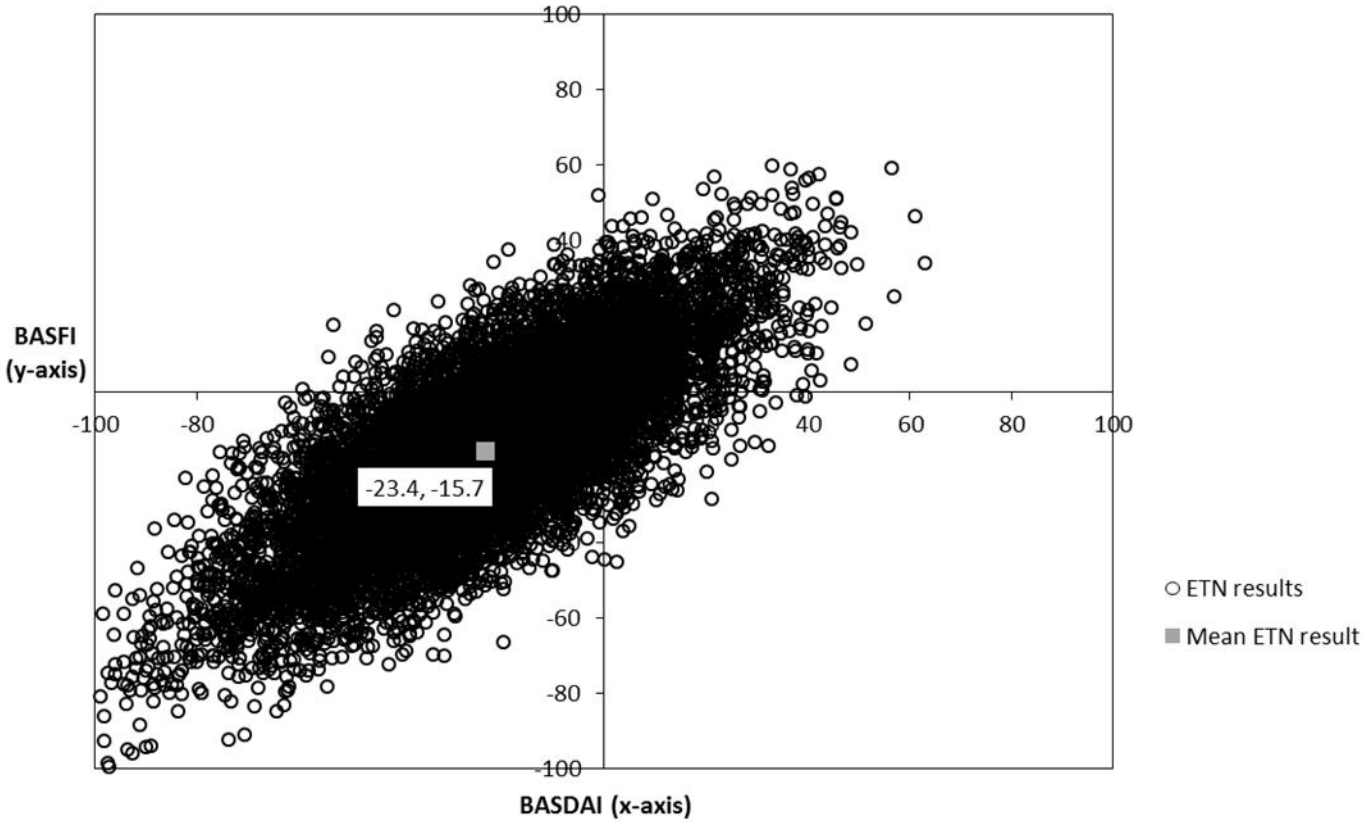
5.6 Model validation

Primary verification was performed by the developer of the economic model. A series of checks were performed including:

- Cross-check that all values within model corresponded to values within original data sources.
- Manually checked all calculations used to derive model inputs.
- Logical checks where model outcomes were checked versus expected outcomes in a range of extreme scenarios.
- Manual checks of all VBA procedures/calculations and excel calculations.
- Individual patients tracked manually through the model by stepping through the VBA code line by line.
- Logical restrictions were set by using 'watches' to ensure illogical outcomes did not occur.

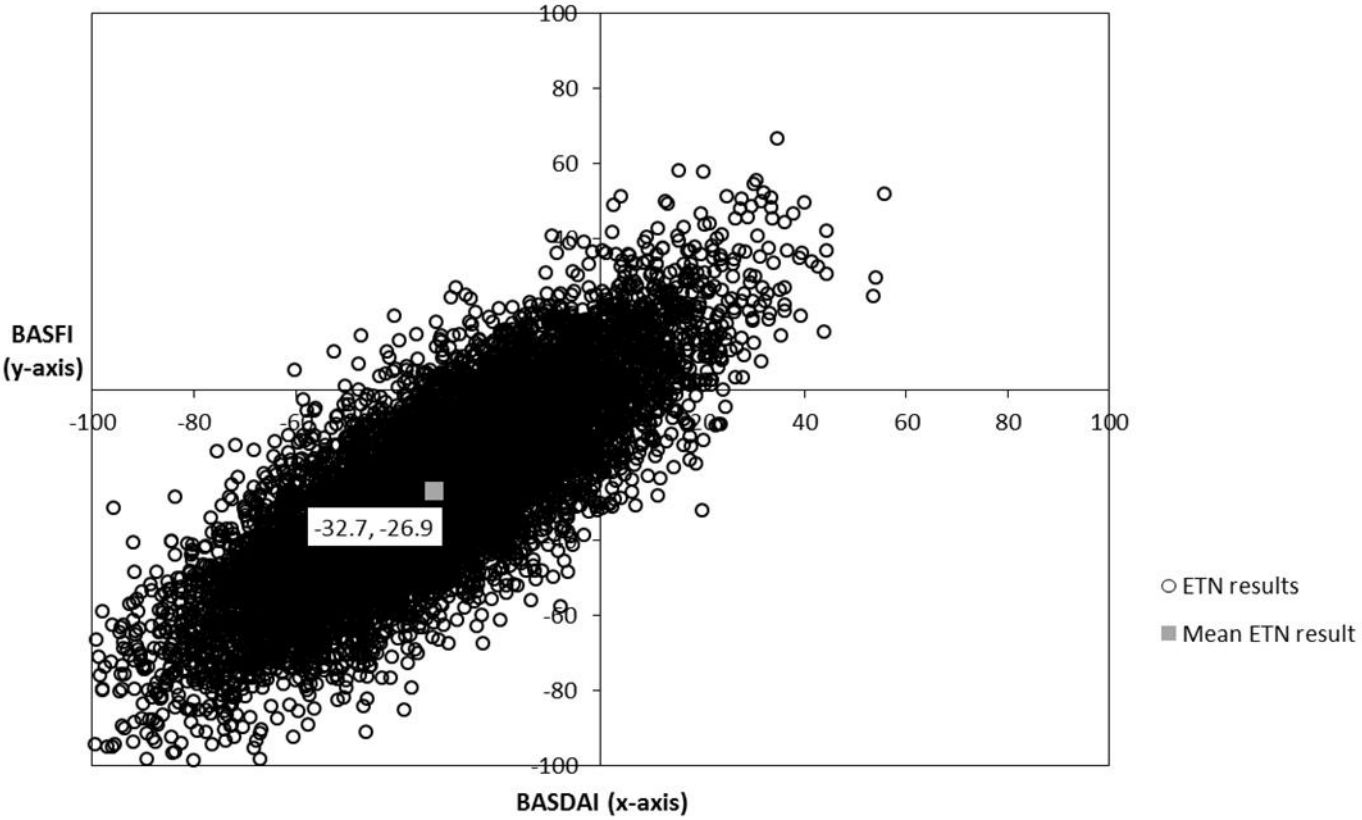
Some examples of the validation checks performed illustrated in Figure 22 – Figure 24. Figure 22 and Figure 23 show the change in BASDAI and BASFI scores from baseline at week 12 for 10,000 patients in the model. The mean changes in BASDAI and BASFI from baseline estimated by the model are plotted on the graphs for the nr-axSpA and AS populations respectively. These closely match the mean changes from study 1031 and study 314-EU (displayed in Table 125 and Table 126) which are -23.3 and -15.6, respectively in the nr-axSpA population and -32.6 and -26.8, respectively in the AS population.

Figure 22: Initial response for etanercept for 10,000 patients (change in BASDAI and BASFI at Week 12 from baseline) in the nr-axSpA population



Abbreviations: BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; ETN, etanercept; nr-axSpA, non-radiographic axial spondyloarthritis; wk, week.

Figure 23: Initial response for etanercept for 10,000 patients (change in BASDAI and BASFI at wk 12 from baseline) in the AS population



Abbreviations: AS, Ankylosing Spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; ETN, etanercept; wk, week.

Table 125: Mean baseline and week 12 BASDAI and BASFI in study 1031 for the nr-axSpA population

	Baseline mean	Baseline S.E.	Week 12 mean	Week 12 S.E.	Mean Δ
BASDAI	59.6	17.5	36.3	24.0	-23.3
BASFI	41.8	24.5	26.2	22.8	-15.6

Abbreviations: AS, Ankylosing Spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; nr-axSpA, non-radiographic axial spondyloarthritis; S.E., standard error.

Table 126: Mean baseline and week 12 BASDAI and BASFI in study 314-EU for the AS population

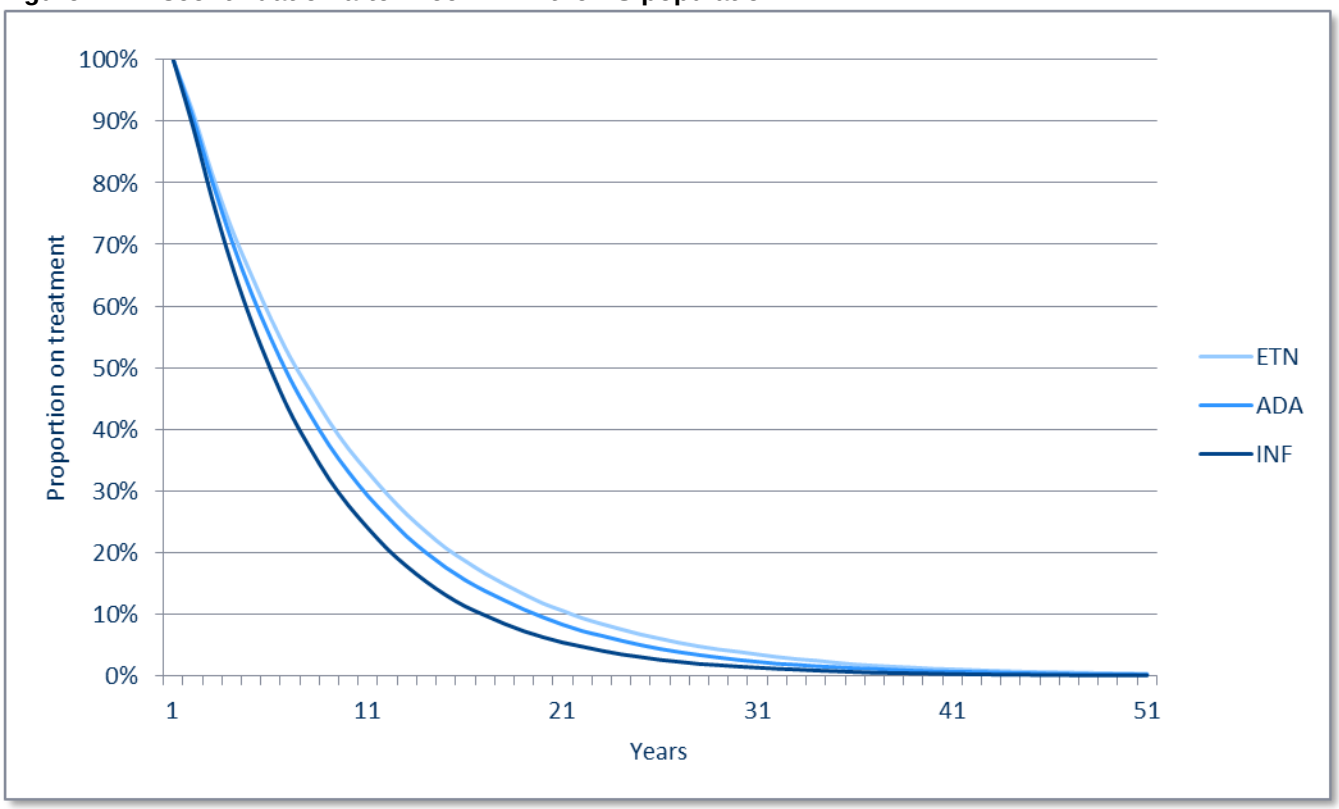
	Baseline mean	Week 12 mean	Mean Δ
BASDAI			
ETN 50mg QW	62.4	27.5	
ETN 25mg BIW	59.4	29.2	
Weighted average†	60.9	28.3	-32.6
BASFI			
ETN 50mg QW	31.8	60.6	
ETN 25mg BIW	32.9	57.7	
Weighted average†	32.3	59.2	-26.8

Abbreviations: AS, Ankylosing Spondylitis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index.

† Weighted average of ETN 50mg QW and ETN 25mg BIW in AS population

Figure 24 demonstrates discontinuation from etanercept, adalimumab and infliximab in the model based on 50,000 patients in the AS population (estimates for golimumab and certolizumab are not shown here as their relative difference versus etanercept was assumed to be the same as for adalimumab). The survival curves show the improved time on treatment with etanercept which is in line with the relative effects versus etanercept used in the model which were 1.12 for adalimumab and 1.3 for infliximab.

Figure 24: Discontinuation after week 12 in the AS population



Abbreviations: ADA, adalimumab; AS, Ankylosing Spondylitis; ETN, etanercept; INF, infliximab.

In addition, an alternative (non-DES) Markov micro-simulation model was built using the same data and assumptions as the base-case analysis. This alternative model utilised a 12 week cycle length. This was used to compare the results of the DES model with a simpler model to identify any major differences.

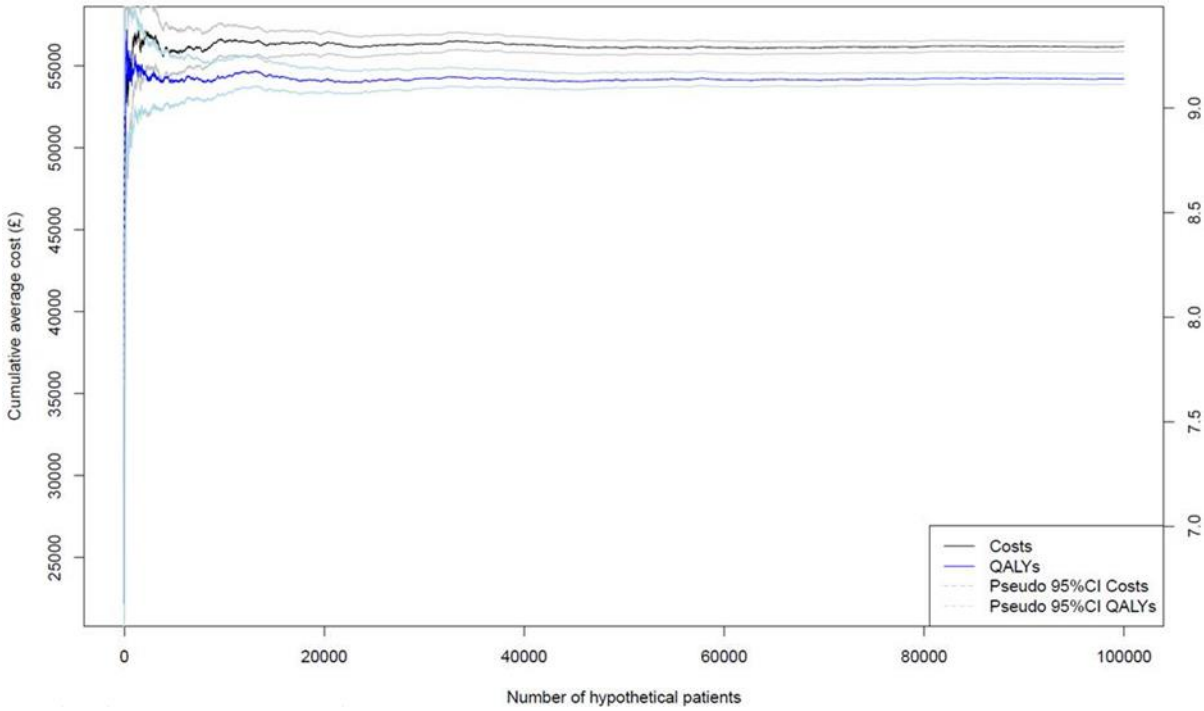
A further verification procedure was performed by an independent health economist. Tests performed included comparisons of expected outcomes versus model outcomes for specific scenarios and manual checks of all spreadsheet formulae.

Internal validity of the model was assessed by comparing model outcomes (such as change in BASDAI and BASFI, time to treatment discontinuation in first- and second-lines of therapy, age at death, and BASFI at death) with original data sources.

In order to determine the appropriate number of hypothetical individuals to pass through the economic model, plots of the cumulative mean costs and QALYs were generated and reviewed to identify the point at which changes to the quasi standard errors intervals became

negligible (Figure 25). 50,000 hypothetical individuals were conservatively chosen. To reduce the total model run time, 10,000 individuals were passed through each 'inner' loop during PSA.

Figure 25: Cumulative mean costs and QALYs for etanercept



Abbreviations: CI, confidence interval; ETN, etanercept; QALY, quality adjusted life year.

5.7 Approach to sensitivity analysis

5.7.1 Probabilistic sensitivity analysis

The robustness of the model was tested using sensitivity analyses. Scenario analyses for key variables and uncertainties were performed as well as PSA. All parameters, for which estimates of uncertainty were available, were assigned distributions and parameter estimates were drawn using Monte Carlo simulation techniques. Where available, known correlation between parameters was preserved[g]. For example, the parameters of the time on treatment survival curves were assumed to follow a multivariate normal distribution and the initial response to anti-TNF agents data were sampled from the WinBUGS CODA[h].

[g] In many instances correlation between parameters is unknown and therefore sampling of parameters is performed independently. However, in the absence of the covariance structure, these parameters are sampled independently, as is common within PSA for economic evaluations.

[h] CODA was sampled 'by row', to preserve correlation

5.7.2 Scenario analysis

A series of scenario analyses were carried out to explore key structural assumptions and specific data inputs. The following scenarios were each run in the model for 50,000 patients in both the nr-axSpA and AS populations:

- Response criteria for continuation of anti-TNF treatment, tested by the reduction in BASDAI of the pre-treatment value of:
    - 25%
    - 30%
    - 35%
    - 40%
    - 45%
- BASDAI 50 response criteria applied at Week 48 (instead of Week 12)
- Mean baseline characteristics of subjects in study AS001 in the AS subgroup receiving certolizumab
    - A weighted average of the characteristics of those receiving 200 mg or 400 mg certolizumab was used
- BSRBR baseline characteristics were used to generate a hypothetical cohort of 1,000 patients[i], maintaining the variance-covariance matrix estimated using etanercept data
- Second line adjustment for initial response to treatment at 12 weeks for BASDAI and BASFI:
    - No adjustment
    - 50% adjustment
- 5 alternative HRQoL algorithms explored:
    1. Algorithm used in Ara 2007, Neilson 2010 and TA143 (Wyeth)
    2. Algorithm used in Marcelluci 2012
    3. Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223
    4. Pfizer 2014 AS algorithm (default algorithm in AS population)
    5. Pfizer 2014 nr-AxSpA algorithm 1
    6. Pfizer 2014 nr-AxSpA algorithm 2 (default used in nr-axSpA)
- SMR of 1 in the AS population, SMR of 1.5 in the nr-axSpA population
- Patient access scheme (PAS) for certolizumab excluded

---

[i] The hypothetical cohort is generated using the method of Tannenbaum et al (214), which assumes multivariate normality to preserve correlation between variables

- BASFI progression after Week 12 while on conventional care
    - 0.4 BASFI progression per year
    - 1 BASFI progression per year
- Monitoring costs included (in addition to administration costs and resource use costs)
- Infliximab administration cost of £153.75 (inflated to 2012/13 costs from Roche submission in TA198 (280) using Curtis, 2013 (289)).
- Adverse events excluded (cost and disutility of serious infections)
- Discount rates: 1.5% for benefits, 6% for costs (based on historical NICE recommendation)
- Time taken to rebound assumed to be instant
- Alternative rebound assumption: return to trajectory of conventional care
- Response assumptions for conventional therapy:
    - No initial response at Week 12; BASDAI and BASFI scores stay at baseline
    - BASDAI and BASFI scores return gradually to baseline from Weeks 12 to 48
    - 50% of patients on CON retain their 12 week response beyond Week 48
    - 20% of patients on CON retain their 12 week response beyond Week 48
- Etanercept regression estimate for 48 week BASDAI and BASFI used for all anti-TNF agents (instead of all anti-TNF agents converging at week 48 to the etanercept absolute scores)
- AS serious infection rate used in the nr-axSpA population (scenario analysis only carried out in nr-axSpA population)
- Discontinuation assumed equal for all anti-TNFs
- Discontinuation probability (after week 12):
    - Half the default rate
    - Double the default rate
- Time horizon:
    - 10 years
    - 20 years
    - 30 years

5.8 Results

5.8.1 Primary analysis

5.8.1.1 AS population

Results for the primary analysis in the AS population are presented in Table 127 and Table 128. ICERs versus conventional care were similar for all treatment strategies with the exception of the infliximab strategy (the range of ICERs was £19,586 – £21,288 for the non-infliximab treatment and £37,741 for the infliximab strategy). The costs and QALYs were very similar for all anti-TNF treatment strategies. The ICER for etanercept versus conventional care was £20,938 in the AS population.

