Adalimumab for treating moderate to severe hidradenitis suppurativa: A Single Technology Appraisal

Produced by School of Health and Related Research (ScHARR), The University of Sheffield

Authors Paul Tappenden, Reader in Health Economic Modelling, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Christopher Carroll, Reader in Health Technology Assessment, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

John W Stevens, Reader in Decision Science, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Andrew Rawdin, Research Associate, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Sabine Grimm, Research Associate in Health Economics, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Mark Clowes, Information Specialist, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Eva Kaltenthaler, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

John R Ingram, Senior Lecturer and Consultant Dermatologist, Institute of Infection and Immunity, Cardiff University, Glamorgan House, University Hospital of Wales, Health Park, Cardiff, CF14 4XN

Fiona Collier, Specialist General Practitioner, Forth Valley Dermatology Centre, Stirling Community Hospital, Livilands, Stirling, FK8 2AU

Mohammad Ghazavi, Consultant Dermatologist, Nottingham NHS Treatment Centre, Lister Road, Nottingham, NG7 2FT

Correspondence to Paul Tappenden, Reader in Health Economic Modelling, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

16th December 2015

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 15/06/09.

Declared competing interests of the authors

John Ingram is the Cardiff Principal Investigator for a hidradenitis suppurativa (HS) observational study sponsored by Abbvie. He is also currently leading the British Association of Dermatologists HS guideline development group and was lead author for a Cochrane review of interventions for HS.

Fiona Collier is an employee of Forth Valley NHS, who are involved in Abbvie's study of HS demographics and secondary care usage. She is a co-investigator at the site.

The other authors have no declared competing interests.

Acknowledgements

We would like to thank Rachid Rafia, ScHARR, for providing comments on the draft version of the report. We also thank Gill Rooney for her help in formatting the report.

Rider on responsibility for report

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

This report should be referenced as follows:

Tappenden P, Carroll C, Stevens J, Rawdin A, Grimm S, Clowes M, Kaltenthaler E, Ingram J, Collier F, Ghazavi M. Adalimumab for treating moderate to severe hidradenitis suppurativa: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2015.

Contributions of authors

Christopher Carroll and Eva Kaltenthaler summarised and critiqued the clinical effectiveness data reported within the company's submission. Paul Tappenden, Andrew Rawdin and Sabine Grimm critiqued the health economic analysis submitted by the company. Mark Clowes critiqued the company's search strategies. John Stevens critiqued the statistical analysis contained within the company's submission. John Ingram, Fiona Collier and Mohammad Ghazavi provided clinical advice to the ERG throughout the project. All authors were involved in drafting and commenting on the final report.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BOXES

ABBREVIATIONS

A&E	Accident and Emergency
AAD	American Academy of Dermatology
ADA	Adalimumab
AE	Adverse event
AN	Abscess and Inflammatory Nodule
AWMSG	All Wales Medicines Strategy Group
BMI	Body Mass Index
BNF	British National Formulary
CASP	Critical Appraisal Skills Programme
CCRCT	Cochrane Central Register of Controlled Trials
CEAC	Cost-effectiveness acceptability curve
CI	Confidence Interval
CRD	Centre for Reviews and Dissemination
CRP	C-reactive protein
CS	Company's submission
CSR	Clinical Study Report
DARE	Database of Abstracts of Reviews of Effectiveness
DLQI	Dermatology Life Quality Index
DSA	Determinstic sensitivity analysis
EMA	European Medicines Agency
EMBASE	Excerpta Medica dataBASE
EOW	Every other week
ERG	Evidence Review Group
ESDR	European Society for Dermatological Research
EW	Every week
FDA	Food and Drug Administration
GLM	Generalised logit model
HADS	Hospital Anxiety and Depression Score
HiSCR	Hidradenitis Suppurativa Clinical Response
HRG	Healthcare resource group
HRQoL	Health-related quality of life
HS	Hidradenitis suppurativa
HS-LASI	HS-lesion, activity and severity
HS-PGA	Hidradenitis Suppurativa Physicians' Global Assessment
HSQoL	Hidradenitis Suppurativa Quality of Life
HSSI	Hidradenitis Suppurativa Severity Index
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITT	Intention-to-treat
LOCF	Last observation carried forward
LOR	Loss of response