Table 127: Results of primary analysis in the AS population

	CON	ETN	ADA	CZP	GOL	INF
Total QALYs	7.318	9.334	9.203	9.040	9.412	9.443
Costs						
Treatment Costs	£274	£45,596	£42,588	£36,600	£48,117	£83,788
Resource Use Costs	£17,848	£14,743	£14,947	£15,243	£14,581	£14,551
Total Costs	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
Clinical outcomes						
Δ BASFI wk 12	-10.982	-26.746	-26.526	-22.518	-30.171	-30.186
Δ BASDAI wk 12	-14.129	-32.672	-32.246	-28.916	-36.191	-40.550
BASDAI 50	22%	54%	53%	47%	61%	68%
Δ BASFI wk 48	0.000	-26.885	-26.396	-24.359	-28.555	-28.674
Δ BASDAI wk 48	0.000	-31.841	-31.194	-29.298	-33.585	-34.992
Δ BASFI at death	21.557	16.440	17.106	17.385	16.196	16.251
Δ BASDAI at death	1.336	-0.661	-0.198	-0.046	-0.664	-0.342
ICER vs CON	-	**£20,938**	**£20,909**	**£19,586**	**£21,288**	**£37,741**

Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; ICER, incremental cost-effectiveness ratio; INF, infliximab; QALY, quality adjusted life year; tx, treatment; wk, week.

Table 128: Incremental results of primary analysis in the AS population

Strategy	Total Cost	Total QALYs	vs. CON		vs. next less costly		Incremental analysis†	ICER vs CON
			Inc. costs	Inc. QALYs	Inc. costs	Inc. QALYs	ICER	
CON	£18,122	7.318						
CZP	£51,843	9.040	£33,720	1.722	£33,720	1.722	£19,586	£19,586
ADA	£57,535	9.203	£39,412	1.885	£5,692	0.163	Extendedly dominated	£20,909
ETN	£60,338	9.334	£42,216	2.016	£2,804	0.131	£28,834	£20,938
GOL	£62,698	9.412	£44,576	2.094	£2,360	0.078	£30,376	£21,288
INF	£98,340	9.443	£80,217	2.125	£35,642	0.032	£1,131,181	£37,741

Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; ICER, incremental cost-effectiveness ratio; Inc., Incremental; INF, infliximab; QALY, quality adjusted life year.
† Extended dominance rules out any intervention that has an incremental cost-effectiveness ratio that is greater than that of a more effective intervention

5.8.1.2 nr-axSpA population

Results in the nr-axSpA population are presented in Table 129 and Table 130. The ICERs versus conventional care for all anti-TNF agents were very similar and ranged from £23,195 (etanercept treatment strategy) to £23,575 (certolizumab treatment strategy).

.This analysis is based on the results of the NMA and therefore the relative treatment effects were potentially biased as a result of differences in baseline characteristics between nr-axSpA clinical studies and the reclassification of nr-axSpA in the certolizumab and adalimumab studies (as discussed in section 4.2.6). Analyses adjusting for differences in study baseline characteristics were performed using data from the MAIC to attempt to illustrate the adjustment required to correct for these issues. The data used to inform these analyses are reported in section 5.4.2.3.

Table 129: Results of primary analysis in the nr-axSpA population

	CON	ETN	ADA	CZP
Total QALYs	10.221	11.903	12.030	12.497
Costs				
Treatment Costs	£314	£42,021	£45,279	£57,676
Resource Use Costs	£20,295	£17,614	£17,388	£16,606
Total Costs	£20,609	£59,635	£62,667	£74,282
Clinical outcomes				
Δ BASFI wk 12	-7.687	-14.599	-15.652	-26.319
Δ BASDAI wk 12	-16.540	-23.353	-26.730	-34.696
BASDAI50	27%	38%	44%	59%

Δ BASFI wk 48	0.000	-13.975	-14.711	-19.504
Δ BASDAI wk 48	0.000	-22.149	-24.570	-28.547
Δ BASFI at death	30.795	25.375	24.921	23.559
Δ BASDAI at death	19.381	17.696	17.735	17.333
ICER vs CON	-	**£23,195**	**£23,242**	**£23,575**

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; ICER, incremental cost-effectiveness ratio; nr-axSpA, non-radiographic axial spondyloarthritis; QALY, quality adjusted life year; tx, treatment; wk, week.

Table 130: Results of primary analysis in the nr-axSpA population

Strategy	Total Cost	Total QALYs	vs CON		vs next less costly		Incremental analysis	ICER vs CON
			Inc. costs	Inc. QALYs	Inc. costs	Inc. QALYs	ICER	
CON	£20,609	10.221	-	-	-	-	-	
ETN	£59,635	11.903	£39,026	1.683	£39,026	1.683	£23,195	£23,195
ADA	£62,667	12.030	£42,058	1.810	£3,033	0.127	£23,871	£23,242
CZP	£74,282	12.497	£53,673	2.277	£11,615	0.467	£24,864	£23,575

Abbreviations: ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; ICER, incremental cost-effectiveness ratio; Inc., Incremental; nr-axSpA, non-radiographic axial spondyloarthritis; QALY, quality adjusted life year.

5.8.2 Analyses adjusting for differences in study baseline characteristics

The results of the MAIC analysis are used to provide the results of economic analyses that take into consideration the differences in classification and baseline characteristics between the trials.

The results of the MAIC analyses were used as opposed to the STC analyses as they were considered overall to provide the most conservative estimate of the treatment benefit of etanercept versus other anti-TNFs. For reasons of consistency, the results of the same analysis method, the MAIC analysis, for both the comparisons between etanercept and adalimumab and etanercept versus certolizumab were used.

5.8.2.1 Etanercept vs certolizumab pegol in AxSpA (using MAIC data)

Results for the etanercept strategy versus the certolizumab strategy in the AxSpA population are presented in Table 131 and Table 132.

In order to compare against the primary analysis which uses the NMA, the adjusted analysis using the MAIC versus CZP was conducted on the AxSpA lave (see section 5.4.2.3). The result of the NMA analysis was that etanercept was associated with an incremental cost saving of £3,076 and an incremental QALY loss of 0.15 versus CZP. Based on the use of the MAIC data, the conclusions of this analysis are reversed; etanercept is now associated with a marginal cost of £4,499 and an incremental QALY gain of 0.093.

This illustrates the considerable effects on the outcomes of the analysis introduced by differences in the baseline characteristics and patient reclassification in the certolizumab and

adalimumab nr-axSpA/AxSpA clinical trials and the methods of evidence synthesis. It must be taken into account, as it has been demonstrated here, that clinical trial population heterogeneity has influenced the NMA results and therefore the incremental analysis performed between therapies for the primary analysis are potentially unreliable.

Table 131: Results of etanercept vs certolizumab pegol in AxSpA (using MAIC data)

	ETN	CZP
Total QALYs	10.636	10.543
Costs		
Treatment Costs	£44,229.02	£39,539
Resource Use Costs	£16,130	£16,322
Total Costs	£60,359	£55,860
Clinical outcomes		
Δ BASFI wk 12	-20.729	-22.687
Δ BASDAI wk 12	-28.073	-28.002
BASDAI50	46%	46%
Δ BASFI wk 48	-20.512	-21.553
Δ BASDAI wk 48	-27.185	-27.108
Δ BASFI at death	20.810	21.045
Δ BASDAI at death	8.743	8.912

Abbreviations: AxSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CZP, certolizumab pegol; ETN, etanercept; MAIC, Matching adjusted indirect comparison; QALY, quality adjusted life year; tx, treatment; wk, week.

Table 132: Incremental results of etanercept vs certolizumab pegol in AxSpA (using MAIC data)

Strategy	Total Cost	Total QALYs	vs next less costly	
			Inc. costs	Inc. QALYs
CZP	£55,860	10.543		
ETN	£60,359	10.636	£4,499	0.093

Abbreviations: AxSpA, axial spondyloarthritis; CZP, certolizumab pegol; ETN, etanercept; Inc., Incremental; MAIC, Matching adjusted indirect comparison; QALY, quality adjusted life year.

Because the true ratio of nr-axSpA to AS patients within the AxSpA population in the UK is a source of uncertainty, two alternative scenarios were considered in which the ratio of nr-axSpA to AS patients was varied between 30:70 and 70:30. These scenarios produced incremental costs of £4,732 and £4,266 and incremental QALY gains of 0.097 and 0.089, respectively. The conclusions of this analysis were unchanged from the original 50:50 ratio.

5.8.2.2 Etanercept vs adalimumab in nr-axSpA (using MAIC data)

Results using the MAIC in the nr-axSpA population for the adalimumab strategy versus the etanercept strategy are presented in Table 133 and Table 134. In the primary analysis, ADA had an incremental cost of £3,033 and 0.127 more QALYs in the nr-axSpA population.

However, by using data from the MAIC, etanercept becomes more costly and is associated with more QALYs. Etanercept has an incremental cost of £11,910 and an incremental QALY gain of 0.455 versus ADA; this is due to the closely linked relationship between cost and QALYs observed in this model. As in the etanercept versus CZP analysis, use of the MAIC data was associated with notable changes when compared against the primary analysis, therefore the same conclusion can be drawn that the incremental analysis performed for the primary analysis should be viewed with caution.

Table 133: Results of etanercept vs adalimumab in nr-axSpA (using MAIC data)

	ETN	ADA
Total QALYs	11.928	11.473
Costs		
Treatment Costs	£42,814	£30,089
Resource Use Costs	£17,590	£18,405
Total Costs	£60,404	£48,494
Clinical outcomes		
Δ BASFI wk 12	-14.794	-11.103
Δ BASDAI wk 12	-23.521	-17.448
BASDAI 50	39%	29%
Δ BASFI wk 48	-13.687	-10.863
Δ BASDAI wk 48	-22.127	-17.808
Δ BASFI at death	25.173	26.702
Δ BASDAI at death	17.971	18.454

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; ETN, etanercept; MAIC, Matching adjusted indirect comparison; nr-axSpA, non-radiographic axial spondyloarthritis; QALY, quality adjusted life year; tx, treatment; wk, week.

Table 134: Incremental results of etanercept vs adalimumab in nr-axSpA (using MAIC data)

Strategy	Total Cost	Total QALYs	vs next less costly	
			Inc. costs	Inc. QALYs
ADA	£48,494	11.473		
ETN	£60,404	11.928	£11,910	0.455

Abbreviations: ADA, adalimumab; ETN, etanercept; nr-axSpA, non-radiographic axial spondyloarthritis; Inc., Incremental; MAIC, Matching adjusted indirect comparison; QALY, quality adjusted life year.

5.8.3 Alternative sequential pathways

Results of the analysis including alternative sequential pathways are presented in Table 135 for the AS population and Table 136 for the nr-axSpA population.

In both populations, the alternative sequential pathways were associated with increased QALYs and increased costs when compared to the treatment strategies based on current NICE guidance. The ICERs versus CON are increased in the alternative pathways versus the conventional pathways for each therapy (with the exception of INF in the AS population).

In the case of the infliximab strategy, the alternative sequential strategy allows an increased use of the other more cost-effective anti-TNF treatments at second line, resulting in a lower ICER versus the current NICE pathway for infliximab.

The ICERs versus CON in the alternative sequential pathways range from £21,823 to £35,766 in the AS population and from £23,877 to £24,041 in the nr-axSpA population. This suggests that the use of a second anti-TNF agent after all-cause discontinuation may be cost-effective when compared against the current NICE guidance. However caution is once again advised as the evidence for second line treatment is based on a reduction in efficacy from the DANBIO registry in AS patients applied to the first-line comparative trial data. The absence of second line efficacy data in anti-TNF treatments is a limitation in this analysis.

Table 135: Results of analysis including alternative sequential pathways in the AS population[j]

Strategy	Current NICE treatment algorithms			Alternative sequential treatment algorithms		
	Total Cost	Total QALYs	ICER vs CON	Total Cost	Total QALYs	ICER vs CON
CON	£18,191	7.278		£18,191	7.278	
ETN	£61,161	9.321	£21,036	£72,910	9.695	£22,641
ADA	£57,750	9.166	£20,953	£70,565	9.603	£22,526
CZP	£52,088	9.010	£19,579	£66,166	9.477	£21,823
GOL	£62,651	9.383	£21,127	£73,328	9.731	£22,480
INF	£98,655	9.415	£37,659	£104,853	9.701	£35,766

Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; Inc., Incremental; ICER, incremental cost-effectiveness ratio; INF, infliximab; QALY, quality adjusted life year.

Table 136: Results of analysis including Alternative sequential pathways in the nr-axSpA population[j]

Strategy	Current NICE treatment algorithms			Alternative sequential treatment algorithms		
	Total Cost	Total QALYs	ICER vs CON	Total Cost	Total QALYs	ICER vs CON
CON	£20,571	10.250		£20,571	10.250	
ETN	£59,641	11.939	£23,140	£79,045	12.699	£23,877
ADA	£62,334	12.059	£23,089	£80,940	12.772	£23,938
CZP	£73,950	12.531	£23,407	£87,470	13.033	£24,041

Abbreviations: ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; ICER, incremental cost-effectiveness ratio; Inc., Incremental; nr-axSpA, non-radiographic axial spondyloarthritis; QALY, quality adjusted life year.

5.8.4 Scenario Analysis

5.8.4.1 AS population

Table 137 - Table 139 presents the results of scenario analysis of the AS population. Descriptions of scenarios are reported in section 5.7.2. The most influential assumptions

[j] The ICERS in the alternative sequential analysis differ slightly to the primary analysis due to random variation caused by the additional pathways in this scenario

were related to the response in the conventional care arm. Assuming that 20% and 50% of the patients on conventional care retain their Week 12 BASDAI and BASFI response resulted in ICERs for ETN of £24,680 and £27,626 respectively (with the exception of the infliximab strategy). Although this scenario was important in testing the effect of the CON care arm assumption on the model results, it is a highly conservative assumption and is not supported by clinical data or clinical opinion (see section 5.2.6). When BASDAI and BASFI were assumed to return to baseline gradually by Week 48 for all patients on conventional care, the ICERs remained stable.

The analysis using a cohort with baseline patient characteristics from the BSRBR, a more representative UK population, resulted in lower ICERs for all strategies (excluding infliximab) in the range of approximately £16,000. Alternative discount rates were also influential and reduced the ICERs versus conventional care to approximately £14,000 – £15,000 (excluding infliximab). For all of the other plausible assumptions the ICERs for anti-TNFs treatments versus CON are very stable.

The analyses exploring the BASDAI response criteria were of particular interest to the advisory board (see section 5.3.3). In the analysis where the BASDAI50 response criteria is evaluated at 48 weeks rather than at 12 weeks patients have more time to show an improvement with anti-TNF treatment, this is in-keeping with the clinical data that continues to show improvement in effect after 12 weeks for all anti-TNFs. The ICERs in this analysis are reduced (with the exception of CZP) compared to the base-case despite the increased costs meaning that week 48 is a far more cost effective point at which to evaluate patients on anti-TNFs with BASDAI50 than at 12 weeks. When the percentage of the pre-treatment BASDAI value required is reduced to less than 50%, the total QALYs for the anti-TNF strategies are greater. Although these scenarios also lead to increased costs, the ICERs stay below £30,000 (with the exception of infliximab) and therefore may be considered cost-effective.

Table 137: Summary of scenario analysis results in AS population

Scenario	ICER vs CON				
	ETN	ADA	CZP	GOL	INF
DEFAULT	£20,938	£20,909	£19,586	£21,288	£37,741
Response criteria 25%	£23,343	£23,019	£22,153	£23,330	£39,005
Response criteria 30%	£22,769	£22,541	£21,587	£22,860	£38,774
Response criteria 35%	£22,195	£22,039	£21,031	£22,346	£38,247
Response criteria 40%	£21,858	£21,734	£20,538	£22,097	£38,445
Response criteria 45%	£21,517	£21,255	£20,002	£21,578	£38,079
BASDAI 50 response criteria applied at week 48	£20,937	£20,651	£20,126	£20,578	£35,186
CZP study baseline characteristics	£20,484	£20,310	£19,115	£20,643	£38,555
BSRBR baseline characteristics	£16,580	£16,452	£15,524	£16,653	£29,480
No second line treatment adjustment	£20,938	£20,909	£19,586	£21,288	£37,741
50% second line treatment adjustment	£20,938	£20,909	£19,586	£21,288	£37,741
HRQoL algorithm used in Ara 2007, Neilson 2010 and TA143	£20,465	£20,429	£19,101	£20,863	£37,013

Scenario	ICER vs CON				
	ETN	ADA	CZP	GOL	INF
(Wyeth)					
Algorithm used in Marcelluci 2012	£22,323	£22,279	£20,838	£22,749	£40,350
Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223	£23,838	£23,807	£22,262	£24,304	£43,103
Pfizer 2014 AS algorithm	£20,433	£20,406	£19,106	£20,785	£36,846
Pfizer 2014 nr-AxSpA algorithm 1	£21,238	£21,180	£19,873	£21,562	£38,225
SMR 1	£20,950	£20,790	£19,423	£20,979	£37,351
Exclude CZP PAS	£21,052	£21,035	£21,663	£21,415	£37,885
0.4 BASFI progression for CON	£22,331	£22,095	£20,642	£22,422	£40,074
1 BASFI progression for CON	£20,330	£20,095	£18,802	£20,387	£36,376
Include monitoring costs	£21,083	£21,057	£19,733	£21,437	£37,948
INF admin cost of £153.75 (inflated from Roche submission in TA198)	£20,773	£20,731	£19,407	£21,096	£35,224
Exclude adverse events	£20,848	£20,819	£19,386	£21,202	£37,626
Discount rates 1.5% (benefits) and 6% (costs)	£14,991	£15,180	£14,013	£15,445	£28,415
Instant rebound	£22,018	£21,837	£20,641	£22,258	£39,403
Rebound to CON trajectory	£21,029	£21,003	£19,673	£21,384	£37,916
No CON initial response	£21,014	£20,696	£19,493	£21,056	£37,586
CON BASDAI and BASFI return to baseline by week 48	£21,127	£21,094	£19,764	£21,472	£38,055
50% of CON retain 12 week response beyond week 48	£27,626	£27,510	£25,787	£28,075	£49,532
20% of CON retain 12 week response beyond week 48	£24,680	£24,593	£23,084	£25,161	£44,407
ETN regression estimate for 48 week BASDAI and BASFI used for all TNFs	£20,904	£20,914	£20,089	£20,819	£36,761
Discontinuation assumed equal for all anti-TNFs	£21,348	£21,026	£19,817	£21,390	£38,710
Discontinuation rate halved	£17,259	£17,221	£7,128	£16,934	£31,342
Discontinuation rate doubled	£20,764	£20,430	£19,705	£20,854	£45,221
10 year time horizon	£23,385	£23,023	£21,209	£23,427	£45,162
20 year time horizon	£22,295	£21,985	£20,566	£22,310	£40,620
30 year time horizon	£21,610	£21,327	£20,096	£21,703	£39,030

Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; HRQoL, health related quality of life; ICER, incremental cost-effectiveness ratio; INF, infliximab; nr-axSpA, non-radiographic axial spondyloarthritis; PAS, patient access scheme; SMR, standardised mortality ratio.

Table 138: Scenario analysis results in AS population: Costs

Scenario	Total costs					
	CON	ETN	ADA	CZP	GOL	INF
DEFAULT	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
Response criteria 25%	£18,122	£83,856	£80,447	£75,708	£83,690	£124,765
Response criteria 30%	£18,201	£79,970	£76,929	£71,629	£80,091	£121,362
Response criteria 35%	£18,194	£75,746	£72,456	£66,253	£76,036	£115,960
Response criteria 40%	£18,140	£70,701	£67,491	£61,680	£71,758	£110,218
Response criteria 45%	£18,162	£66,254	£62,687	£56,138	£67,107	£105,040
BASDAI50 response criteria applied at week 48	£18,114	£69,028	£66,520	£65,096	£66,056	£96,785
CZP study baseline characteristics	£18,638	£62,140	£59,153	£52,869	£63,784	£104,552
BSRBR baseline characteristics	£19,456	£63,329	£60,576	£54,672	£65,117	£101,581
No second line treatment adjustment	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
50% second line treatment adjustment	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
HRQoL algorithm used in Ara 2007, Neilson 2010 and TA143 (Wyeth)	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
Algorithm used in Marcelluci 2012	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
Pfizer 2014 AS algorithm	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
Pfizer 2014 nr-AxSpA algorithm 1	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
SMR 1	£19,023	£62,147	£59,326	£53,143	£63,491	£99,786
Exclude CZP PAS	£18,122	£60,569	£57,772	£55,418	£62,964	£98,646
0.4 BASFI progression for CON	£18,476	£61,194	£57,778	£52,118	£62,760	£98,865
1 BASFI progression for CON	£17,892	£60,658	£57,773	£51,133	£61,940	£98,515
Include monitoring costs	£18,122	£60,632	£57,813	£52,095	£63,011	£98,779
INF admin cost of £153.75 (inflated from Roche submission in TA198)	£18,122	£60,007	£57,199	£51,534	£62,296	£92,990
Exclude adverse events	£18,122	£60,184	£57,390	£51,546	£62,546	£98,153
Discount rates 1.5% (benefits) and 6% (costs)	£12,866	£48,604	£46,421	£41,211	£50,812	£83,240
Instant rebound	£18,206	£61,465	£58,194	£52,602	£62,742	£98,908
Rebound to CON trajectory	£18,122	£60,338	£57,535	£51,843	£62,698	£98,340
No CON initial response	£18,184	£60,670	£57,807	£51,692	£62,601	£98,577

Scenario	Total costs					
	CON	ETN	ADA	CZP	GOL	INF
CON BASDAI and BASFI return to baseline by week 48	£17,963	£60,234	£57,425	£51,730	£62,594	£98,236
50% of CON retain 12 week response beyond week 48	£15,775	£59,122	£56,179	£50,167	£61,029	£97,231
20% of CON retain 12 week response beyond week 48	£17,183	£59,852	£56,992	£51,173	£62,031	£97,896
ETN regression estimate for 48 week BASDAI and BASFI used for all TNFs	£18,122	£60,334	£57,532	£51,868	£62,664	£98,252
Discontinuation assumed equal for all anti-TNFs	£18,190	£62,429	£61,311	£55,107	£66,083	£111,504
Discontinuation rate halved	£18,164	£22,274	£22,179	£19,701	£22,368	£26,299
Discontinuation rate doubled	£18,186	£115,196	£112,477	£99,333	£123,513	£265,178
10 year time horizon	£7,813	£38,517	£37,108	£31,881	£40,264	£72,662
20 year time horizon	£12,858	£52,726	£50,146	£44,208	£54,392	£90,250
30 year time horizon	£15,813	£58,083	£55,065	£49,091	£59,396	£95,454

Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; HRQoL, Health Related Quality of Life; ICER, incremental cost-effectiveness ratio; INF, infliximab; nr-axSpA, non-radiographic axial spondyloarthritis; PAS, patient access scheme; SMR, standardised mortality ratio.

Table 139: Scenario analysis results in AS population: QALYs

Scenario	Total QALYs					
	CON	ETN	ADA	CZP	GOL	INF
DEFAULT	7.318	9.334	9.203	9.040	9.412	9.443
Response criteria 25%	7.298	10.114	10.006	9.898	10.109	10.032
Response criteria 30%	7.310	10.023	9.915	9.785	10.017	9.970
Response criteria 35%	7.287	9.880	9.749	9.573	9.876	9.844
Response criteria 40%	7.310	9.715	9.581	9.430	9.737	9.705
Response criteria 45%	7.332	9.567	9.427	9.230	9.600	9.613
BASDAI50 response criteria applied at week 48	7.304	9.736	9.648	9.639	9.634	9.540
CZP study baseline characteristics	7.142	9.266	9.137	8.933	9.329	9.370
BSRBR baseline characteristics	4.143	6.789	6.642	6.411	6.885	6.929
No second line treatment adjustment	7.318	9.334	9.203	9.040	9.412	9.443
50% second line treatment adjustment	7.318	9.334	9.203	9.040	9.412	9.443
HRQoL algorithm used in Ara 2007, Neilson 2010 and TA143 (Wyeth)	7.218	9.281	9.147	8.983	9.354	9.385
Algorithm used in Marcelluci 2012	8.719	10.610	10.488	10.337	10.679	10.707
Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223	9.386	11.157	11.042	10.901	11.220	11.247
Pfizer 2014 AS algorithm	7.710	9.776	9.641	9.474	9.854	9.887
Pfizer 2014 nr-AxSpA algorithm 1	7.837	9.824	9.698	9.533	9.904	9.935
SMR 1	7.596	9.654	9.535	9.353	9.716	9.758
Exclude CZP PAS	7.318	9.334	9.203	9.040	9.412	9.443
0.4 BASFI progression for CON	7.699	9.612	9.478	9.329	9.674	9.705
1 BASFI progression for CON	7.008	9.111	8.992	8.776	9.168	9.224
Include monitoring costs	7.318	9.334	9.203	9.040	9.412	9.443
INF admin cost of £153.75 (inflated from Roche submission in TA198)	7.318	9.334	9.203	9.040	9.412	9.443
Exclude adverse events	7.318	9.336	9.204	9.042	9.413	9.445
Discount rates 1.5% (benefits) and 6%	10.277	12.660	12.487	12.299	12.733	12.753

Scenario	Total QALYs					
	CON	ETN	ADA	CZP	GOL	INF
(costs)						
Instant rebound	7.308	9.273	9.139	8.974	9.309	9.356
Rebound to CON trajectory	7.318	9.326	9.195	9.032	9.403	9.434
No CON initial response	7.295	9.317	9.210	9.014	9.405	9.434
CON BASDAI and BASFI return to baseline by week 48	7.357	9.358	9.228	9.066	9.436	9.467
50% of CON retain 12 week response beyond week 48	8.399	9.968	9.868	9.733	10.011	10.044
20% of CON retain 12 week response beyond week 48	7.750	9.588	9.469	9.317	9.652	9.684
ETN regression estimate for 48 week BASDAI and BASFI used for all TNFs	7.318	9.337	9.202	8.998	9.457	9.498
Discontinuation assumed equal for all anti-TNFs	7.302	9.374	9.352	9.164	9.541	9.712
Discontinuation rate halved	7.294	7.532	7.527	7.510	7.542	7.553
Discontinuation rate doubled	7.290	11.962	11.905	11.408	12.340	12.752
10 year time horizon	3.345	4.658	4.617	4.480	4.730	4.781
20 year time horizon	5.328	7.116	7.024	6.852	7.190	7.233
30 year time horizon	6.467	8.423	8.307	8.123	8.475	8.507

Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; HRQoL, health related quality of life; ICER, incremental cost-effectiveness ratio; nr-axSpA, non-radiographic axial spondyloarthritis; PAS, patient access scheme; QALY, quality adjusted life year; SMR, standardised mortality ratio.

5.8.4.2 nr-axSpA population

Table 140 - Table 142 presents the results of scenario analysis of the nr-axSpA population. Descriptions of scenarios are reported in section 5.7.2. As in the scenario analyses performed for the AS population, the most influential assumptions were related to the assumptions for the conventional care arm. Assuming that 20% and 50% of the patients on conventional care retain their Week 12 BASDAI and BASFI response resulted in ICERs for ETN of £26,125 and £28,901 respectively. As mentioned in section 5.8.4.1, this assumption is not clinically plausible. When BASDAI and BASFI were assumed to return to baseline gradually by Week 48 for all patients on conventional care, the ICERs remained stable.

Alternative discount rates were also influential; and reduced ICERs versus conventional care to less than £15,000 for all treatment strategies. All other plausible assumptions resulted in the ICERs for anti-TNFs treatments versus CON remaining stable.

As mentioned in section 5.8.4.1, the analyses related to the BASDAI response criteria were of particular interest to the advisory board (see section 5.3.3) and also produced more QALYs compared to the base case in the nr-axSpA population. As with the AS population, the ICER is reduced compared to the base case in the scenario where the BASDAI50 criteria is applied at Week 48. This again emphasises the cost effectiveness of evaluating anti-TNF usage at 48 weeks with BASDAI50 rather than at 12 weeks.

Table 140: Summary of scenario analysis results in nr-axSpA population

Scenario	ICER vs CON		
	ETN	ADA	CZP
DEFAULT	£23,195	£23,242	£23,575
Response criteria 25%	£24,571	£24,717	£25,415
Response criteria 30%	£24,251	£24,404	£24,905
Response criteria 35%	£23,774	£24,007	£24,455
Response criteria 40%	£23,567	£23,699	£24,115
Response criteria 45%	£23,239	£23,367	£23,741
BASDAI50 response criteria applied at week 48	£22,300	£21,899	£20,944
CZP study baseline characteristics	£20,148	£20,105	£20,225
No second line treatment adjustment	£23,195	£23,242	£23,575
50% second line treatment adjustment	£23,195	£23,242	£23,575
HRQoL algorithm used in Ara 2007, Neilson 2010 and TA143 (Wyeth)	£22,917	£23,048	£23,447
Algorithm used in Marcelluci 2012	£24,766	£24,897	£25,335
Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223	£26,236	£26,393	£26,839
Pfizer 2014 nr-AxSpA algorithm 1	£22,480	£22,574	£22,821
Pfizer 2014 nr-AxSpA algorithm 2	£22,156	£22,254	£22,517
SMR 1.5	£23,511	£23,262	£23,458
Exclude CZP PAS	£23,238	£23,288	£25,145
0.4 BASFI progression for CON	£24,924	£24,824	£25,075
1 BASFI progression for CON	£21,965	£21,929	£22,110
Include monitoring costs	£23,357	£23,406	£23,743
INF admin cost of £153.75 (inflated from Roche submission in TA198)	£23,195	£23,242	£23,575
Exclude adverse events	£23,195	£23,242	£23,575
Discount rates 1.5% (benefits) and 6% (costs)	£14,660	£14,936	£14,955
Instant rebound	£24,085	£24,103	£24,284
Rebound to CON trajectory	£23,294	£23,347	£23,686
No CON initial response	£23,107	£23,245	£23,296
CON BASDAI and BASFI return to baseline by week 48	£23,332	£23,378	£23,704

Scenario	ICER vs CON		
	ETN	ADA	CZP
50% of CON retain 12 week response beyond week 48	£28,901	£28,874	£29,011
20% of CON retain 12 week response beyond week 48	£26,125	£26,127	£26,323
ETN regression estimate for 48 week BASDAI and BASFI used for all TNFs	£23,128	£22,829	£21,474
AS SI rate used for nr-AxSpA	£23,272	£23,321	£23,748
Discontinuation assumed equal for all anti-TNFs	£23,330	£23,316	£23,507
Discontinuation rate halved	£19,509	£18,308	£5,209
Discontinuation rate doubled	£24,314	£24,233	£25,532
10 year time horizon	£27,506	£27,413	£27,067
20 year time horizon	£25,365	£25,238	£25,533
30 year time horizon	£24,248	£24,340	£24,420

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; HRQoL, health related quality of life; ICER, incremental cost-effectiveness ratio; nr-axSpA, non-radiographic axial spondyloarthritis; PAS, patient access scheme; SMR, standardised mortality ratio.

Table 141: Scenario analysis results in nr-axSpA population: Costs

Scenario	Total costs			
	CON	ETN	ADA	CZP
DEFAULT	£20,609	£59,635	£62,667	£74,282
Response criteria 25%	£20,623	£86,994	£87,860	£97,165
Response criteria 30%	£20,608	£81,640	£83,145	£92,767
Response criteria 35%	£20,608	£76,083	£77,940	£88,077
Response criteria 40%	£20,615	£71,220	£73,390	£83,742
Response criteria 45%	£20,575	£64,952	£67,720	£78,936
BASDAI50 response criteria applied at week 48	£20,583	£75,269	£71,584	£69,736
CZP study baseline characteristics	£20,807	£63,640	£65,775	£77,007
No second line treatment adjustment	£20,609	£59,635	£62,667	£74,282
50% second line treatment adjustment	£20,609	£59,635	£62,667	£74,282
HRQoL algorithm used in Ara 2007, Neilson 2010 and TA143 (Wyeth)	£20,609	£59,635	£62,667	£74,282
Algorithm used in Marcelluci 2012	£20,609	£59,635	£62,667	£74,282
Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223	£20,609	£59,635	£62,667	£74,282
Pfizer 2014 nr-AxSpA algorithm 1	£20,609	£59,635	£62,667	£74,282
Pfizer 2014 nr-AxSpA algorithm 2	£20,609	£59,635	£62,667	£74,282
SMR 1.5	£19,870	£58,909	£61,303	£72,394
Exclude CZP PAS	£20,609	£59,707	£62,750	£77,857
0.4 BASFI progression for CON	£20,996	£60,316	£62,738	£73,964
1 BASFI progression for CON	£20,229	£60,235	£62,322	£74,205

Scenario	Total costs			
Include monitoring costs	£20,609	£59,908	£62,964	£74,665
INF admin cost of £153.75 (inflated from Roche submission in TA198)	£20,609	£59,635	£62,667	£74,282
Exclude adverse events	£20,609	£59,635	£62,667	£74,282
Discount rates 1.5% (benefits) and 6% (costs)	£13,869	£45,178	£47,872	£56,677
Instant rebound	£20,610	£59,751	£62,818	£74,172
Rebound to CON trajectory	£20,609	£59,635	£62,667	£74,282
No CON initial response	£20,600	£59,950	£62,505	£73,773
CON BASDAI and BASFI return to baseline by week 48	£20,418	£59,485	£62,520	£74,145
50% of CON retain 12 week response beyond week 48	£17,547	£57,369	£60,412	£72,191
20% of CON retain 12 week response beyond week 48	£19,384	£58,728	£61,765	£73,446
ETN regression estimate for 48 week BASDAI and BASFI used for all TNFs	£20,609	£59,628	£62,600	£73,981
AS SI rate used for nr-AxSpA	£20,609	£59,744	£62,786	£74,611
Discontinuation assumed equal for all anti-TNFs	£20,610	£59,405	£65,432	£77,282
Discontinuation rate halved	£20,629	£23,701	£23,711	£21,643
Discontinuation rate doubled	£20,599	£99,503	£107,996	£137,197
10 year time horizon	£7,817	£31,601	£33,927	£39,998
20 year time horizon	£13,101	£46,742	£48,833	£58,509
30 year time horizon	£16,543	£53,842	£56,341	£67,100

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; HRQoL, health related quality of life; ICER, incremental cost-effectiveness ratio; nr-axSpA, non-radiographic axial spondyloarthritis; PAS, patient access scheme; SMR, standardised mortality ratio.

Table 142: Scenario analysis results in nr-axSpA population: QALYs

Scenario	Total QALYs			
	CON	ETN	ADA	CZP
DEFAULT	10.221	11.903	12.030	12.497
Response criteria 25%	10.202	12.903	12.922	13.213
Response criteria 30%	10.218	12.735	12.781	13.116
Response criteria 35%	10.228	12.562	12.617	12.987
Response criteria 40%	10.205	12.353	12.432	12.823
Response criteria 45%	10.230	12.139	12.247	12.688
BASDAI50 response criteria applied at week 48	10.219	12.671	12.547	12.565
CZP study baseline characteristics	8.231	10.357	10.468	11.010
No second line treatment adjustment	10.221	11.903	12.030	12.497
50% second line treatment adjustment	10.221	11.903	12.030	12.497
HRQoL algorithm used in Ara 2007, Neilson 2010 and TA143 (Wyeth)	9.907	11.610	11.732	12.196
Algorithm used in Marcelluci 2012	11.322	12.898	13.011	13.440

Scenario	Total QALYs			
Algorithm used in TA143 (Schering-plough), McLeod 2007 and TA223	11.906	13.394	13.500	13.906
Pfizer 2014 nr-AxSpA algorithm 1	9.835	11.571	11.698	12.187
Pfizer 2014 nr-AxSpA algorithm 2	10.367	12.129	12.257	12.751
SMR 1.5	9.920	11.581	11.701	12.159
Exclude CZP PAS	10.221	11.903	12.030	12.497
0.4 BASFI progression for CON	10.733	12.310	12.414	12.845
1 BASFI progression for CON	9.798	11.620	11.718	12.240
Include monitoring costs	10.221	11.903	12.030	12.497
INF admin cost of £153.75 (inflated from Roche submission in TA198)	10.221	11.903	12.030	12.497
Exclude adverse events	10.221	11.903	12.030	12.497
Discount rates 1.5% (benefits) and 6% (costs)	15.153	17.288	17.429	18.015
Instant rebound	10.260	11.885	12.011	12.466
Rebound to CON trajectory	10.221	11.896	12.022	12.487
No CON initial response	10.214	11.917	12.017	12.497
CON BASDAI and BASFI return to baseline by week 48	10.255	11.929	12.056	12.521
50% of CON retain 12 week response beyond week 48	11.402	12.780	12.887	13.286
20% of CON retain 12 week response beyond week 48	10.693	12.254	12.373	12.813
ETN regression estimate for 48 week BASDAI and BASFI used for all TNFs	10.221	11.908	12.060	12.706
AS SI rate used for nr-AxSpA	10.221	11.902	12.029	12.495
Discontinuation assumed equal for all anti-TNFs	10.242	11.905	12.165	12.653
Discontinuation rate halved	10.220	10.377	10.388	10.414
Discontinuation rate doubled	10.231	13.477	13.838	14.798
10 year time horizon	4.278	5.143	5.230	5.467
20 year time horizon	6.903	8.229	8.319	8.681
30 year time horizon	8.540	10.078	10.175	10.610

Abbreviations: ADA, adalimumab; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; HRQoL, health related quality of life; ICER, incremental cost-effectiveness ratio; nr-axSpA, non-radiographic axial spondyloarthritis; PAS, patient access scheme; SMR, standardised mortality ratio.