LS	Least squares
LSCF	Last state carried forward
MCID	Minimal clinically important difference
MDI	Major Depression Inventory
MEDLINE	Medical Literature Analysis and Retrieval System Online
MSS	Modified Sartorius Scale
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NR	Not reported
NRI	Non responder imputation
NRS30	Patient's Global Assessment of Skin Pain
NS	Not significant
NSAID	Non-steroidal anti-inflammatory drug
NYHA	New York Heart Association
OLE	Open-label extension
PAS	Patient Access Scheme
PBO	Placebo
PHQ-9	Patient Health Questionnaire-9
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROMs	Patient Reported Outcome Measures
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
s.c.	Subcutaneous
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SF-36	Short-Form 36
SIGN	Scottish Intercollegiate Guidelines Network
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
TB	Tuberculosis
TEAE	Treatment-emergent adverse event
TNF	Tumour necrosis factor
TSQM	Treatment Satisfaction Questionnaire for Medicine
VAS	Visual Analogue Scale
WCD	World Congress of Dermatology
WOAI	Worsening or absence of improvement
WPAI-SHP	Work Productivity and Activity Impairment-Specific Health Problem
WTP	Willingness-to-pay

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

Hidradenitis suppurativa (HS), also known as acne inversa, is a chronic, inflammatory, recurrent, debilitating, skin follicular disease that usually presents after puberty with painful deep-seated, inflamed lesions. In patients with HS, hair follicles in the apocrine gland-bearing regions (axilla, genital area, groin, infra-mammary region, peri-anal region and buttocks) become blocked and inflamed, resulting in painful recurrent deep-seated boils and nodules. Boils and nodules may progress to abscesses, sinus tracts and scarring. In most patients, disease flares occur at varying intervals, often pre-menstrually in women. Disease flares are characterised by increased pain and suppuration with a foul smelling discharge which stains clothing. Studies have suggested that active disease can have a substantial impairment on a patient's health-related quality of life (HRQoL), exceeding that of other skin diseases that are generally perceived to have a high burden and substantial disability, for example, alopecia, acne, mild to moderate psoriasis, vascular anomalies of the face and atopic dermatitis.

The decision problem required an assessment of the clinical effectiveness and cost-effectiveness of adalimumab compared with established clinical management of active moderate to severe HS in adults whose disease has not responded to conventional systemic HS therapy.

Adalimumab is a recombinant human IgG1 monoclonal antibody expressed in Chinese Hamster Ovary cells. Adalimumab inhibits the activity of the pro-inflammatory cytokine tumour necrosis factor alpha (TNF-α), a key component in the inflammatory process. Adalimumab has a marketing authorisation for the treatment of active moderate to severe HS in adult patients with an inadequate response to conventional systemic HS therapy. Adalimumab also holds a European marketing authorisation for a number of other conditions including rheumatoid arthritis, psoriasis, Crohn's disease and ulcerative colitis. In the management of HS, the recommended adalimumab dose regimen for adult patients with HS is 160mg initially at Day 1 (four 40mg injections in one day or two 40mg injections per day for two consecutive days), followed by 80mg two weeks later at Day 15 (two 40mg injections on the same day). From Day 29 onwards, the recommended dose regimen is 40mg every week (EW). As of December 2015, the NHS indicative price for adalimumab 40mg/0.8ml solution as two pre-filled syringes or auto-injection pens is £704.28.

The population defined in the final NICE scope relates to *"adults with active moderate to severe HS which has not responded to conventional therapy."* This is in line with the marketing authorisation for adalimumab and reflects the populations of the PIONEER I/II studies which form the main basis of

the clinical evidence presented within the company's submission (CS). The health economic model submitted by the company is largely based on evidence relating to the relative efficacy and safety of adalimumab versus placebo within the PIONEER I/II trials. The model also includes additional data on long-term responders to adalimumab 40mg EW who were initially enrolled into the PIONEER I/II trials and who were subsequently enrolled into the M12-555 open-label extension (OLE) study.