5.8.5 Probabilistic Sensitivity analysis

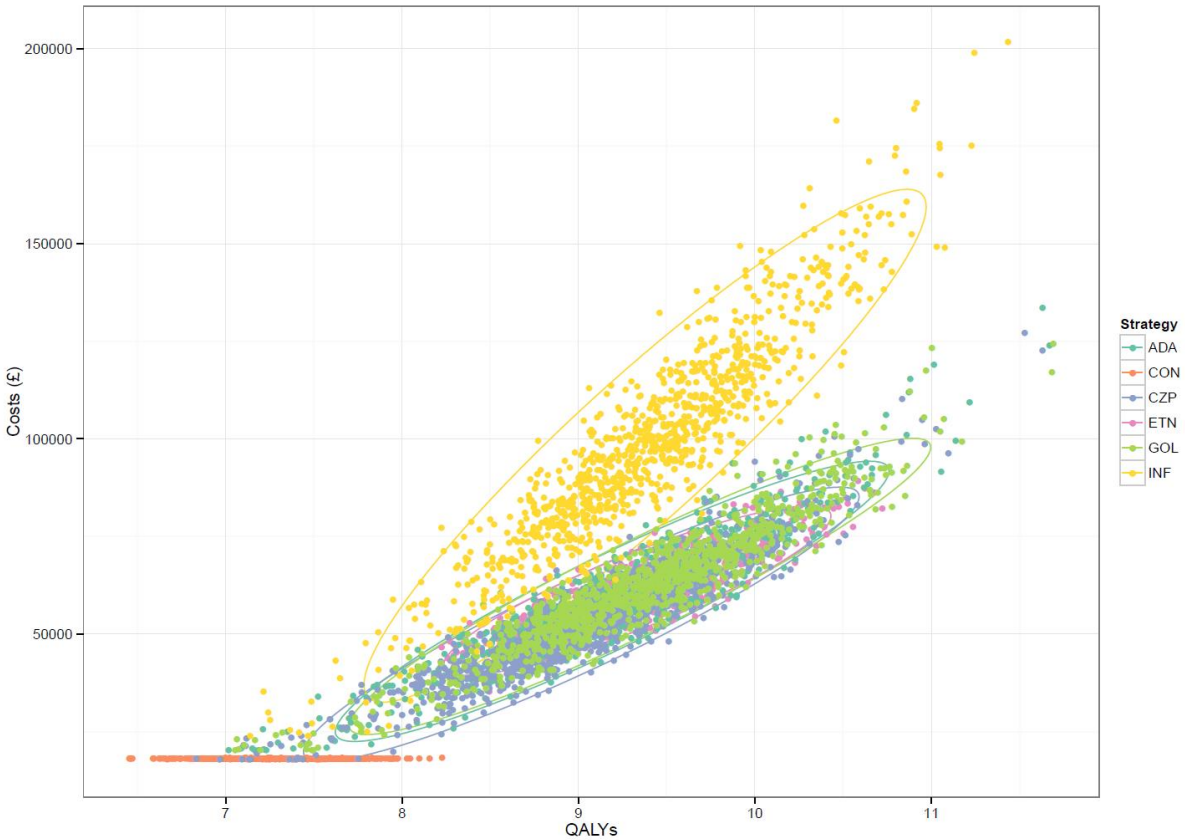
PSA was run for both nr-axSpA and AS populations. On each PSA simulation, the model samples all parameters from their respective distributions, and uses these parameters to run the model for 20,000 patients. A total of 1,000 simulations ('outer' loops) were run for each of the populations.

5.8.5.1 AS population

The results of probabilistic sensitivity analysis are presented in Figure 26 - Figure 28. Figure 26 reveals that the differences between the anti-TNF agent strategies is very small and

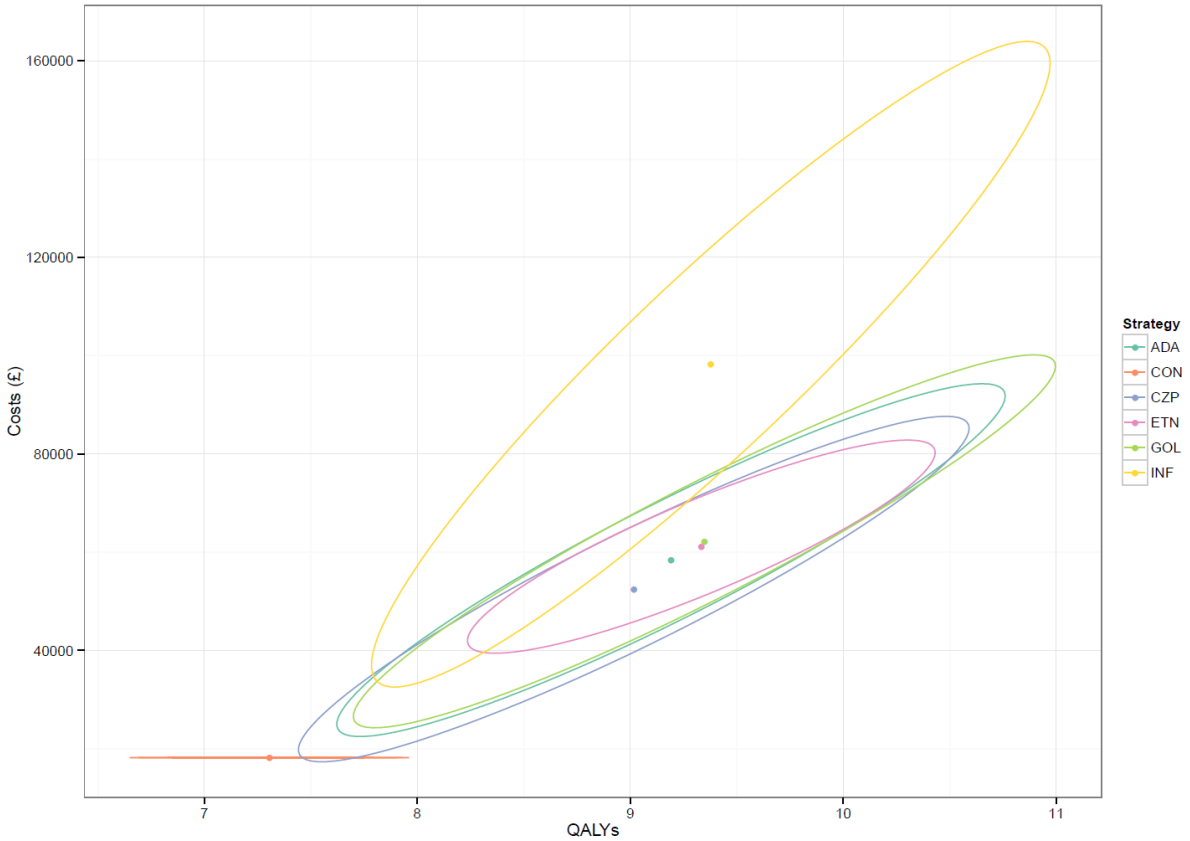
subject to considerable overlap in terms of costs and QALYs, with the exception of infliximab, which is associated with higher costs than other strategies. Compared to conventional care, the PSA results mirror those of the deterministic analysis. The striking correlation between costs and outcomes is a consequence of the imposed treatment discontinuation rule; higher efficacy is associated with more patients continuing treatment beyond Week 12, accruing a greater number of QALYs and greater cost. Figure 27 removes the noise associated with the scatter plot and shows that the mean estimates of costs and QALYs is very similar for adalimumab, certolizumab, etanercept and golimumab. Figure 28 demonstrates in all simulations, the etanercept strategy was cost-effective versus conventional care at a cost-effectiveness threshold of £30,000.

Figure 26: Cost-effectiveness plane for AS population with 95% confidence ellipses



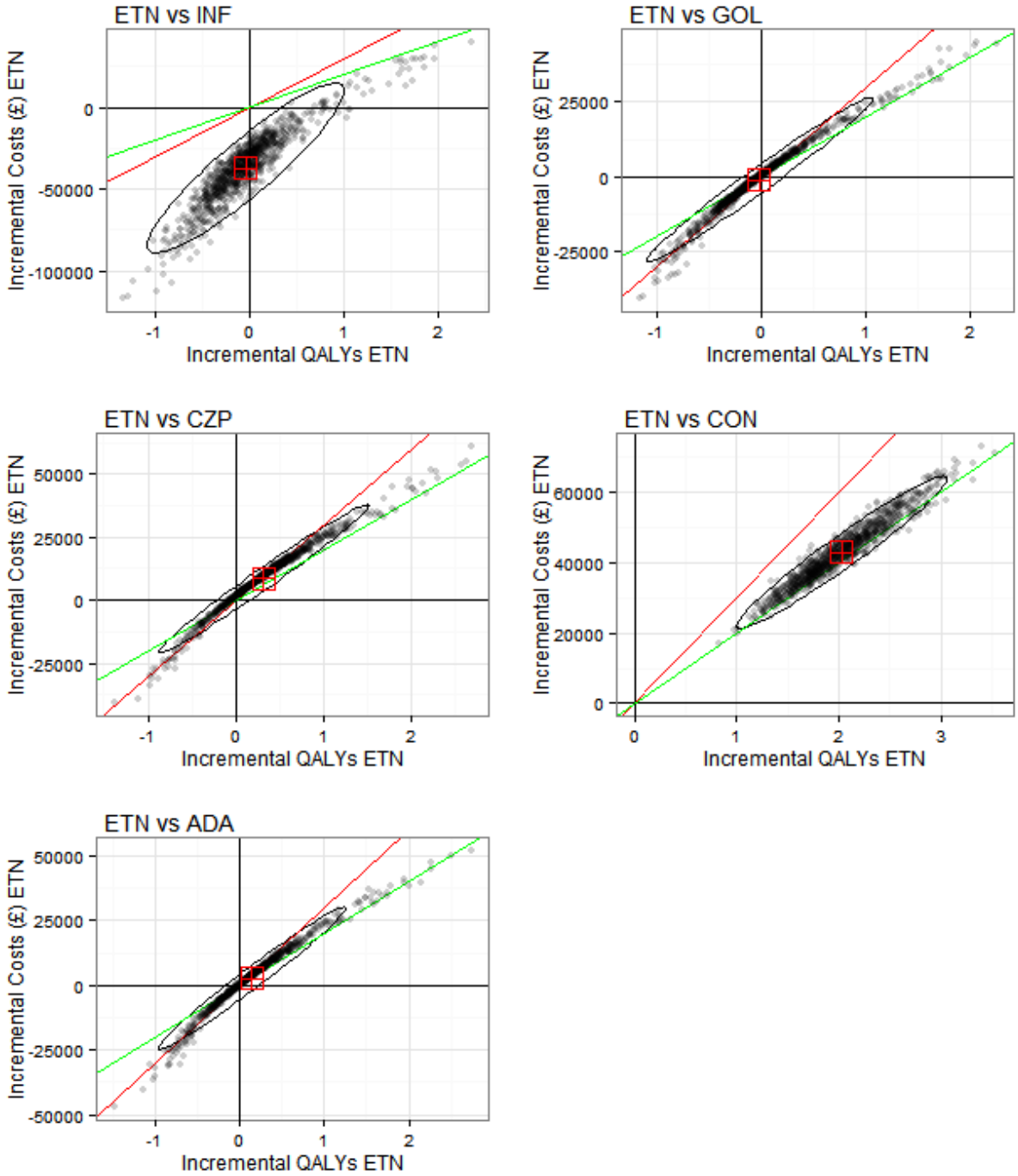
Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; QALY, quality-adjusted life-years.

Figure 27: Cost-effectiveness plane for AS population with 95% confidence ellipses and mean costs and QALYs



Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; QALY, quality-adjusted life-years.

Figure 28: Pairwise cost-effectiveness comparisons for AS with mean and 95% confidence ellipses



Abbreviations: AS, Ankylosing Spondylitis; ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; GOL, golimumab; INF, infliximab; QALY, quality-adjusted life-years.
Red line represents a cost-effectiveness threshold of £30,000
Green line represents a cost-effectiveness threshold of £20,000

5.8.5.2 nr-axSpA population

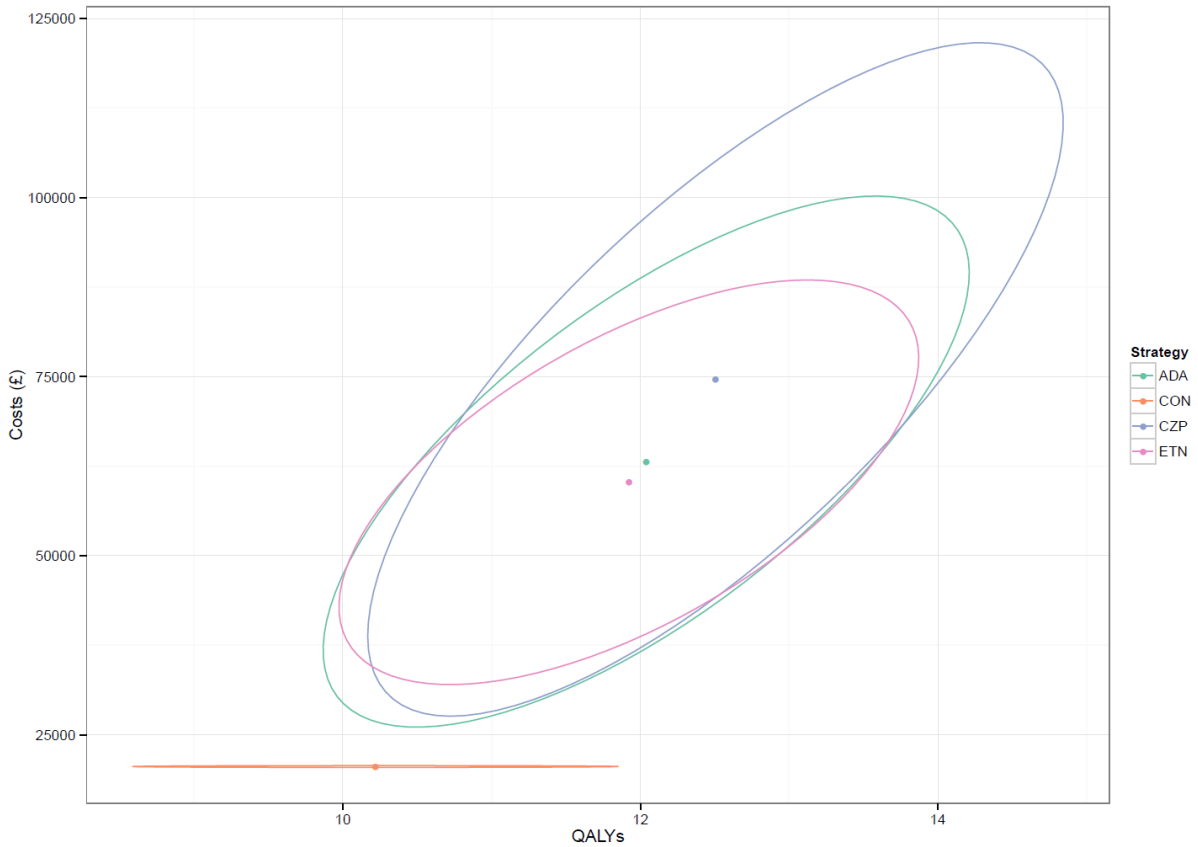
The results of probabilistic sensitivity analysis for the nr-axSpA population are presented in Figure 29 - Figure 31. Compared to conventional care, the PSA results mirror those of the deterministic analysis. Figure 29 and Figure 30 reveals that the differences between the anti-TNF agent strategies are again very small. Figure 31 demonstrates that the etanercept strategy was considered cost-effective versus conventional care at a cost-effectiveness threshold of £30,000 in 100% of simulations.

It must be noted that the results of this analysis are based on the NMA, and therefore continue to be subject to the limitations of this approach to evidence synthesis, particularly in the nr-axSpA population. As such, comparison of anti-TNF agents using these approaches should be treated with a high degree of caution.

Figure 29: Cost-effectiveness plane for primary analysis for nr-axSpA population with 95% confidence ellipses

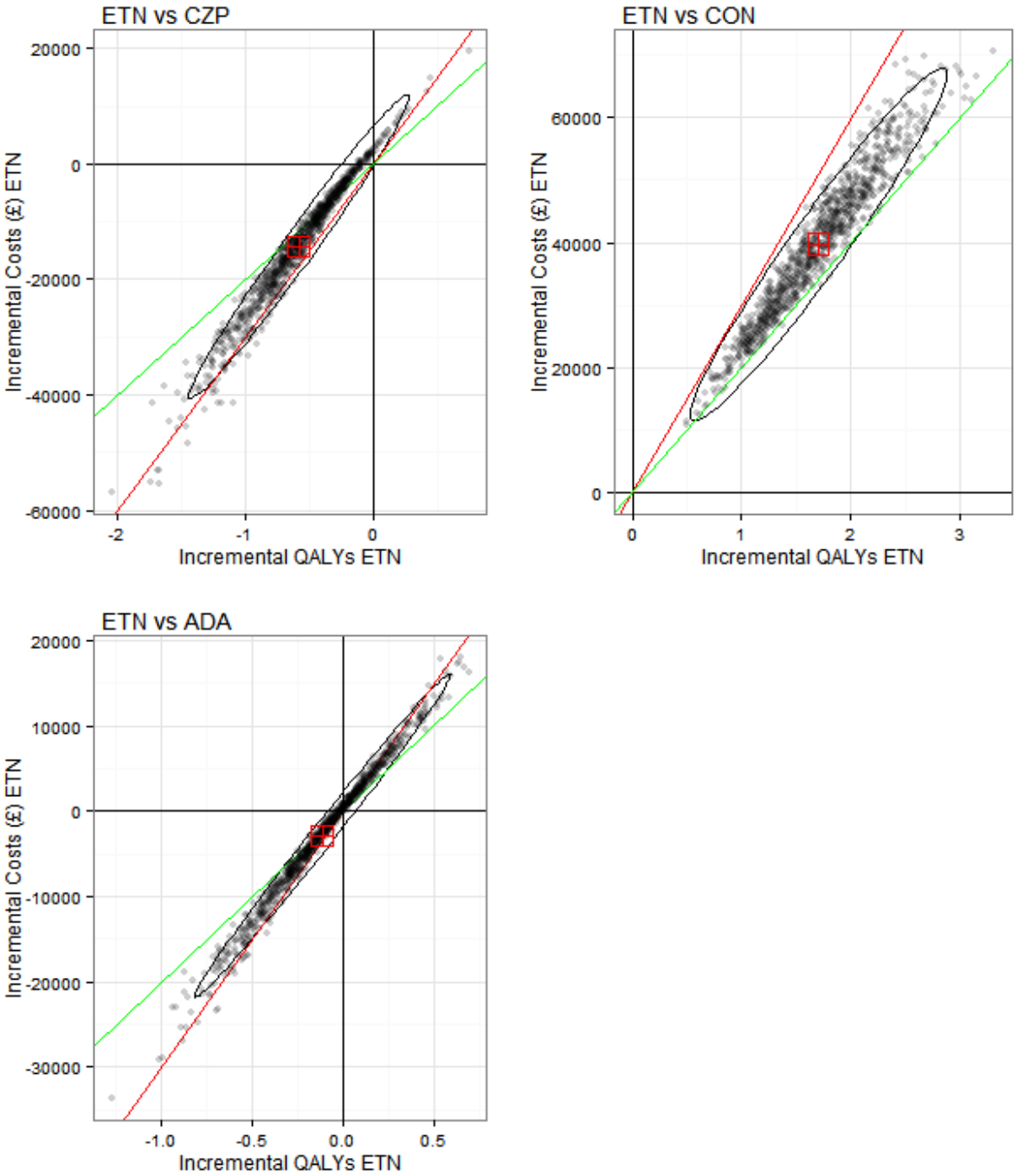
Abbreviations:; ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; QALY, quality-adjusted life-years; nr-axSpA, non-radiographic axial spondyloarthritis.

Figure 30: Cost-effectiveness plane for primary analysis for nr-axSpA population and mean costs and QALYs†



Abbreviations: ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; QALY, quality-adjusted life-year; nr-axSpA, non-radiographic axial spondyloarthritis.

Figure 31: Pairwise cost-effectiveness comparisons for nr-AxSpA with mean and 95% confidence ellipses



Abbreviations: ADA, adalimumab; CON, conventional care; CZP, certolizumab pegol; ETN, etanercept; QALY, quality-adjusted life-year; nr-axSpA, non-radiographic axial spondyloarthritis.
Red line represents a cost-effectiveness threshold of £30,000
Green line represents a cost-effectiveness threshold of £20,000

5.9 Discussion

A discrete event simulation model was developed to consider the cost-effectiveness of anti-TNF agents in nr-axSpA and AS populations. Events considered included Week 12 treatment response, Week 48 treatment response, treatment discontinuation, BASFI progression, and death. Baseline risks of events in the model were populated with *de novo* analyses using patient-level data from the etanercept clinical trial programme, with relative effects for other anti-TNF agents and conventional care taken from a systematic review and NMA. The NMA was deemed appropriate for the primary AS analysis due to the group being relatively homogeneous and more clearly defined, i.e. there were no issues with misclassification of patients (with the possible exception of certolizumab pegol).

The results of the primary analysis demonstrate that anti-TNF therapy, including etanercept, is cost-effective at £30,000 versus conventional care in both the nr-axSpA and AS populations (with the exception of the infliximab treatment strategy in AS). The ICERs for etanercept versus conventional care were £23,195 in the nr-axSpA population and £20,938 in the AS population. The other anti-TNF treatment strategies, excluding infliximab, demonstrated similar cost-effectiveness versus conventional care.

Some of the main strengths of the analysis include:

- The de novo patient-level analyses using long-term data for etanercept
- The systematic approach to evidence generation and synthesis
- The use of nr-axSpA and AS-specific population inputs for many aspects of the model
- The adoption of a model structure congruent with that used previously in the literature and within NICE technology appraisals.

The key areas of uncertainty in the analysis include:

- The failure to explicitly model the progression from nr-axSpA to AS (for which only very limited data were identified)
- The uncertainty around assumptions regarding response in patients treated with conventional care (it is assumed that patients on conventional care return immediately to baseline at week 12 but this assumption is explored in sensitivity analysis)
- The inclusion of some AS input parameters for the nr-axSpA populations:
  - Adjustment for second line treatment response
  - Relative effects for adverse events for the anti-TNFs compared to etanercept
  - Relative effects for time to discontinuation for the anti-TNFs compared to etanercept
  - Time taken to rebound to baseline on discontinuation from anti-TNF treatment
  - BASFI progression per year for patients on conventional therapy
  - Resource use cost algorithm
- As per the OMERACT guidelines, BASDAI and BASFI were used to model costs and outcomes. However, the model was not based on objective signs of inflammation as complementary to the clinical signs and the presence or absence of structural damage was not captured as this was inconsistently reported across trials

As discussed in section 4, there are differences in the baseline characteristics (including hypothesised treatment effect modifiers such as CRP, age and duration of onset) between study 1031 (etanercept), RAPID-axSpA (certolizumab pegol) and ABILITY-1 (adalimumab). This means that the results of synthesising these studies for the nr-axSpA analysis may

have methodological limitations which are likely to influence the comparison of cost-effectiveness between anti-TNF agents in the nr-axSpA population and introduce bias. In an attempt to overcome these limitations, the results of a matched adjusted indirect comparison (MAIC) were used to inform analyses which adjust for differences in baseline characteristics. The results of the MAIC were used in the economic evaluation as, overall across all outcomes and all comparisons, these results represented more conservative estimates of the treatment effect size for etanercept than those estimated by the STC. For comparison against certolizumab, this required analysis at the AxSpA level, which was possible due to the patient population included within the RAPID-axSpA study (90).

As expected, the results in terms of outcomes and costs for the analysis in the AxSpA population were intermediate between those for the nr-axSpA and AS populations. However, the results of this analysis were qualitatively different and this analysis highlights the importance of considering the differences in baseline characteristics between studies. The results were not meaningfully different between any of the anti-TNF agents when this analysis was carried out (either etanercept versus certolizumab at the AxSpA level or adalimumab at the nr-axSpA level) and it aided the conclusion that all the anti-TNFs have very similar efficacy.

The MAIC has certain methodological limitations as discussed in section 4.2 but it adds value to the analysis in allowing the populations to be brought closer together, enabling more evenly matched comparisons to be made. It demonstrates the heterogeneity in the trials for nr-axSpA and highlights the difficulties in making precise comparisons between the anti-TNF treatments.

Another analysis explored the assumption around the relative discontinuation rates for the anti-TNF agents. The scenario where all other anti-TNF agents had the same discontinuation rate as etanercept resulted in the ICERs versus conventional care remaining stable in both the nr-axSpA and AS populations.

The results of PSA showed that, versus conventional care only, etanercept is cost-effective in 100% of the simulations for both the nr-axSpA and AS populations based on a cost-effectiveness threshold of £30,000.

The PSA on the primary analysis (using the NMA) results also demonstrated that there is significant overlap between the incremental costs and QALYs for all anti-TNF agents. In both populations, 95% of the ETN ICERs simulated in the PSA are contained within the ICER ranges of other anti-TNFs. This overwhelming overlap in cost effectiveness is heavily impacted but the crossing confidence intervals in the effectiveness of the various treatments. It can, on this basis, be concluded concluded that all anti-TNF agents, (with the exception of infliximab) have similar cost-effectiveness versus conventional care. As demonstrated earlier, had the results of the MAIC been used in the PSA, the overlap would be even more complete. The smaller ICER range with ETN also demonstrates the greater volume of data available for ETN versus the other anti-TNFs.

Deterministic sensitivity analysis showed that the results of the primary analysis are very stable to the majority of changes in model assumptions. The most influential scenario analysis was related to an assumption for the conventional care arm. In the base case, it is assumed that patients on conventional care return immediately to baseline at Week 12, which was also concluded in a review by the NICE DSU of TA143 where they stated that the 0% spontaneous long term response in the control arm is a very favourable assumption (1). In scenario analysis it was assumed that 20% and 50% of the patients on conventional care

retain their Week 12 BASDAI and BASFI response which resulted in increased ICERs which remained below £30,000. However, these assumptions are very conservative and are not considered clinically plausible (see section 5.2.6). When BASDAI and BASFI were assumed to return to baseline gradually by Week 48 for all patients on conventional care (rather than immediately at Week 12), the ICERs remained stable.

One analysis that was of particular interest to the advisory board with key opinion leading rheumatologists (see section 5.3.3) was the impact of evaluating the BASDAI 50 response criteria at a later date than 12 weeks. Although the trials only have comparative data up to 12 weeks, there is continued improvement up to at least 1 year in the open label follow-up periods. Assessing the BASDAI50 criteria at 48 weeks resulted in lower ICERs for the anti-TNF agents versus conventional care despite increased cost. Assessing treatment response with BASDAI50 at 48 weeks allows patients longer to show an improvement with anti-TNF treatment, in keeping with the clinical data, and represents a more cost effective treatment paradigm than the current 12 week assessment. When the percentage of the pre-treatment BASDAI value required is reduced to less than 50%, the total QALYs for the anti-TNF strategies are greater. Although these scenarios also lead to increased costs, the ICERs stay below £30,000 (with the exception of infliximab in the AS population) and therefore may be considered cost-effective versus conventional care.

An alternative sequential use of anti-TNF agents was considered in a sensitivity analysis. The sequential use is defined as treatment strategies which permitted the use of second-line anti-TNF agents following all cause treatment failure, rather than just treatment failure due to intolerance as is currently permitted by NICE. In both the nr-axSpA and AS populations, the results of this analysis demonstrated higher costs associated with increased anti-TNF use and also more QALY benefits, leading to higher ICERs versus conventional care. Despite this, the alternative sequential treatment strategies were still considered cost-effective at a cost-effectiveness threshold of £30,000 (with the exception of infliximab) (see Appendix K).

With regards to immunogenicity, the formation of neutralising antibodies with anti-TNF therapies can lead to diminished drug levels, discontinuation of therapy due to lack of efficacy and response failure over time (7, 84, 87, 88). Potential consequences of immunogenicity are the requirement for dose increases which subsequently impacts on costs. There appears to be no correlation between antibody development and clinical response or adverse events with etanercept compared with other anti-TNF agents.

The analysis presented here excludes significant societal benefits associated with effective treatment of AS and nr-axSpA. The high burden of axial spondyloarthritis on productivity, unemployment and absenteeism has been demonstrated previously (191). The omission of such costs and health effects may underestimate the economic benefit of anti-TNF agents; if the additional societal costs were factored into the analysis, the cost-effectiveness of etanercept would be even greater. Rafia et al, 2012, (191) explored the relationship between disease severity and total costs in 612 AS patients. Direct NHS funded healthcare costs contributed to just 15% of total costs while indirect costs in the form of unemployment, absenteeism from work and reduced productivity at work accounted for 63.2%, 1.4% and 19.0% of total costs, respectively. Physical functioning and disease activity were found to be significant predictors of total costs. Boonen et al, 2013, (290) assessed the effect of etanercept after 6 and 12 months on work-related outcomes in 80 patients with active AS and in paid employment. After 6 months of etanercept treatment, significant improvements from baseline ($p<0.0001$) were observed in all Work Productivity & Activity Impairment (WPAI) components (which includes absenteeism, paid work impairment, overall work

impairment, and activity impairment)[k] and were sustained for up to 12 months. ($p<0.0001$ for all). Husain et al, 2014 (292) observed 482 UK subjects and concluded that AS results in significant work-related costs; these were predominantly due to early retirement and, in those still working, lost productivity. The total cost per patient was estimated as £11,943 annually.

In summary, the results of this cost-effectiveness analysis demonstrate that etanercept is a cost-effective treatment option for both the nr-axSpA and AS populations at a threshold of £30,000 versus conventional care. The cost-effectiveness of all anti-TNF treatments (with the exception of infliximab) are very similar and also have ICERs below £30,000. In addition, this economic analysis supports the continued recommendation of etanercept in patients with severe AS and also demonstrates that etanercept is a cost-effective treatment option for patients with nr-axSpA.

5.9.1 Relevance to England and Wales

The analysis adopts the costing perspective of the NHS and personal social services as prescribed in the NICE guides to the methods of technology appraisal (173). Whilst etanercept data is derived from multinational clinical studies, these included UK and European patients and resource use is derived from a published UK study. The analysis was defined to consider treatment strategies in line with current NICE guidance and includes measures of disease activity and functioning used within UK clinical practice.

[k] The WPAI questionnaire is a validated instrument to measure impairments in work and activities. It measures absenteeism, presenteeism, and impairments in unpaid activity because of health problems over the past seven days. It has been validated for numerous diseases such as asthma, psoriasis, irritable bowel syndrome (IBS), ankylosing spondylitis (AS) and Crohn's disease (291).

6. Wider NHS Implications

NHS Budget Impact calculations for etanercept in nr-axSpA

As AS patients already have access to etanercept as a result of previous NICE Guidance (TA143), a recommendation for etanercept in the AS population would not have any additional budget impact. Therefore, this section is only concerned with the budget impact in the nr-axSpA population.

Estimated prevalence of nr-axSpA patients

The population of England and Wales is 56,567,800 (293), of which 76% are adults (42,991,528) (294). It has been estimated that AxSpA has a prevalence in the adult population of 0.3%, and that approximately 50% of AxSpA patients are classified as nr-axSpA (64,487).

In order for a patient to receive anti-TNF treatment, these nr-axSpA patients will have to be referred to secondary care. Based on an assumption previously used for AS patients, this is assumed to happen in 60% of cases (38,692 patients) (295), of which only 38% are assumed to present with a BASDAI >4 (296). Based on previous NICE Guidance in RA, 15% of these patients (2,205) will be contraindicated to anti-TNF agents (297).

Etanercept is indicated in patients with an elevated CRP and/or a positive MRI. In the 1031 trial, this represented, [redacted] of all patients (141), meaning that the prevalent population of nr-axSpA patients eligible for treatment with anti-TNF agents is approximately [redacted] patients.

Estimated incidence of nr-axSpA patients

The annual incidence of nr-axSpA is assumed to be 0.0069% (assumed to be the same as the AS incidence from the NICE costing template in TA143 (4)). The incidence rate is multiplied by the adult population of the UK (42,991,528) to give 2,986 patients. The same factors affect the incidence as the prevalence giving an incident population of nr-axSpA patients eligible for treatment with anti-TNF agents of [redacted] per year.

Based on the combination of the incident and prevalent populations, [redacted] patients are expected to receive anti-TNF agents in the first year. Two further assumptions are applied to this figure based on the clinical data within the submission:

1. Proportion of patients reaching BASDAI 50 at 12 weeks: 44% (applied at Week 12)
2. Discontinuation rate: 10% per year (applied annually)

The annual cost of Enbrel is given in Table 122. For those patients who do not achieve a BASDAI 50 at 12 weeks and therefore discontinue etanercept, a cost of £2,152 is applied.

Table 143 shows the budget impact calculations for the next five years based on the prevalence, incidence, costs and assumptions outlined above. The calculations below include the predicted market share over the next five years:

Table 143: Budget Impact calculations for use of etanercept in nr-axSpA patients over 5 years

	Year 1	Year 2	Year 3	Year 4	Year 5
Etanercept Market share	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients					
Number of nr-axSpA patients in licenced population	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Number of nr-axSpA patients expected to receive treatment with etanercept - incident	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Number of nr-axSpA patients expected to receive treatment with etanercept - prevalent	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of nr-axSpA patients expected to receive etanercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost					
Cost of treating nr-axSpA patients with etanercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cumulative cost of treating nr-axSpA patients with etanercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

7. References

1. Hamilton L, Macgregor A, Toms A, Warmington V, Pinch E, Gaffney K. The prevalence of axial spondyloarthropathy in the UK: A cross sectional cohort study in a primary care population. BSR abstract 012. 2014.
2. Braun J, Pham T, Sieper J, Davis J, van der Linden S, Dougados M, et al. International ASAS consensus statement for the use of anti-tumour necrosis factor agents in patients with ankylosing spondylitis. Ann Rheum Dis. 2003 Sep;62(9):817-24.
3. Keat A, Barkham N, Bhalla A, Gaffney K, Marzo-Ortega H, Paul S, et al. BSR guidelines for prescribing TNF-alpha blockers in adults with ankylosing spondylitis. Report of a working party of the British Society for Rheumatology. Rheumatology (Oxford). 2005 Jul;44(7):939-47.
4. National Institute for Health and Clinical Excellence. Adalimumab, etanercept and infliximab for ankylosing spondylitis. October 2010. NICE technology appraisal guidance 143. Available from: http://www.nice.org.uk/nicemedia/pdf/TA143Guidance.pdf (Accessed May 2014).
5. National Institute for Health and Clinical Excellence. Golimumab for the treatment of ankylosing spondylitis. NICE technology appraisal guidance 233. August 2011. Available from: http://www.nice.org.uk/nicemedia/live/13560/56001/56001.pdf (Accessed 2 June 2014)
6. Haroon N, Inman RD, Learch TJ, Weisman MH, Lee M, Rahbar MH, et al. The impact of tumor necrosis factor alpha inhibitors on radiographic progression in ankylosing spondylitis. Arthritis Rheum. 2013 Oct;65(10):2645-54.
7. Arends S, Lebbink HR, Spoorenberg A, Bungener LB, Roozendaal C, van der Veer E, et al. The formation of autoantibodies and antibodies to TNF-alpha blocking agents in relation to clinical response in patients with ankylosing spondylitis. Clin Exp Rheumatol. 2010 Sep-Oct;28(5):661-8.
8. European Medicines Agency (EMA). Enbrel. Summary of product characteristics. Available from : http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000262/WC50 (Accessed 10 April 2014).
9. Baraliakos X, Haibel H, Fritz C, Listing J, Heldmann F, Braun J, et al. Long-term outcome of patients with active ankylosing spondylitis with etanercept-sustained efficacy and safety after seven years. Arthritis Res Ther. 2013;15(3):R67.
10. Davis JC, Jr., Van Der Heijde D, Braun J, Dougados M, Cush J, Clegg DO, et al. Recombinant human tumor necrosis factor receptor (etanercept) for treating ankylosing spondylitis: a randomized, controlled trial. Arthritis Rheum. 2003 Nov;48(11):3230-6.
11. Calin A, Dijkmans BA, Emery P, Hakala M, Kalden J, Leirisalo-Repo M, et al. Outcomes of a multicentre randomised clinical trial of etanercept to treat ankylosing spondylitis. Ann Rheum Dis. 2004 Dec;63(12):1594-600.
12. van der Heijde D, Da Silva JC, Dougados M, Geher P, van der Horst-Bruinsma I, Juanola X, et al. Etanercept 50 mg once weekly is as effective as 25 mg twice weekly in patients with ankylosing spondylitis. Ann Rheum Dis. 2006 Dec;65(12):1572-7.
13. Dougados M, Braun J, Szanto S, Combe B, Elbaz M, Geher P, et al. Efficacy of etanercept on rheumatic signs and pulmonary function tests in advanced ankylosing spondylitis: results of a randomised double-blind placebo-controlled study (SPINE). Ann Rheum Dis. 2011 May;70(5):799-804.
14. Braun J, van der Horst-Bruinsma IE, Huang F, Burgos-Vargas R, Vlahos B, Koenig AS, et al. Clinical efficacy and safety of etanercept versus sulfasalazine in patients

with ankylosing spondylitis: a randomized, double-blind trial. Arthritis Rheum. 2011 Jun;63(6):1543-51.

15. Song IH, Hermann K, Haibel H, Althoff CE, Listing J, Burmester G, et al. Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): a 48-week randomised controlled trial. Ann Rheum Dis. 2011 Apr;70(4):590-6.
16. Song IH, Weiss A, Hermann KG, Haibel H, Althoff CE, Poddubnyy D, et al. Similar response rates in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis after 1 year of treatment with etanercept: results from the ESTHER trial. Ann Rheum Dis. 2013 Jun;72(6):823-5.
17. Gottlieb AB, Gordon K, Giannini EH, Mease P, Li J, Chon Y, et al. Clinical trial safety and mortality analyses in patients receiving etanercept across approved indications. J Drugs Dermatol. 2011 Mar;10(3):289-300.
18. Sieper J, Koenig A, Baumgartner S, Wishneski C, Foehl J, Vlahos B, et al. Analysis of uveitis rates across all etanercept ankylosing spondylitis clinical trials. Ann Rheum Dis. 2010 Jan;69(1):226-9.
19. Braun J, Sieper J. Ankylosing spondylitis. Lancet. 2007 Apr 21;369(9570):1379-90.
20. van der Weijden MA, Claushuis TA, Nazari T, Lems WF, Dijkmans BA, van der Horst-Bruinsma IE. High prevalence of low bone mineral density in patients within 10 years of onset of ankylosing spondylitis: a systematic review. Clin Rheumatol. 2012 Nov;31(11):1529-35.
21. Zochling J, Braun J. Mortality in ankylosing spondylitis. Clin Exp Rheumatol. 2008 Sep-Oct;26(5 Suppl 51):S80-4.
22. Boonen A, Patel V, Traina S, Chiou CF, Maetzel A, Tsuji W. Rapid and sustained improvement in health-related quality of life and utility for 72 weeks in patients with ankylosing spondylitis receiving etanercept. J Rheumatol. 2008 Apr;35(4):662-7.
23. Dagfinrud H, Mengshoel AM, Hagen KB, Loge JH, Kvien TK. Health status of patients with ankylosing spondylitis: a comparison with the general population. Ann Rheum Dis. 2004 Dec;63(12):1605-10.
24. Barkham N, Kong KO, Tennant A, Fraser A, Hensor E, Keenan AM, et al. The unmet need for anti-tumour necrosis factor (anti-TNF) therapy in ankylosing spondylitis. Rheumatology (Oxford). 2005 Oct;44(10):1277-81.
25. Van der Heijde D, Braun J, Rudwaleit M, Purcaru O, Kavanaugh A. Improvements in work and household productivity after 24 weeks of certolizumab pegol in treatment of axial spondyloarthritis patients, including patients with ankylosing spondylitis: Results of rapid-axspa study. Paper presented at: European League Against Rheumatism (EULAR) 2013; Madrid, Spain.
26. Healey EL, Haywood KL, Jordan KP, Garratt A, Packham JC. Impact of ankylosing spondylitis on work in patients across the UK. Scand J Rheumatol. 2011 Jan;40(1):34-40.
27. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis. 2009 Jun;68(6):777-83.
28. Feldtkeller E, Khan MA, van der Heijde D, van der Linden S, Braun J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. Rheumatol Int. 2003 Mar;23(2):61-6.
29. Smolen JS, Braun J, Dougados M, Emery P, Fitzgerald O, Helliwell P, et al. Treating spondyloarthritis, including ankylosing spondylitis and psoriatic arthritis, to target: recommendations of an international task force. Ann Rheum Dis. 2014 Jan;73(1):6-16.