The comparator within all three randomised controlled trials (RCTs) was placebo. No head-to-head data are available for adalimumab versus any other therapy. The CS argues that neither surgery nor antibiotics represent relevant comparators for adalimumab. Surgery is argued to be an inappropriate comparator because adalimumab and surgery are not alternative or exclusive treatment choices and because within the PIONEER I/II trials, patients were allowed to undergo surgery to control symptoms (although it is unclear whether this was actually the case). Antibiotics are argued to be inappropriate comparators because they are used alongside adalimumab and because the use of oral antibiotics was allowed in both the intervention and control arms of the PIONEER II trial and as rescue therapy in the PIONEER I trial. The CS also argues that dapsone, retinoids and immunomodulators are not relevant comparators because these are prescribed before adalimumab. According to the company's network meta-analysis (NMA) feasibility assessment, a comparison with infliximab would not be feasible due to evidence limitations and heterogeneity between studies with respect to C-Reactive Protein (CRP) levels and disease severity. As such, the CS argues that the main comparator for the analysis is standard care, as represented by the placebo arms in the PIONEER I/II trials.

The company's clinical review includes data on a large range of outcomes relating to disease severity, clinical response, inflammation and fibrosis, discomfort and pain, adverse events (AEs) and HRQoL. The ERG notes that the primary efficacy endpoint in the PIONEER trials is the Hidradenitis Suppurativa Clinical Response (HSiCR) measure, which was developed by the company.

The CS highlights that there is little research around the treatment of HS, hence the evidence base supporting existing treatment options is limited. The CS also notes that the use of unlicensed treatments exposes patients to potential safety risks and results in variations in clinical practice and inequities with respect to access to effective HS therapies.

End-of-Life criteria were not relevant to this submission and no Patient Access Scheme (PAS) was submitted by the company. The Evidence Review Group (ERG) considers that the evidence presented in the submission was therefore generally consistent with the decision problem.

1.2 Summary of clinical effectiveness evidence submitted by the company

The CS consists of three separate reviews: (1) a review of the clinical efficacy evidence from RCTs of treatments for HS, specifically RCTs comparing adalimumab with placebo; (2) a review of the evidence from an non-controlled OLE study, and; (3) a review of safety evidence from the RCTs of adalimumab versus placebo and the OLE study.

The principal clinical efficacy review included three relevant RCTs comparing adalimumab with placebo in adults with moderate to severe HS: these were a Phase II "dosing" trial, M10-467, and two Phase III trials, PIONEER I and II. The three trials all have two periods: an initial period (weeks 0-12 in the PIONEER I/II trials and weeks 0-16 in the M10-467 trial) comparing adalimumab 40mg EW with placebo, and a second period (weeks 12-36 in the PIONEER trials), initiated by re-randomisation of patients at week 12 to arms of adalimumab 40mg EW, placebo or adalimumab 40mg every other week (EOW, PIONEER trials only). The three RCTs and the OLE study were all found by the company to be at low risk of bias following quality assessment using critical appraisal tools. In the M10-467 trial, significantly more patients in the adalimumab 40mg EW group achieved a clinical response (defined as achieving a HS-PGA score of clear, minimal or mild with at least a 2 grade improvement relative to baseline at week 16) than patients receiving placebo: 17.6% versus 3.9% ($p<0.025$). Significant improvements compared with placebo were also seen at week 16 in individual symptoms, overall disease severity and pain scores with adalimumab 40mg EW.

In PIONEER I and II, significantly more patients in the adalimumab 40mg EW group achieved a clinical response (defined as achieving HiSCR, that is, at least a 50% reduction in the total abscess and inflammatory nodule [AN] count with no increase in abscess count and no increase in draining fistula count relative to baseline at week 12) than patients receiving placebo: 41.8% for adalimumab vs 26.0% for placebo ($p=0.003$) in PIONEER I, and 58.9% for adalimumab vs 27.6% for placebo ($p<0.001$) in PIONEER II.

Significant improvements were also seen in symptoms, disease severity (according to the Modified Sartorius Severity [MSS] score) and pain. All outcomes were significant in PIONEER II. However, in PIONEER I, some of the improvements with adalimumab 40mg EW were numerically but not significantly better than placebo. Subgroup analyses indicated that patients achieved benefit with adalimumab 40mg EW regardless of their baseline characteristics, although some subgroups were subject to small patient numbers. In PIONEER I and II, adalimumab 40mg EW significantly improved quality of life as measured by the EQ-5D, the physical components of the Short-Form 36 (SF-36), and the Dermatology Life Quality Index (DLQI) compared with placebo, but the improvements were not significant across some other components of SF-36. The treatment effect varied between the trials. This might be explained in part by differences in patient demographics and

study design between trials. The company is conducting ongoing analyses of the data from the PIONEER trials and the OLE study to understand these differences. The CS did not include a pairwise meta-analysis of the PIONEER I/II trials. An NMA was not considered feasible.