30. Braun J, van den Berg R, Baraliakos X, Boehm H, Burgos-Vargas R, Collantes-Estevez E, et al. 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis. 2011 Jun;70(6):896-904.
31. Molto A, Paternotte S, Claudepierre P, Breban M, Dougados M. Effectiveness of Tumor Necrosis Factor alpha Blockers in Early Axial Spondyloarthritis: Data From the DESIR Cohort. Arthritis Rheumatol. 2014 Jul;66(7):1734-44.
32. Sieper J, Van der Heijde D, Elewaut D, Pangan A, Anderson JK. Disease Burden Is Comparable in Patients with Non-Radiographic Axial Spondyloarthritis and Ankylosing Spondylitis. Poster.
33. AMGEN. Enbrel® (etanercept). About Enbrel. Available from: http://www.enbrel.com/RheumPro/anti-tumor-necrosis-factor.jspx (Accessed 9 June 2014).
34. Zochling J, Brandt J, Braun J. The current concept of spondyloarthritis with special emphasis on undifferentiated spondyloarthritis. Rheumatology (Oxford). 2005 Dec;44(12):1483-91.
35. van der Horst-Bruinsma IE, Nurmohamed MT. Management and evaluation of extra-articular manifestations in spondyloarthritis. Ther Adv Musculoskelet Dis. 2012 Dec;4(6):413-22.
36. Huhtinen M, Karma A. HLA-B27 typing in the categorisation of uveitis in a HLA-B27 rich population. Br J Ophthalmol. 2000 Apr;84(4):413-6.
37. Forsblad-d'Elia H, Wallberg H, Klingberg E, Carlsten H, Bergfeldt L. Cardiac conduction system abnormalities in ankylosing spondylitis: a cross-sectional study. BMC Musculoskelet Disord. 2013;14:237.
38. Kazmierczak J, Peregud-Pogorzelska M, Biernawska J, Przepiera-Bedzak H, Goracy J, Brzosko I, et al. Cardiac arrhythmias and conduction disturbances in patients with ankylosing spondylitis. Angiology. 2007 Dec-2008 Jan;58(6):751-6.
39. Dincer U, Cakar E, Kiralp MZ, Bozkanat E, Kilac H, Dursun H. The pulmonary involvement in rheumatic diseases: pulmonary effects of ankylosing spondylitis and its impact on functionality and quality of life. Tohoku J Exp Med. 2007 Aug;212(4):423-30.
40. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum. 1984 Apr;27(4):361-8.
41. Baraliakos X, Davis J, Tsuji W, Braun J. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis before and after therapy with the tumor necrosis factor alpha receptor fusion protein etanercept. Arthritis Rheum. 2005 April;52(4):1216-23.
42. Bennett AN, McGonagle D, O'Connor P, Hensor EM, Sivera F, Coates LC, et al. Severity of baseline magnetic resonance imaging-evident sacroiliitis and HLA-B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. Arthritis Rheum. 2008 Nov;58(11):3413-8.
43. Poddubnyy D. Axial spondyloarthritis: is there a treatment of choice? Ther Adv Musculoskelet Dis. 2013 Feb;5(1):45-54.
44. Rudwaleit M, Listing J, Brandt J, Braun J, Sieper J. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor alpha blockers in ankylosing spondylitis. Ann Rheum Dis. 2004 Jun;63(6):665-70.
45. Electronic Medicines Compendium. Arcoxia 30mg, 60 mg, 90 mg & 120 mg Film-coated Tablets: SPC. Available from: https://www.medicines.org.uk/emc/medicine/8734#CONTRAINDICATIONS (Accessed August 2014).

46. Dean LE, Jones GT, MacDonald AG, Downham C, Sturrock RD, Macfarlane GJ. Global prevalence of ankylosing spondylitis. Rheumatology (Oxford). 2014 Apr;53(4):650-7.
47. Strand V, Rao SA, Shillington AC, Cifaldi MA, McGuire M, Ruderman EM. Prevalence of axial spondyloarthritis in United States rheumatology practices: Assessment of SpondyloArthritis International Society criteria versus rheumatology expert clinical diagnosis. Arthritis Care Res (Hoboken). 2013 Aug;65(8):1299-306.
48. Poddubnyy D, Brandt H, Vahldiek J, Spiller I, Song IH, Rudwaleit M, et al. The frequency of non-radiographic axial spondyloarthritis in relation to symptom duration in patients referred because of chronic back pain: results from the Berlin early spondyloarthritis clinic. Ann Rheum Dis. 2012 Dec;71(12):1998-2001.
49. Davis JC, van der Heijde D, Dougados M, Woolley JM. Reductions in health-related quality of life in patients with ankylosing spondylitis and improvements with etanercept therapy. Arthritis Rheum. 2005 Aug 15;53(4):494-501.
50. Farren W, Goodacre L, Stigant M. Fatigue in ankylosing spondylitis: causes, consequences and self-management. Musculoskeletal Care. 2013 Mar;11(1):39-50.
51. Rudwaleit M, Haibel H, Baraliakos X, Listing J, Marker-Hermann E, Zeidler H, et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum. 2009 Mar;60(3):717-27.
52. Stone MA, Pomeroy E, Keat A, Sengupta R, Hickey S, Dieppe P, et al. Assessment of the impact of flares in ankylosing spondylitis disease activity using the Flare Illustration. Rheumatology (Oxford). 2008 Aug;47(8):1213-8.
53. Salaffi F, Carotti M, Gasparini S, Intorcia M, Grassi W. The health-related quality of life in rheumatoid arthritis, ankylosing spondylitis, and psoriatic arthritis: a comparison with a selected sample of healthy people. Health Qual Life Outcomes. 2009;7:25.
54. Kiltz U, Baraliakos X, Karakostas P, Igelmann M, Kalthoff L, Klink C, et al. Do patients with non-radiographic axial spondylarthritis differ from patients with ankylosing spondylitis? Arthritis Care Res (Hoboken). 2012 Sep;64(9):1415-22.
55. Ara RM, Packham JC, Haywood KL. The direct healthcare costs associated with ankylosing spondylitis patients attending a UK secondary care rheumatology unit. Rheumatology (Oxford). 2008 Jan;47(1):68-71.
56. Boonen A, van der Heijde D, Landewe R, Guillemin F, Rutten-van Molken M, Dougados M, et al. Direct costs of ankylosing spondylitis and its determinants: an analysis among three European countries. Ann Rheum Dis. 2003 Aug;62(8):732-40.
57. Hussain MJ, Brophy S, Cooksey R, Rahman MA, Phillips CJ, Siebert S. The work-related costs of ankylosing spondylitis in a UK cohort. BSR 2014 abstract.
58. Heuft-Dorenbosch L, Landewe R, Weijers R, Wanders A, Houben H, van der Linden S, et al. Combining information obtained from magnetic resonance imaging and conventional radiographs to detect sacroiliitis in patients with recent onset inflammatory back pain. Ann Rheum Dis. 2006 Jun;65(6):804-8.
59. Rudwaleit M, Landewe R, van der Heijde D, Listing J, Brandt J, Braun J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. Ann Rheum Dis. 2009 Jun;68(6):770-6.
60. van der Heijde D, Sieper J, Maksymowych WP, Dougados M, Burgos-Vargas R, Landewe R, et al. 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. Ann Rheum Dis. 2011 Jun;70(6):905-8.
61. Sieper J, Klopsch T, Richter M, Kapelle A, Rudwaleit M, Schwank S, et al. Comparison of two different dosages of celecoxib with diclofenac for the treatment of active ankylosing spondylitis: results of a 12-week randomised, double-blind, controlled study. Ann Rheum Dis. 2008 Mar;67(3):323-9.

62. Wanders A, Heijde D, Landewe R, Behier JM, Calin A, Olivieri I, et al. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum. 2005 Jun;52(6):1756-65.

63. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Marker-Hermann E, Zeidler H, et al. Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Ann Rheum Dis. 2012 Oct;71(10):1616-22.

64. Song IH, Poddubnyy DA, Rudwaleit M, Sieper J. Benefits and risks of ankylosing spondylitis treatment with nonsteroidal antiinflammatory drugs. Arthritis Rheum. 2008 Apr;58(4):929-38.

65. European Medicines Agency. Remicade. Summary of product characteristics. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000240/WC500050888.pdf (Accessed August 2014).

66. Mercer LK, Lunt M, Low AL, Dixon WG, Watson KD, Symmons DP, et al. Risk of solid cancer in patients exposed to anti-tumour necrosis factor therapy: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. Ann Rheum Dis. 2014 Mar 31.

67. Braun J, Zochling J, Baraliakos X, Alten R, Burmester G, Grasedyck K, et al. Efficacy of sulfasalazine in patients with inflammatory back pain due to undifferentiated spondyloarthritis and early ankylosing spondylitis: a multicentre randomised controlled trial. Ann Rheum Dis. 2006 Sep;65(9):1147-53.

68. Dougados M, van der Heijde D, Sieper J, Braun J, Maksymowych WP, Citera G, et al. The symptomatic efficacy and effect on objective signs of inflammation of etanercept in early nonradiographic axial spondyloarthritis. Arthritis Rheumatol. 2014 May 28.

69. Inman RD, Davis JC, Jr., Heijde D, Diekman L, Sieper J, Kim SI, et al. Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis & Rheumatism. 2008 Nov;58(11):3402-12.

70. Revicki DA, Luo MP, Wordsworth P, Wong RL, Chen N, Davis JC, Jr., et al. Adalimumab reduces pain, fatigue, and stiffness in patients with ankylosing spondylitis: results from the adalimumab trial evaluating long-term safety and efficacy for ankylosing spondylitis (ATLAS).[Erratum appears in J Rheumatol. 2011 Apr;38(4):788]. Journal of Rheumatology. 2008 Jul;35(7):1346-53.

71. Van Der Heijde D, Dijkmans B, Geusens P, Sieper J, DeWoody K, Williamson P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: Results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum. 2005 February;52(2):582-91.

72. Baraliakos X, Haibel H, Listing J, Sieper J, Braun J. Continuous long-term anti-TNF therapy does not lead to an increase in the rate of new bone formation over 8 years in patients with ankylosing spondylitis. Ann Rheum Dis. 2014 Apr;73(4):710-5.

73. Van der Heijde D. Effect of certolizumab pegol on inflammation of spine and sacroiliac joints in patients with axial spondyloarthritis: 12-week magnetic resonance imaging results of rapid-axspa study. . BSR. 2013.

74. Wyeth Research. Final report: A randomized double-blind, placebo-controlled trial evaluating safety and efficacy of etanercept 50 mg once weekly compared with 25 mg twice weekly in subjects with ankylosing spondylitis. Protocol number 0881A3-314-EU. CSR59339. 11-10-2005. Data on file.

75. Vastesaeger N, van der Heijde D, Inman RD, Wang Y, Deodhar A, Hsu B, et al. Predicting the outcome of ankylosing spondylitis therapy. Ann Rheum Dis. 2011 Jun;70(6):973-81.

76. Pradeep DJ, Keat AC, Gaffney K, Brooksby A, Leeder J, Harris C. Switching anti-TNF therapy in ankylosing spondylitis. Rheumatology (Oxford). 2008 Nov;47(11):1726-7.

77. Conti F, Ceccarelli F, Marocchi E, Magrini L, Spinelli FR, Spadaro A, et al. Switching tumour necrosis factor alpha antagonists in patients with ankylosing spondylitis and psoriatic arthritis: an observational study over a 5-year period. Ann Rheum Dis. 2007 Oct;66(10):1393-7.

78. Coates LC, Marzo-Ortega H, Bennett AN, Emery P. Anti-TNF Therapy in Ankylosing Spondylitis: Insights for the Clinician. Ther Adv Musculoskelet Dis. 2010 Feb;2(1):37-43.

79. Furst DE, Keystone EC, So AK, Braun J, Breedveld FC, Burmester GR, et al. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2012. Ann Rheum Dis. 2013 Apr;72 Suppl 2:ii2-34.

80. NICE Decision Support Unit. Comparisons of submitted models for the appraisal of etanercept, adalimumab and infliximab for ankylosing spondylitis. 2007.

81. European Medicines Agency. Humira. Summary of product characteristics. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000481/WC500050870.pdf (Accessed August 2014).

82. Vincent FB, Morand EF, Murphy K, Mackay F, Mariette X, Marcelli C. Antidrug antibodies (ADAb) to tumour necrosis factor (TNF)-specific neutralising agents in chronic inflammatory diseases: a real issue, a clinical perspective. Ann Rheum Dis. 2013 Feb;72(2):165-78.

83. Pascual-Salcedo D, Plasencia C, Ramiro S, Nuno L, Bonilla G, Nagore D, et al. Influence of immunogenicity on the efficacy of long-term treatment with infliximab in rheumatoid arthritis. Rheumatology (Oxford). 2011 Aug;50(8):1445-52.

84. Plasencia C, Pascual-Salcedo D, Nuno L, Bonilla G, Villalba A, Peiteado D, et al. Influence of immunogenicity on the efficacy of longterm treatment of spondyloarthritis with infliximab. Ann Rheum Dis. 2012 Dec;71(12):1955-60.

85. Schellekens H. Factors influencing the immunogenicity of therapeutic proteins. Nephrol Dial Transplant. 2005 Jun;20 Suppl 6:vi3-9.

86. de Vries MK, Brouwer E, van der Horst-Bruinsma IE, Spoorenberg A, van Denderen JC, Jamnitski A, et al. Decreased clinical response to adalimumab in ankylosing spondylitis is associated with antibody formation. Ann Rheum Dis. 2009 Nov;68(11):1787-8.

87. de Vries MK, Wolbink GJ, Stapel SO, de Vrieze H, van Denderen JC, Dijkmans BA, et al. Decreased clinical response to infliximab in ankylosing spondylitis is correlated with anti-infliximab formation. Ann Rheum Dis. 2007 Sep;66(9):1252-4.

88. de Vries MK, van der Horst-Bruinsma IE, Nurmohamed MT, Aarden LA, Stapel SO, Peters MJ, et al. Immunogenicity does not influence treatment with etanercept in patients with ankylosing spondylitis. Ann Rheum Dis. 2009 Apr;68(4):531-5.

89. Brandt J, Khariouzov A, Listing J, Haibel H, Sorensen H, Grassnickel L, et al. Six-month results of a double-blind, placebo-controlled trial of etanercept treatment in patients with active ankylosing spondylitis. Arthritis Rheum. 2003 Jun;48(6):1667-75.

90. Landewe RBM, Rudwaleit M, Van Der Heijde D, Dougados M, Maksymowych WP, Braun J, et al. Effect of certolizumab pegol on signs and symptoms of ankylosing spondyltitis and non-radiographic axial spondyloarthritis: 24 week results of a double blind randomized placebo-controlled phase 3 axial spondyloarthritis study. Arthritis Rheum. 2012 October;64:S336-S7.

91. van der Heijde D, Kivitz A, Schiff MH, Sieper J, Dijkmans BA, Braun J, et al. Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis & Rheumatism. 2006 Jul;54(7):2136-46.

92. National Institute for Health and Care Excellence. Multiple Technology Appraisal: TNF-alpha inhibitors for ankylosing spondylitis and axial spondyloarthritis without radiographic evidence of ankylosing spondylitis (including a review of technology appraisal 143 and technology appraisal 233): Final scope. National Institute for Health and Care Excellence. 2014.

93. Gorman JD, Sack KE, Davis JC, Jr. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor alpha. N Engl J Med. 2002 May 2;346(18):1349-56.

94. Barkham N, Keen HI, Coates LC, O'Connor P, Hensor E, Fraser AD, et al. Clinical and imaging efficacy of infliximab in HLA-B27-positive patients with magnetic resonance imaging-determined early sacroiliitis. Arthritis Rheum. 2009 April;60(4):946-54.

95. Braun J, Brandt J, Listing J, Zink A, Alten R, Golder W, et al. Treatment of active ankylosing spondylitis with infliximab: A randomised controlled multicentre trial. Lancet. 2002 06 Apr;359(9313):1187-93.

96. Braun J, Deodhar A, Dijkmans B, Geusens P, Sieper J, Williamson P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis over a two-year period. Arthritis Care and Research. 2008 15 Sep;59(9):1270-8.

97. Braun J, Landewe R, Hermann KG, Han J, Yan S, Williamson P, et al. Major reduction in spinal inflammation in patients with ankylosing spondylitis after treatment with infliximab: results of a multicenter, randomized, double-blind, placebo-controlled magnetic resonance imaging study. Arthritis & Rheumatism. 2006 May;54(5):1646-52.

98. Braun J, McHugh N, Singh A, Wajdula JS, Sato R. Improvement in patient-reported outcomes for patients with ankylosing spondylitis treated with etanercept 50 mg once-weekly and 25 mg twice-weekly. Rheumatology (Oxford). 2007 Jun;46(6):999-1004.

99. Braun J, Pavelka K, Ramos-Remus C, Dimic A, Vlahos B, Freundlich B, et al. Clinical efficacy of etanercept versus sulfasalazine in ankylosing spondylitis subjects with peripheral joint involvement. Journal of Rheumatology. 2012 April;39(4):836-40.

100. Breban M, Ravaud P, Claudepierre P, Baron G, Henry YD, Hudry C, et al. Maintenance of infliximab treatment in ankylosing spondylitis: Results of a one-year randomized controlled trial comparing systematic versus on-demand treatment. Arthritis Rheum. 2008 January;58(1):88-97.

101. Cantini F, Niccoli L, Cassara E, Kaloudi O, Nannini C. Duration of remission after halving of the etanercept dose in patients with ankylosing spondylitis: A randomized, prospective, long-term, follow-up study. Biologics: Targets and Therapy. 2013;7(1):1-6.

102. Davis Jr JC, Van Der Heijde D, Braun J, Dougados M, Cush J, Clegg DO, et al. Recombinant Human Tumor Necrosis Factor Receptor (Etanercept) for Treating Ankylosing Spondylitis: A Randomized, Controlled Trial. Arthritis Rheum. 2003 November;48(11):3230-6.

103. Dougados M, Braun J, Szanto S, Combe B, Elbaz M, Geher P, et al. Efficacy of etanercept on rheumatic signs and pulmonary function tests in advanced ankylosing spondylitis: results of a randomised double-blind placebo-controlled study (SPINE).[Erratum appears in Ann Rheum Dis. 2011 Jul;70(7):1349]. Ann Rheum Dis. 2011 May;70(5):799-804.

104. Giardina AR, Ferrante A, Ciccia F, Impastato R, Miceli MC, Principato A, et al. A 2-year comparative open label randomized study of efficacy and safety of etanercept and infliximab in patients with ankylosing spondylitis. Rheumatol Int. 2010 September;30(11):1437-40.

105. Haibel H, Rudwaleit M, Listing J, Heldmann F, Wong RL, Kupper H, et al. Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis: Results of a twelve-week randomized, double-blind, placebo-controlled

trial followed by an open-label extension up to week fifty-two. Arthritis Rheum. 2008 July;58(7):1981-91.

106. Inman RD, Maksymowych WP. A double-blind, placebo-controlled trial of low dose infliximab in ankylosing spondylitis. Journal of Rheumatology. 2010 June;37(6):1203-10.

107. Lambert RG, Salonen D, Rahman P, Inman RD, Wong RL, Einstein SG, et al. Adalimumab significantly reduces both spinal and sacroiliac joint inflammation in patients with ankylosing spondylitis: a multicenter, randomized, double-blind, placebo-controlled study. Arthritis & Rheumatism. 2007 Dec;56(12):4005-14.

108. Maksymowych WP, Salonen D, Inman RD, Rahman P, Lambert RG, Group CS. Low-dose infliximab (3 mg/kg) significantly reduces spinal inflammation on magnetic resonance imaging in patients with ankylosing spondylitis: a randomized placebo-controlled study. Journal of Rheumatology. 2010 Aug 1;37(8):1728-34.

109. Marzo-Ortega H, McGonagle D, Jarrett S, Haugeberg G, Hensor E, O'Connor P, et al. Infliximab in combination with methotrexate in active ankylosing spondylitis: A clinical and imaging study. Ann Rheum Dis. 2005 November;64(11):1568-75.

110. Navarro-Sarabia F, Fernandez-Sueiro JL, Torre-Alonso JC, Gratacos J, Queiro R, Gonzalez C, et al. High-dose etanercept in ankylosing spondylitis: results of a 12-week randomized, double blind, controlled multicentre study (LOADET study). Rheumatology (Oxford). 2011 Oct;50(10):1828-37.

111. Park W, Hrycaj P, Jeka S, Kovalenko V, Lysenko G, Miranda P, et al. A randomised, double-blind, multicentre, parallel-group, prospective study comparing the pharmacokinetics, safety, and efficacy of CT-P13 and innovator infliximab in patients with ankylosing spondylitis: the PLANETAS study. Ann Rheum Dis. 2013 Oct;72(10):1605-12.

112. Sieper J, Van Der Heijde D, Dougados M, Mease P, Maksymowych WP, Brown MA, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: Results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis. 2013 June;72(6):815-22.

113. Song IH, Hermann K, Haibel H, Althoff CE, Althoff C, Listing J, et al. Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): a 48-week randomised controlled trial. Ann Rheum Dis. 2011 Apr;70(4):590-6.

114. van der Heijde DM, Revicki DA, Gooch KL, Wong RL, Kupper H, Harnam N, et al. Physical function, disease activity, and health-related quality-of-life outcomes after 3 years of adalimumab treatment in patients with ankylosing spondylitis. Arthritis Res Ther. 2009;11(4):R124.

115. Song IH, Weis A, Hermann KG, Haibel H, Althoff CE, Poddubnyy D, et al. Similar response rates in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis after 1 year of treatment with etanercept: results from the ESTHER trial. Ann Rheum Dis. 2013 Jun;72(6):823-5.

116. Brandt J, Listing J, Haibel H, Sorensen H, Schwebig A, Rudwaleit M, et al. Long-term efficacy and safety of etanercept after readministration in patients with active ankylosing spondylitis. Rheumatology (Oxford). 2005 Mar;44(3):342-8.

117. Davis J, Jr., Webb A, Lund S, Sack K. Results from an open-label extension study of etanercept in ankylosing spondylitis. Arthritis Rheum. 2004 Apr 15;51(2):302-4.

118. Davis JC, Jr., van der Heijde DM, Braun J, Dougados M, Clegg DO, Kivitz AJ, et al. Efficacy and safety of up to 192 weeks of etanercept therapy in patients with ankylosing spondylitis. Ann Rheum Dis. 2008 Mar;67(3):346-52.

119. Davis JC, van der Heijde DM, Braun J, Dougados M, Cush J, Clegg D, et al. Sustained durability and tolerability of etanercept in ankylosing spondylitis for 96 weeks. Ann Rheum Dis. 2005 Nov;64(11):1557-62.

120. Dijkmans B, Emery P, Hakala M, Leirisalo-Repo M, Mola EM, Paolozzi L, et al. Etanercept in the longterm treatment of patients with ankylosing spondylitis. J Rheumatol. 2009 Jun;36(6):1256-64.

121. Dougados M, Braun J, Szanto S, Combe B, Geher P, Leblanc V, et al. Continuous efficacy of etanercept in severe and advanced ankylosing spondylitis: results from a 12-week open-label extension of the SPINE study. Rheumatology (Oxford). 2012 Sep;51(9):1687-96.

122. Martin-Mola E, Sieper J, Leirisalo-Repo M, Dijkmans BA, Vlahos B, Pedersen R, et al. Sustained efficacy and safety, including patient-reported outcomes, with etanercept treatment over 5 years in patients with ankylosing spondylitis. Clin Exp Rheumatol. 2010 Mar-Apr;28(2):238-45.

123. Song IH, Althoff CE, Haibel H, Hermann KG, Poddubnyy D, Listing J, et al. Frequency and duration of drug-free remission after 1 year of treatment with etanercept versus sulfasalazine in early axial spondyloarthritis: 2 year data of the ESTHER trial. Ann Rheum Dis. 2012 Jul;71(7):1212-5.

124. van der Heijde D, Landewe R, Einstein S, Ory P, Vosse D, Ni L, et al. Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. Arthritis Rheum. 2008 May;58(5):1324-31.

125. van der Heijde D, Zack D, Wajdula J, Sridharan S, Koenig A. Rates of serious infections, opportunistic infections, inflammatory bowel disease, and malignancies in subjects receiving etanercept vs. controls from clinical trials in ankylosing spondylitis: a pooled analysis. Scand J Rheumatol. 2014;43(1):49-53.

126. Dougados M, van der Heijde D, Sieper J. . Safety and Efficacy Of Etanercept In Early Non-Radiographic Axial Spondyloarthritis: A Randomized, Double-Blind, Placebo-Controlled Trial At 24 Weeks. Arthritis & Rheumatism. 2013;65(10 (supplement)).

127. Dougados M TW, Saaibi D, . Patient-Reported Outcomes of Etanercept in Early Non-Radiographic Axial Spondyloarthritis: A 12-Week, Randomized, Double-Blind, Placebo-Controlled Trial. Arthritis & Rheumatism. 2013;65(10 (supplement)).

128. Dougados M, van der Heijde D, Sieper J, Braun J. Clinical and imaging efficacy of etanercept in early non-radiographic axial spondyloarthritis: A 12-week, randomized, double-blind, placebo-controlled trial. EULAR. 2013.

129. Hermann K, Braun J, Investigators. G-RS. The effect of golimumab on structural spinal changes in ankylosing spondylitis: Magnetic resonance imaging results of the placebo-controlled GO-RAISE study. EULAR. 2011.

130. Park W, Jaworski J, Brzezicki J. A randomised, double-blind, parallel-group, phase 1 study comparing the pharmacokinetics, safety and efficacy of CT-P13 and infliximab in patients with active ankylosing spondylitis: 54 week results from the PLANETAS study. EULAR. 2013.

131. Sieper J RM, Lenaerts J. ROLE OF BASELINE C-REACTIVE PROTEIN IN RESPONSE TO INFLIXIMAB PLUS NAPROXEN VS NAPROXEN ALONE IN PATIENTS WITH AXIAL SPONDYLOARTHRITIS IN THE INFAST STUDY. EULAR. 2013.

132. Song I-H, Hermann, Kay-Geert A. Haibel, Hildrun. Stable Clinical and MRI Response In Patients With Early Axial Spondyloarthritis After 3 Years Of Continuous Treatment With Etanercept, Data Of The Esther Trial. ACR. 2013.

133. Braun J, Baraliakos X, Hermann KGA, Deodhar A, Van Der Heijde D, Inman R, et al. The effect of two golimumab doses on radiographic progression in ankylosing spondylitis: Results through 4 years of the GO-RAISE trial. Ann Rheum Dis. 2014 June;73(6):1107-13.

134. Van der Heijde D. Effect of certolizumab pegol on inflammation of spine and sacroiliac joints in patients with axial spondyloarthritis: 12-week magnetic resonance imaging results of rapid-axspa study. . BSR. 2014.

135. Gaffney K. An open label, pilot, multi-centre, stepdown, randomized controlled trial to examine whether etanercept 25mg once weekly is effective in maintaining a clinical response in patients with ankylosing spondylitis who have responded to 50mg once weekly. BSR. 2014.
136. Maksymowych W. Clinical and imaging efficacy of etanercept in early non-radiographic axial spondyloarthritis: 48-week treatment data. . EULAR. 2014.
137. Van der Heijde D. Asas40 and asdas responses are associated with improved physical function, hrqol, and work productivity in patients with non-radiographic axial spondyloarthritis. . EULAR. 2014.
138. Sieper J. Quality of life with etanercept in early non-radiographic axial spondyloarthritis: 24 and 48-week data from a randomized, double-blind, placebo-controlled trial. . EULAR. 2014.
139. FDA: Clinical Review for the Arthritis Advisory Committee Meeting. Humira® (Adalimumab) BLA 125057/323: For active non-radiographic axial spondyloarthritis (nr-axSpA). Food & Drug Administration. 2013.
140. FDA: Clinical Review for the Arthritis Advisory Committee Meeting. Cimzia® (Certolizumab) BLA 125160/215: Active axial spondyloarthritis, including patients with ankylosing spondylitis. Food & Drug Administration. 2013.
141. Dougados M. A multicentre, 12-week double blind placebo controlled randomized study of etanercept on a background NSAID in the treatment of adult subjects with non radiographic axial spondyloarthritis with a 92-week open label extension. Clinical study report. Data on file. . 2013.
142. Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. Arthritis Rheum. 2001 Aug;44(8):1876-86.
143. European Medicines Agency. Committee for medicinal products for human use (CHMP). Guideline on clinical investigation of medicinal products for the treatment of ankylosing spondylitis. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003424.pdf (Accessed 16 June 2014). 2009.
144. Lukas C, Landewe R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis. 2009 Jan;68(1):18-24.
145. National Ankylosing Spondylitis Society (NASS). Bath Indices. The Bath Ankylosing Spondylitis Disease Activity Index (BASDAI). Available from: http://www.nass.co.uk/loose-leaf-pages/resources-for-health-professionals-2/bath-indices/?keywords=Bath+indices (Accessed 21/02/14).
146. National Ankylosing Spondylitis Society (NASS). Bath Indices. The Bath Ankylosing Spondylitis Functional Index (BASFI). Available from: http://www.nass.co.uk/loose-leaf-pages/resources-for-health-professionals-2/bath-indices/?keywords=Bath+indices (Accessed 21/02/14).
147. Moncur C. Ankylosing spondylitis measures. Arthritis & Rheumatism (Arthritis Care & Research). 2003;49(5S):S197-S209.
148. National Ankylosing Spondylitis Society (NASS). Bath Indices. The Bath Ankylosing Spondylitis Metrology Index (BASMI). Available from: http://www.nass.co.uk/loose-leaf-pages/resources-for-health-professionals-2/bath-indices/?keywords=Bath+indices (Accessed 21/02/14).
149. Gould D. Visual Analogue Scale (VAS). Available from: http://www.blackwellpublishing.com/specialarticles/jcn_10_706.pdf (Accessed 21/02/14).
150. Viitanen JV, Kokko ML, Heikkila S, Kautiainen H. Neck mobility assessment in ankylosing spondylitis: a clinical study of nine measurements including new tape

methods for cervical rotation and lateral flexion. Br J Rheumatol. 1998 Apr;37(4):377-81.

151. Creemers MC, Franssen MJ, van't Hof MA, Gribnau FW, van de Putte LB, van Riel PL. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis. 2005 Jan;64(1):127-9.

152. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. J Rheumatol. 2007 Apr;34(4):862-70.

153. Electronic Medicines Compendium (eMC). Etopan XL 600 mg Tablets: SPC. Available from: https://www.medicines.org.uk/emc/medicine/27532#CONTRAINDICATIONS (Accessed August 2014).

154. Cowling P, Ebringer R, Cawdell D, Ishii M, Ebringer A. C-reactive protein, ESR, and klebsiella in ankylosing spondylitis. Ann Rheum Dis. 1980 Feb;39(1):45-9.

155. de Vries MK, van Eijk IC, van der Horst-Bruinsma IE, Peters MJ, Nurmohamed MT, Dijkmans BA, et al. Erythrocyte sedimentation rate, C-reactive protein level, and serum amyloid a protein for patient selection and monitoring of anti-tumor necrosis factor treatment in ankylosing spondylitis. Arthritis Rheum. 2009 Nov 15;61(11):1484-90.

156. EuroQol Group. EQ-5D. What is EQ-5D? Available from: www.euroqol.org (Accessed 11 April 2014)

157. Ware JE. SF-36.org. A community for measuring health outcomes using SF tools. Available from: www.sf-36.org/tools/sf36.shtml (Accessed 11 April 2014).

158. GL Assessment. Hospital Anxiety and Depression Scale (HADS). Available from: http://www.gl-assessment.co.uk/products/hospital-anxiety-and-depression-scale-0 (Accessed 05/03/14).

159. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis. 2005 Jan;64(1):29-33.

160. Maksymowych WP, Gooch K, Dougados M, Wong RL, Chen N, Kupper H, et al. Thresholds of patient-reported outcomes that define the patient acceptable symptom state in ankylosing spondylitis vary over time and by treatment and patient characteristics. Arthritis Care Res (Hoboken). 2010 Jun;62(6):826-34.

161. Dougados M. Interim analysis at Week 48 of open-label extension: A multicentre, 12-week double-blind placebo-controlled randomized study of etanercept on a background NSAID in the treatment of adult subjects with non-radiographic axial spondyloarthritis with a 92-week open-label extension. Clinical study report. Data on file. 2014.

162. Song IH, Althoff CE, Haibel H, Hermann KG, Poddubnyy D, Listing J, et al. Frequency and duration of drug-free remission after 1 year of treatment with etanercept versus sulfasalazine in early axial spondyloarthritis: 2 year data of the ESTHER trial. Ann Rheum Dis. 2012 Jul;71(7):1212-5.

163. Wyeth Research. Final report. An open-label, multicenter, long-term extension study of etanercept in ankylosing spondylitis. Protocol 0881A3-101615. CSR-70708. Data on file.

164. Pavelka K, Forejtova S, Stolfa J, Chroust K, Buresova L, Mann H, et al. Anti-TNF therapy of ankylosing spondylitis in clinical practice. Results from the Czech national registry ATTRA. Clin Exp Rheumatol. 2009 Nov-Dec;27(6):958-63.

165. van der Heijde D, Zack D, Wajdula J, Sridharan S, Koenig AS. Rates of serious infections, opportunistic infections, inflammatory bowel disease, and malignancies in

subjects receiving etanercept vs. controls from clinical trials in ankylosing spondylitis: a pooled analysis. Scand J Rheumatol. 2014;43(1):49-53.

166. Dixon WG, Hyrich KL, Watson KD, Lunt M, Galloway J, Ustianowski A, et al. Drug-specific risk of tuberculosis in patients with rheumatoid arthritis treated with anti-TNF therapy: results from the British Society for Rheumatology Biologics Register (BSRBR). Ann Rheum Dis. 2010 Mar;69(3):522-8.

167. Tubach F, Salmon D, Ravaud P, Allanore Y, Goupille P, Breban M, et al. Risk of tuberculosis is higher with anti-tumor necrosis factor monoclonal antibody therapy than with soluble tumor necrosis factor receptor therapy: The three-year prospective French Research Axed on Tolerance of Biotherapies registry. Arthritis Rheum. 2009 Jul;60(7):1884-94.

168. Manadan AM, Joyce K, Sequeira W, Block JA. Etanercept therapy in patients with a positive tuberculin skin test. Clin Exp Rheumatol. 2007 Sep-Oct;25(5):743-5.

169. The European Parliament and the Council of the European Union. Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:121:0034:0044:en:PDF (Accessed September 2014).

170. MHRA. Good Pharmacovigilance Practice Guide. Page 189. Pharmaceutical Press, London. 2008.

171. The European Parliament and the Council of the European Union. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:311:0067:0128:en:PDF (Accessed September 2014).

172. Davis J, Gorman JD, Fye K, Sack K, Wofsy D. Anti-tumor necrosis factor (TNFR:Fc) in ankylosing spondylitis (a Phase 2 trial). Appendix VII. Results of the 6 month open-label extension. Clinical study report. Data on file. .

173. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. Available from: http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9 (Accessed June 2014).

174. Dempster AP. The direct use of likelihood for significance testing. Statistics and Computing. 1997;7 (4):247-52.

175. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. Stat Med. 2009 Nov 10;28(25):3049-67.

176. Spiegelhalter D, Thomas A, Best N, Lunn D. WinBUGS user manual version 1.4. 2003.

177. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010 Mar 30;29(7-8):932-44.

178. Braun J, Huang F, Burgos-Vargas R, van der Horst-Bruinsma I, Freundlich B, Vlahos B. Assessment of clinical efficacy in a randomized, double-blind study of etanercept and sulphasalazine in patients with ankylosing spondylitis. . Arthritis & Rheumatism. 2008;58 (Suppl):S415.

179. Martin-Mola E, Van der Linden S, Wishneski C, Robertson D, Freundlich B, Koenig A. Long term efficacy and safety, including patient-reported outcomes, with etanercept for 5 years in ankylosing spondylitis. Ann Rheum Dis. 2008;67 (Suppl II):510.

180. Braun J, McHugh N, Singh A, Wajdula JS, Sato R. Improvement in patient-reported outcomes for patients with ankylosing spondylitis treated with etanercept 50 mg once-weekly and 25 mg twice-weekly. Rheumatology. 2007 June;46(6):999-1004.

181. Braun J, Van Der Horst-Bruinsma IE, Huang F, Burgos-Vargas R, Vlahos B, Koenig AS, et al. Clinical efficacy and safety of etanercept versus sulfasalazine in patients with ankylosing spondylitis: A randomized, double-blind trial. Arthritis and Rheumatism. 2011 June;63(6):1543-51.

182. Gorman JD, Sack KE, Davis JC, Jr. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor alpha. New England Journal of Medicine. 2002 May 2;346(18):1349-56.

183. Navarro-Sarabia F, Fernandez-Sueiro JL, Torre-Alonso JC, Gratacos J, Queiro R, Gonzalez C, et al. High-dose etanercept in ankylosing spondylitis: Results of a 12-week randomized, double blind, controlled multicentre study (LOADET study). Rheumatology. 2011 October;50(10):1828-37.

184. Van Der Heijde D, Da Silva JC, Dougados M, Geher P, Van Der Horst-Bruinsma I, Juanola X, et al. Etanercept 50 mg once weekly is as effective as 25 mg twice weekly in patients with ankylosing spondylitis. Annals of the rheumatic diseases. 2006 December;65(12):1572-7.

185. Dougados M, Tsai WC, Saaibi DL, Bonin R, Bukowski J, Pedersen R, et al. Patient-reported outcomes of etanercept in early non-radiographic axial spondyloarthritis: A 12-week, randomized, double-blind, placebo-controlled trial. . Arthritis Rheum. 2013;65(10 (Suppl)):S656-7.

186. European Medicines Agency. Assessment report: Cimzia. 2013;Procedure no. EMEA/H/C/001037/II/29.

187. Riemsma R, Joore M, van Asselt T, Armstrong H, Kleijnen J. Golimumab for the treatment of ankylosing spondylitis: A Single Technology Appraisal. York: Kleijnen Systematic Reviews Ltd. . 2011.

188. Scottish Medicines Consortium. Certolizumab pegol 200mg/mL solution for injection in pre-filled syringe (Cimzia®). SMC No. (960/14). UCB Pharma UK. 04 April 2014. Available from: http://www.scottishmedicines.org.uk/files/advice/certolizumab_pegol_Cimzia_FINAL_April_2014_for_website.pdf (Accessed June 2014).

189. Signorovitch JE, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao Y, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. PharmacoEconomics. 2010;28(10):935-45.

190. Caro JJ, Ishak KJ. No head-to-head trial? simulate the missing arms. PharmacoEconomics. 2010;28(10):957-67.

191. Rafia R, Ara R, Packham J, Haywood KL, Healey E. Healthcare costs and productivity losses directly attributable to ankylosing spondylitis. Clin Exp Rheumatol. 2012 Mar-Apr;30(2):246-53.

192. Lord PA, Farragher TM, Lunt M, Watson KD, Symmons DP, Hyrich KL, et al. Predictors of response to anti-TNF therapy in ankylosing spondylitis: results from the British Society for Rheumatology Biologics Register. Rheumatology (Oxford). 2010 Mar;49(3):563-70.

193. Ara RM, Reynolds AV, Conway P. The cost-effectiveness of etanercept in patients with severe ankylosing spondylitis in the UK. Rheumatology (Oxford). 2007 Aug;46(8):1338-44.

194. Botteman MF, Hay JW, Luo MP, Curry AS, Wong RL, van Hout BA. Cost effectiveness of adalimumab for the treatment of ankylosing spondylitis in the United Kingdom. Rheumatology (Oxford). 2007 Aug;46(8):1320-8.

195. Farrell J, Muszbek N, Sheppard O, Chaudhary M, Naci H, Kachroo S. The long term cost-effectiveness of golimumab for the treatment of severe, active ankylosing spondylitis in adults who have responded inadequately to conventional therapy. . Value in Health. 2011;14 (7):A311.

196. Fautrel B, Benhamou M, Breban M, Roy C, Lenoir C, Trape G, et al. Cost effectiveness of two therapeutic regimens of infliximab in ankylosing spondylitis: economic evaluation within a randomised controlled trial. Ann Rheum Dis. 2010 Feb;69(2):424-7.

197. Kobelt G, Andlin-Sobocki P, Brophy S, Jonsson L, Calin A, Braun J. The burden of ankylosing spondylitis and the cost-effectiveness of treatment with infliximab (Remicade). Rheumatology (Oxford). 2004 Sep;43(9):1158-66.

198. Kobelt G, Sobocki P, Mulero J, Gratacos J, Collantes-Estevez E, Braun J. The cost-effectiveness of infliximab in the treatment of ankylosing spondylitis in Spain. Comparison of clinical trial and clinical practice data. Scand J Rheumatol. 2008 Jan-Feb;37(1):62-71.

199. Kobelt G, Sobocki P, Sieper J, Braun J. Comparison of the cost-effectiveness of infliximab in the treatment of ankylosing spondylitis in the United Kingdom based on two different clinical trials. Int J Technol Assess Health Care. 2007 Summer;23(3):368-75.

200. Marcellusi A, Botteman MF, Rao S, Cifaldi M, Solem CT, Gitto L, et al. Cost effectiveness of adalimumab versus golimumab and placebo in ankylosing spondylitis in Italy. . Value in Health. 2012;15(7):A444.

201. McLeod C, Bagust A, Boland A, Dagenais P, Dickson R, Dundar Y, et al. Adalimumab, etanercept and infliximab for the treatment of ankylosing spondylitis: a systematic review and economic evaluation. Health Technol Assess. 2007 Aug;11(28):1-158, iii-iv.

202. Muszbek N, Punekar Y, Buchanan J, Prasad M, Vioix H, Brown R. Cost-effectiveness of golimumab in ankylosing spondylitis from the UK payer perspective. Value in Health. 2012 November;13 (7):A467.

203. Neilson AR, Sieper J, Deeg M. Cost-effectiveness of etanercept in patients with severe ankylosing spondylitis in Germany. Rheumatology (Oxford). 2010 Nov;49(11):2122-34.