Some improvements were maintained into the second period of the PIONEER trials up to 36 weeks. The company stated that re-randomisation at week 12, at the beginning of this second period (Period B), and protocol-driven discontinuation during Period B for patients with Loss of Response (LOR) or Worsening or Absence of Improvement (WOAI), accounted for low patient numbers in the group receiving adalimumab 40mg EW for the total study duration (n=21 in PIONEER I and n=20 in PIONEER II). In the second period, there was a loss of effect for patients re-randomised to placebo or adalimumab 40mg EOW. Outcomes were maintained in patients who went on to enter the M12-555 OLE study.

The review of the safety evidence included the three key RCTs and the single OLE cohort study (M12-555 OLE). Adalimumab 40mg EW was well-tolerated in all three RCTs. The proportion of patients experiencing serious adverse events (SAEs) or discontinuing treatment attributable to AEs was low and similar in both the adalimumab and placebo arms. In an integrated summary of PIONEER I and II (n=633), six patients receiving placebo (1.9%) and three receiving adalimumab 40mg EW (0.9%) gave AEs as their primary reason for discontinuation during Period A. The most common AEs were exacerbation of HS, nasopharyngitis and headache. Rates of infectious AEs were similar for patients receiving adalimumab and for those receiving placebo. The CS states that the M12-555 OLE is the only ongoing study of adalimumab in this indication. Final data from this study are expected to be available in 2016.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The principal efficacy review is a poorly-reported systematic review of relevant RCTs (M10-467 and PIONEER I and II). The trials are generally consistent with the final NICE scope. The primary efficacy outcome was clinical response, principally measured using the HiSCR measured developed by the company. Clinical advice received by the ERG confirms that the HiSCR measure has been validated but, in terms of clinical decision-making, its findings must be viewed alongside the results of patient-reported outcome measures, in particular quality of life assessed by the DLQI and a pain measure. In the trials, secondary outcomes included assessments of disease severity and symptoms, using the MSS score and AN counts, pain and quality of life (various measures).

The ERG considers the M10-467 trial to be at low risk of bias across all domains for the relevant Period 1 (up to week 16). The ERG also considers the results from Period A (i.e. up to week 12) in PIONEER I and II to be generally at low risk of bias: only the domains of attrition and reporting have

a low-to-moderate risk of bias. However, the ERG considers there to be a moderate or unclear risk of selection and attrition bias affecting the results of Period B in the PIONEER trials. There is also a low-to-moderate risk of reporting bias in Period B in the two trials. It should also be noted that whilst M10-467 has been published, the PIONEER trials have not.

Across all three RCTs, the percentage of patients achieving clinical response according to the HiSCR measure on adalimumab 40mg EW compared with placebo at week 12 or week 16 was significantly higher than in the placebo groups ($p<0.01$), although the treatment effect varied between the trials. In addition, significant or clinically relevant differences in favour of adalimumab 40mg EW that were reported for secondary outcomes in PIONEER II were not always found for those outcomes in PIONEER I, especially for AN count, MSS score, pain and some components of quality of life measured by the SF-36. An arm-based integrated summary, which breaks randomisation, was conducted for the two PIONEER trials to tabulate Period B response (for all patients and for a group of HiSCR "responders" and "partial responders"). This "partial responder" group (defined as HiSCR responders with ≥25% reduction rather than ≥50% reduction) are a *post hoc* analysis group. This group was not defined in protocols or published descriptions of study design or pre-specified analysis methods for the PIONEER trials. It was also not considered in the published validation study for the HiSCR measure, nor was it justified or explained in the company's clinical review. According to this analysis, improvements in response were maintained or reduced in this second period. A small number of secondary outcomes were reported for Period B of PIONEER I and II, but only for patients who had had a clinical response at week 12. The results were based on analyses with small sample sizes (range of 15 to 22 patients across all outcomes for both PIONEER trials).

These trials were supplemented by a single, unpublished, non-randomised, non-controlled, unblinded cohort study, which was an OLE study of the PIONEER trials (M12-555 OLE). In terms of efficacy, the results suggested [redacted]. Details of the results for secondary outcomes such as MSS and NRS30 were not reported. The ERG considers these efficacy results to be subject to uncertainty because they are drawn from interim analyses of unpublished study data. The study also only potentially offers efficacy data for up to 72 weeks for a drug that might be taken for many years by patients with moderate to severe HS.