204. Glintborg B, Østergaard M, Krogh NS, Tarp U, Manilo N, Loft AGR, et al. Clinical response, drug survival and predictors thereof in 432 ankylosing spondylitis patients after switching tumour necrosis factor α inhibitor therapy: results from the Danish nationwide DANBIO registry. Ann Rheum Dis. 2013;72:1149–55.

205. Sutton A, Ades AE, Cooper N, Abrams K. Use of Indirect and Mixed Treatment Comparisons for Technology Assessment. PharmacoEconomics. 2008;26(9):753-67.

206. HM treasury. The Green Book: Appraisal and Evaluation in Central Government. London: TSO; 2003.

207. Scottish Medicines Consortium. Adalimumab 40 mg pre-filled syringe (Humira®). (No. 300/06). Abbott Laboratories Ltd. November 2006.

208. Scottish Medicines Consortium. Etanercept 25 mg vial of powder for subcutaneous injection (Enbrel®). (No. 212/05). Wyeth. 4 October 2005.

209. Brophy S, Mackay K, Al-Saili A, Taylor G, Calin A. The natural history of ankylosing spondylitis as defined by radiological progression. J Rheumatol. 2002;29(6):1236-43.

210. Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, et al. Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. Rheumatology (Oxford). 2007 Aug;46(8):1345-54.

211. NICE Decision Support Unit. Patient-level simulation TSD. TSD 15: Cost-effectiveness modelling using patient-level simulation. Available from:

http://www.nicedsu.org.uk/Patientlevel-simulation-TSD(2892880).htm (Accessed June 2014).

212. Isdale A, Keat A, Barkham N, Bennett AN, Gaffney K, Marzo-Ortega H, et al. Expanding the spectrum of inflammatory spinal disease: AS it was, as it is now. Rheumatology (Oxford). 2013 Dec;52(12):2103-5.

213. van Tubergen A, Weber U. Diagnosis and classification in spondyloarthritis: identifying a chameleon. Nat Rev Rheumatol. 2012 May;8(5):253-61.

214. Tannenbaum SJ, Holford NH, Lee H, Peck CC, Mould DR. Simulation of correlated continuous and categorical variables using a single multivariate distribution. J Pharmacokinet Pharmacodyn. 2006 Dec;33(6):773-94.

215. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis. 2009 Jun;68 Suppl 2:ii1-44.

216. All Wales Therapeutics and Toxicology Centre. AWMSG Secretariat Assessment Report. Adalimumab (Humira®). 40 mg prefilled pen or 40 mg prefilled syringe. Reference number: 1381. Full submission. May 2013.

217. Ariza-Ariza R, Hernandez-Cruz B, Lopez-Antequera G, Navarro-Sarabia F. Variables related to utility in patients with ankylosing spondylitis. Clin Rheumatol. 2009 Feb;28(2):207-11.

218. Ariza-Ariza R, Hernandez-Cruz B, Navarro-Compan V, Leyva Pardo C, Juanola X, Navarro-Sarabia F. A comparison of telephone and paper self-completed questionnaires of main patient-related outcome measures in patients with ankylosing spondylitis and psoriatic arthritis. Rheumatol Int. 2013 Nov;33(11):2731-6.

219. Ariza-Ariza R, Hernandez-Cruz B, Navarro-Sarabia F. Physical function and health-related quality of life of Spanish patients with ankylosing spondylitis. Arthritis Rheum. 2003 Aug 15;49(4):483-7.

220. Boonen A, van der Heijde D, Landewe R, Guillemin F, Spoorenberg A, Schouten H, et al. Costs of ankylosing spondylitis in three European countries: the patient's perspective. Ann Rheum Dis. 2003 Aug;62(8):741-7.

221. Boonen A, van der Heijde D, Landewe R, Spoorenberg A, Schouten H, Rutten-van Molken M, et al. Work status and productivity costs due to ankylosing spondylitis: comparison of three European countries. Ann Rheum Dis. 2002 May;61(5):429-37.

222. Boonen A, van der Heijde D, Severens JL, Boendermaker A, Landewe R, Braun J, et al. Markov model into the cost-utility over five years of etanercept and infliximab compared with usual care in patients with active ankylosing spondylitis. Ann Rheum Dis. 2006 Feb;65(2):201-8.

223. Choi CB, Kim TJ, Park HJ, Uhm WS, Jun JB, Bae SC, et al. Safety and clinical responses in ankylosing spondylitis after three months of etanercept therapy. J Korean Med Sci. 2008 Oct;23(5):852-6.

224. Ciurea A, Scherer A, Exer P, Bernhard J, Dudler J, Beyeler B, et al. Tumor necrosis factor alpha inhibition in radiographic and nonradiographic axial spondyloarthritis: results from a large observational cohort. Arthritis Rheum. 2013 Dec;65(12):3096-106.

225. Goycochea-Robles MV, Arce-Salinas CA, Guzman-Vazquez S, Cardiel-Rios MH. Prescription rheumatology practices among Mexican specialists. Arch Med Res. 2007 Apr;38(3):354-9.

226. Haglund E, Bremander A, Bergman S, Jacobsson LT, Petersson IF. Work productivity in a population-based cohort of patients with spondyloarthritis. Rheumatology (Oxford). 2013 Sep;52(9):1708-14.

227. Haglund e, Petersson IF, Strombeck B, Bergman S, Bremander A. Self reported productivity at work in people with ankylosing spondylitis. EULAR abstract. Ann Rheum Dis. 2011;70 (Suppl 3):328.

228. Haywood KL, Garratt AM, Dziedzic K, Dawes PT. Generic measures of health-related quality of life in ankylosing spondylitis: reliability, validity and responsiveness. Rheumatology (Oxford). 2002 Dec;41(12):1380-7.

229. Healey EL, Haywood KL, Jordan KP, Garratt AM, Packham JC. Patients with well-established ankylosing spondylitis show limited deterioration in a ten-year prospective cohort study. Clin Rheumatol. 2013 Jan;32(1):67-72.

230. Kobelt G, Andlin-Sobocki P, Collantes-Estevez E, Gratacos-Masmitja J, Povoci A, Mulero-Medoza J. The cost of ankylosing spondylitis in Spain. An observational study in 600 patients. Abstract. Ann Rheum Dis. 2006;65 (Suppl III):283.

231. Kobelt G, Andlin-Sobocki P, Maksymowych WP. Costs and quality of life of patients with ankylosing spondylitis in Canada. J Rheumatol. 2006 Feb;33(2):289-95.

232. Kobelt G, Sobocki P, Mulero J, Gratacos J, Pocovi A, Collantes-Estevez E. The burden of ankylosing spondylitis in Spain. Value Health. 2008 May-Jun;11(3):408-15.

233. Kvamme MK, Kristiansen IS, Lie E, Kvien TK. Identification of cutpoints for acceptable health status and important improvement in patient-reported outcomes, in rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. J Rheumatol. 2010 Jan;37(1):26-31.

234. Lee TJ, Park BH, Kim JW, Shin K, Lee EB, Song YW. Cost-of-illness and quality of life in patients with ankylosing spondylitis at a tertiary hospital in Korea. J Korean Med Sci. 2014 Feb;29(2):190-7.

235. Lee TJ, Park BH, Son HK, Kim JW, Shin K, Lee EB, et al. Differences in cost-of-illness and quality of life between rheumatoid arthritis and ankylosing spondylitis in South Korea. Value in Health. 2011;14 (7):A306.

236. Lie E, Lillegraven S, van der Heijde D, Kvamme MK, Uhlig T, Kvien T. EQ-5D and SF-6D perform differently in ankylosing spondylitis (AS): A follow-up study of patients receiving disease modifying therapy. Abstract. Scand J Rheumatol. 2010;39:31.

237. Lindstrom U, Bremander A, Bergman S, Haglund E, Petersson IF, Jacobsson LT. Patients with non-AS axial SpA have similar prevalence compared to AS, but worse perceived health. Results from a population based study. . Ann Rheum Dis. 2013;72 (Suppl 3):667.

238. Miranda L, Negreiro F, Queiroz MJ, Silva C. [RAISE study observational and cross sectional study to evaluate the actual reality of the socio-economic impact of ankylosing spondylitis]. Acta Reumatol Port. 2008 Apr-Jun;33(2):189-97.

239. Fernandez M, Lazaro S, Blasco P, Gracia AJ, Cano-Garcia SF, Ivorra L, et al. Quality of life and adherence to treatment in patients managed in nursing clinics in rheumatology. Arthritis Rheum. 2013;65(Abstract Supplement).

240. Gordeev VS, Maksymowych WP, Evers SM, Ament A, Schachna L, Boonen A. Role of contextual factors in health-related quality of life in ankylosing spondylitis. Ann Rheum Dis. 2010 Jan;69(1):108-12.

241. Mustur D, Vesovic-Potic V, Stanisavljevic D, Ille T, Ille M. Assessment of functional disability and quality of life in patients with ankylosing spondylitis. Srp Arh Celok Lek. 2009 Sep-Oct;137(9-10):524-8.

242. Park YS, Kim HS, Baek SW, Oh JH. Preoperative computer-based simulations for the correction of kyphotic deformities in ankylosing spondylitis patients. Spine J. 2014 Feb 15.

243. Pimentel-Santos FM, Pinto T, Santos H, Barcelos A, Cunha I, Branco JC, et al. Portuguese version of the bath indexes for ankylosing spondylitis patients: a cross-cultural adaptation and validation. Clin Rheumatol. 2012 Feb;31(2):341-6.

244. Poddubnyy D, Callhoff J, Listing J, Sieper J. Ustekinumab for the treatment of patients with active ankylosing spondylitis: Results of a 28-week, prospective, open-label, proof-of- concept study (TOPAS). ACR Abstract. 2013.

245. Sarka F, Herman M, Tomas P, Karel C, Karel V, Karel P. Quality of life in the patients with ankylosing spondylitis in the Czech Republic. Journal of Clinical Rheumatology. 2010;16:S71.

246. Van Tubergen A, Boonen A, Landewe R, Rutten-Van Molken M, Van Der Heijde D, Hidding A, et al. Cost effectiveness of combined spa-exercise therapy in ankylosing spondylitis: a randomized controlled trial. Arthritis Rheum. 2002 Oct 15;47(5):459-67.

247. Aktekin LA, Eser F, Malhan S, Oksuz E, Keskin D, Bodur H. A comparison of four different HRQoL generic questionnaire in five different patient groups. Rheumatol Int. 2009 Nov;30(1):63-7.

248. Gooch K, Feeny D, Wong RL, Kupper H, Pangan AL, Revicki DA, et al. Is the Health Utilities Index 3 valid for patients with ankylosing spondylitis? Value Health. 2011 Jan;14(1):160-5.

249. Gooch K, Feeny DH, Wong RL, Pangan A, Revicki D, Van der Heijde D. Is the health utilities index valid and responsive in assessing patients with ankylosing spondylitis? Value in Health. 2009;12 (3):A71.

250. Heiberg MS, Nordvag BY, Mikkelsen K, Rodevand E, Kaufmann C, Mowinckel P, et al. The comparative effectiveness of tumor necrosis factor-blocking agents in patients with rheumatoid arthritis and patients with ankylosing spondylitis: a six-month, longitudinal, observational, multicenter study. Arthritis Rheum. 2005 Aug;52(8):2506-12.

251. Kimel M, Revicki D, Rao S, Fryback D, Feeny D, Harnam N, et al. Norms-based assessment of patient-reported outcomes associated with adalimumab monotherapy in patients with ankylosing spondylitis. Clin Exp Rheumatol. 2011 Jul-Aug;29(4):624-32.

252. Maksymowych WP, Gooch KL, Wong RL, Kupper H, van der Heijde D. Impact of age, sex, physical function, health-related quality of life, and treatment with adalimumab on work status and work productivity of patients with ankylosing spondylitis. J Rheumatol. 2010 Feb;37(2):385-92.

253. Verstappen SM, Jacobs JW, van der Heijde DM, van der Linden S, Verhoef CM, Bijlsma JW, et al. Utility and direct costs: ankylosing spondylitis compared with rheumatoid arthritis. Ann Rheum Dis. 2007 Jun;66(6):727-31.

254. Wade AG, Crawford GM, Pumford N, Koscielny V, Maycock S, McConnachie A. Baseline characteristics and patient reported outcome data of patients prescribed etanercept: web-based and telephone evaluation. BMC Med Res Methodol. 2011;11:91.

255. Reilly MC, Gooch KL, Wong RL, Kupper H, van der Heijde D. Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. Rheumatology (Oxford). 2010 Apr;49(4):812-9.

256. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. Value Health. 2013 Jun;16(4):686-95.

257. Van der Heijde D, Da Silva J, Dougados M, Geher P, Van der Horst-Bruinsma I, Juanola X, et al. Once-weekly 50-mg dosing of etanercept is as effective as twice-weekly dosing in patients with ankylosing spondylitis. Abstract. Ann Rheum Dis. 2006;65 (Suppl II):509.

258. Boonen A, van der Heijde D, Landewe R, van Tubergen A, Mielants H, Dougados M, et al. How do the EQ-5D, SF-6D and the well-being rating scale compare in patients with ankylosing spondylitis? Ann Rheum Dis. 2007 Jun;66(6):771-7.

259. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997 Nov;35(11):1095-108.

260. Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ. 1998 Mar 7;316(7133):736-41.

261. Dolan P, Kind p, Williams A. A social tariff for EuroQol: results from a UK general population survey. York: Centre for Health Economics, University of York. 1995.
262. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Econ. 1996 Mar-Apr;5(2):141-54.
263. Muszbek N, Punekar Y, Buchanan J, Prasad M, Vioix H, Brown R. Cost-effectiveness of golimumab in ankylosing spondylitis from the UK payer perspective. . Value in Health.13(7):A467.
264. Scottish Medicines Consortium. Golimumab 50 mg solution for injections prefilled pen (auto-injector) or pre-filled syringe (Simponi®). SMC No. (721/11). August 2011.
265. Scottish Medicines Consortium. Adalimumab, 40 mg/0.8 mL, solution for injection (Humira®). SMC No. (858/13). AbbVie Ltd. (previously part of Abbott). 8 March 2013.
266. Tran-Duy A, Boonen A, van de Laar MA, Franke AC, Severens JL. A discrete event modelling framework for simulation of long-term outcomes of sequential treatment strategies for ankylosing spondylitis. Ann Rheum Dis. 2011 Dec;70(12):2111-8.
267. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996 Aug 3;313(7052):275-83.
268. Glintborg B, Østergaard M, Krogh NS. DANBIO registry surveillance in the Danish nationwide necrosis factor: results from 8 years' spondylitis treated with anti-tumour ankylosing continuation in 842 patients with Predictors of treatment response and drug. Ann Rheum Dis. 2010;69:2002-8.
269. European Medicines Agency. Simponi. Summary of product characteristics. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000992/WC500052368.pdf (Accessed August 2014).
270. Arends S, van der Veer E, Kallenberg CG, Brouwer E, Spoorenberg A. Baseline predictors of response to TNF-alpha blocking therapy in ankylosing spondylitis. Curr Opin Rheumatol. 2012 May;24(3):290-8.
271. Dean LE. The Natural History of Ankylosing Spondylitis: Results from the Scotland and Ireland Registry for Ankylosing Spondylitis.
272. Pavy S, Brophy S, Calin A. Establishment of the minimum clinically important difference for the bath ankylosing spondylitis indices: a prospective study. J Rheumatol. 2005;32(1):80-5.
273. Kobelt G, Andlin-Sobocki P, Maksymowych WP. The Cost-Effectiveness of Infliximab (Remicade®) in the Treatment of Ankylosing Spondylitis in Canada. Journal of Rheumatology. 2006;33(4):732-40.
274. Arora A, Mahajan A, Spurden D, Boyd H, Porter D. Long-Term Drug Survival of TNF Inhibitor Therapy in RA Patients: A Systematic Review of European National Drug Registers. Int J Rheumatol. 2013;2013:764518.
275. Nell-Duxneuner V, Schroeder Y, Reichardt B, Bucsics A. The use of TNF-inhibitors in ankylosing spondylitis in Austria from 2007 to 2009 - a retrospective analysis. Int J Clin Pharmacol Ther. 2012 Dec;50(12):867-72.
276. Latimer N. NICE DSU Technical support document 14: survival analysis for economic evaluation alongside clinical trials - extrapolation with patient-level data. Report by the Decision Support Unit. 2013.
277. European Medicines Agency. Cimzia. Summary of product characteristics. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/001037/WC500069763.pdf (Accessed August 2014).
278. Oppong R, Kaambwa B, Nuttall J, Hood K, Smith RD, Coast J. The impact of using different tariffs to value EQ-5D health state descriptions: an example from a study of acute cough/lower respiratory tract infections in seven countries. The European journal of health economics : HEPAC : health economics in prevention and care. 2011 Nov 5.

279. Singh JA, Wells GA, Christensen R, Tanjong Ghogomu E, Maxwell L, Macdonald JK, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. Cochrane Database Syst Rev. 2011(2):CD008794.

280. Roche. RoActemra® (tocilizumab) NICE STA submission. Manufacturer's submission TA198. 2009 (February).

281. Rowen D, Brazier J, Roberts J. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? Health Qual Life Outcomes. 2009;7(27).

282. National Institute for Health and Care Excellence (NICE). Certolizumab pegol (CIMZIA®) for the treatment of Rheumatoid Arthritis, Patient Access Scheme. Available from: http://www.nice.org.uk/guidance/ta186/resources/cimzia-pas-submission2 (Accessed 31/07/14). 2009.

283. National Institute for Health and Care Excellence. Golimumab Patient Access Scheme. October 2009. Available from: http://www.nice.org.uk/guidance/ta225/resources/patient-access-scheme-submitted-by-merck-sharpe-dohme-msd2 (Accessed 31/07/14).

284. Department of Health. National Schedule of Reference Costs 2012-13 for NHS trusts and NHS foundation trusts. 2013.

285. National Institute for Health and Care Excellence (NICE). Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. TA130. 2011.

286. Husain MJ, Brophy S, Cooksey R, Rahman MA, Phillips CJ, Siebert S. The cost of ankylosing spondylitis to the UK NHS. BSR 2014 abstract 211. .

287. Lekander I, Borgstrom F, Svarvar P, Ljung T, Carli C, van Vollenhoven RF. Cost-effectiveness of real-world infliximab use in patients with rheumatoid arthritis in Sweden. Int J Technol Assess Health Care. 2010 Jan;26(1):54-61.

288. Office of National Statistics. National Life Tables, England & Wales 2010-12. 2012.

289. Curtis L. Unit Costs of Health and Social Care 2013. Canterbury: Personal Social Services Research Unit, University of Kent. 2013.

290. Boonen A, Boone C, Albert A, Mielants H. Work productivity in a cohort of employed ankylosing spondylitis patients treated with etanercept. EULAR abstract 2013. Available from: http://www.abstracts2view.com/eular/view.php?nu=EULAR13L_THU0365-A (Accessed June 2014).

291. Zhang W, Bansback N, Boonen A, Young A, Singh A, Anis AH. Validity of the work productivity and activity impairment questionnaire--general health version in patients with rheumatoid arthritis. Arthritis Res Ther. 2010;12(5):R177.

292. Husain MJ, Brophy S, Cooksey R, Rahman MA, Phillips CJ, Siebert S. The work-related costs of ankylosing spondylitis in a UK cohort. BSR 2014 abstract 216.

293. Office for National Statistics. Estimates for England and Wales mid-2012. Available from: http://ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2012/rft---mid-2012-population-estimates-for-england-and-wales.zip (Accessed August 2014).

294. Office for National Statistics. Mid-year population estimates for the UK mid-2012. Office for National Statistics, National Records of Scotland, Northern Ireland Statistics and Research Agency. Available from: http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/2013/stb---mid-2013-uk-population-estimates.html (Accessed August 2014).

295. Office for National Statistics. National Life Tables, England & Wales 2010-12. 2012.

296. National Institute for Health and Care Excellence. Golimumab Patient Access Scheme. http://www.nice.org.uk/nicemedia/live/12076/54406/54406.pdf. 2009.

297. National Institute for Health and Care Excellence (NICE). Certolizumab pegol (CIMZIA®) for the treatment of Rheumatoid Arthritis, Patient Access Scheme. http://www.nice.org.uk/nicemedia/live/11903/45808/45808.pdf. 2009.

298. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Stat Med. 2009 Jun 30;28(14):1861-81.

299. Law A. M. Simulation Modeling and Analysis. 4 ed. Singapore: McGraw-Hill; 2007.