The submission of safety evidence was a review of the three generally good quality RCTs, supplemented by the single arm cohort study. There were no obvious safety concerns, with most AEs being balanced across adalimumab 40mg EW and placebo trial arms, and small numbers of SAEs.

Longer-term data are required to determine whether reported AE rates are maintained for patients on long-term maintenance doses of adalimumab 40mg EW; whether or not certain subgroups of patients are at a higher risk of certain events; and to confirm whether or not there are any differences between the interrupted and uninterrupted regimens.

1.4 Summary of cost-effectiveness submitted evidence by the company

The CS includes a systematic review of economic evaluations of treatments for HS together with a *de novo* model-based economic evaluation of adalimumab versus standard care in adult patients with an inadequate response to conventional systemic HS therapy.

The company's systematic review of existing economic evaluations did not identify any relevant studies for inclusion.

The company's *de novo* economic model adopts a Markov approach to estimate costs and health outcomes for adalimumab and standard care from the perspective of the NHS and Personal Social Services (PSS) over a lifetime horizon. All analyses presented in the CS relate to the full population specified in the marketing authorisation for adalimumab; no subgroup analyses are presented within the CS. The company's model includes five mutually exclusive health states, based on depth of HiSCR response: (i) high response; (ii) response; (iii) partial response; (iv) no response, and; (v) dead. The model uses a 2-week cycle length for the first 2 cycles, and a 4-week cycle length thereafter. Health state transitions are modelled up to week 36 using data from PIONEER I/II, including a discontinuation rule for patients who do not achieve at least a partial response by week 12. The long-term HiSCR trajectory of adalimumab responders (including partial responders) beyond 36 weeks is subsequently modelled using a time-invariant generalised logit model (GLM) fitted to last observation carried forward (LOCF)-imputed data from the M12-555 OLE study. The long-term HiSCR trajectories for patients receiving standard care and for those who have previously discontinued adalimumab beyond 36 weeks are modelled using separate time-invariant GLMs fitted to data from weeks 12-36 from the PIONEER I/II trials. Health utilities are modelled according to depth of HiSCR response using a *post hoc* analysis of EQ-5D data collected within PIONEER II. Resource use estimates, which were again differentiated by depth of HiSCR response, were based on a survey of UK physicians and were assumed to include inpatient visits due to HS surgery, outpatient visits due to HS surgery, visits to wound care due to HS surgery, non-surgical inpatient visits, non-surgical outpatient visits, visits to wound care not due to HS surgery, Accident and Emergency (A&E) visits and costs associated with AEs. Unit costs were taken from the British National Formulary (BNF), the Personal Social Services Research Unit (PSSRU) and NHS Reference Costs. AEs are not assumed to have an additional impact on HRQoL.

Based on the probabilistic version of the company's base case model, adalimumab is expected to produce an additional 1.02 quality-adjusted life years (QALYs) at an additional cost of £[redacted] as compared with standard care; the probabilistic incremental cost-effectiveness ratio (ICER) for adalimumab versus standard care is expected to be £[redacted] per QALY gained. The results of the deterministic model are similar, with adalimumab yielding an ICER of £[redacted] per QALY gained compared with standard care. The company's probabilistic sensitivity analysis (PSA) suggests that assuming a willingness-to-pay (WTP) threshold of £20,000 per QALY gained, the probability that adalimumab produces more net benefit than standard care is approximately [redacted]. Assuming a WTP threshold of £30,000 per QALY gained, the probability that adalimumab produces more net benefit than standard care is approximately [redacted]. Within the company's deterministic scenario analysis, the ICER for adalimumab was greater than £30,000 per QALY gained in four scenarios: (i) when the time horizon was truncated to 20 years; (ii) when the model was based only on data from PIONEER II; (iii) when the last state carried forward (LSCF) imputation rule was used, and; (iv) when the discontinuation rate for adalimumab non-responders after week 36 was based on the OLE study.

[redacted]

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's model. The ERG's critical appraisal identified a number of issues relating to the company's model and analysis. The most pertinent of these relate to: (i) the use of a model structure in which health gains and treatment continuation rules are defined according to depth of response, which does not reflect the pre-planned and validated HiSCR endpoint used in the PIONEER trials; (ii) the likely overestimation of the lifetime costs of HS surgery predicted by the company's model; (iii) the incorrect implementation of a continuation rule for adalimumab non-responders which does not mathematically reflect the actual assumptions stated in the CS; (iv) the use of arm-based aggregate data from the PIONEER I/II trials rather than a meta-analysis of relative treatment effects, and; (v) uncertainty surrounding the long-term transition probabilities derived from the PIONEER I/II trials and the M12-555 OLE study.

The ERG undertook eight sets of exploratory analyses based on the company's submitted model. The first three of these analyses represent the ERG's base case analysis. These include: (i) correction of technical programming errors in the company's model; (ii) applying structural amendments to the model to correctly reflect the company's intended adalimumab non-responder continuation rule during the maintenance phase; (iii) re-estimation of the costs of HS surgery. Further analyses were also undertaken to explore uncertainty surrounding the transition probabilities employed in the model, the likely impact of discontinuing non-responders and partial responders to adalimumab (during the induction phase only) and the potential structural uncertainty around the company's adopted

modelling approach. The latter two analyses could not however be fully implemented due to the limitations of the company's model structure.

The ERG's exploratory analyses indicate that the technical programming errors have only a minor impact on the model results and lead to a small increase in the ICER for adalimumab versus standard care. The incorporation of tunnel states for adalimumab non-responders within the maintenance phase of the corrected model increases the ICER for adalimumab versus standard care more substantially (ICER=£[redacted] per QALY gained). The ERG's base case, which comprises a scenario whereby these two sets of corrections are combined with a lower cost of HS surgery, results in an estimated deterministic ICER for adalimumab versus standard care of £[redacted] per QALY gained. The probabilistic ICER for this analysis is slightly higher (£[redacted] per QALY gained). The ERG's base case ICER for adalimumab versus standard care is markedly less favourable than that presented within the CS.

1.6 ERG commentary on the robustness of evidence submitted by the company

The ERG considers the RCT evidence to be robust for the initial trial periods up to 12 or 16 weeks: it generally satisfied the requirements of the decision problem, with some minor exceptions, and was good quality. The treatment effect did vary between studies, which might be explained by differences in patient characteristics and study design between trials. The efficacy results from the second period of the PIONEER trials are at a higher risk of bias across some domains, and are affected by the merging of "responders" with "partial responders", the latter being a *post hoc* analysis group which is neither justified nor explained in the submission. The safety evidence is generally at low risk of bias but is limited, and several questions remain over AE rates for patients on "continuous" or long-term adalimumab 40mg EW.

The ERG has concerns regarding the company's implemented model structure, in particular, the incorrect implementation of the adalimumab non-responder discontinuation rule during the maintenance phase and the definition of health states and treatment continuation rules based on depth of HiSCR response rather than the ≥50% AN reduction threshold. In addition, the cost savings due to HS surgery avoided predicted by the company's model are likely to be overestimated. The ERG has further concerns regarding the use of arm-based summaries to aggregate data from the PIONEER I/II trials and the uncertainty surrounding the long-term transition probabilities used to inform the model.

1.6.1 Strengths

The ERG recognises that the submission included three RCTs at low risk of bias for the initial study period (up to 12 weeks for the PIONEER trials and 16 weeks for M10-467) comparing the study drug, adalimumab at its licensed dose, with placebo. All of the required outcomes were assessed and

reported: clinical response, , as well as disease severity, symptoms, pain and quality of life. The ERG considers the efficacy results for up to 12 weeks and the safety data for up to 36 weeks to be at a low risk of bias.

1.6.2 Weaknesses and areas of uncertainty

The ERG noted that the principal areas of uncertainty in the clinical evidence related to potential treatment effect modifiers and the short follow-up. These uncertainties exist due to observed differences in certain outcomes or level of outcome between trials, differences in disease severity and other baseline characteristics between trials, and the amount of missing data and imputed results beyond 12 weeks in the PIONEER I/II trials and the OLE study. In addition, the ERG notes that there is uncertainty with respect to whether the achievement of a partial HiSCR represents a clinically meaningful treatment benefit sufficient to warrant continuing adalimumab, and around the expected impact of adalimumab on the use of other healthcare resources (for example surgery and other pharmacological treatments used to manage HS).

2 BACKGROUND

2.1 Critique of company's description of underlying health problem

Hidradenitis suppurativa (HS), also known as acne inversa, is a chronic, inflammatory, recurrent, debilitating, skin follicular disease that usually presents after puberty with painful deep-seated, inflamed lesions. In patients with HS, hair follicles in the apocrine gland-bearing regions (axilla, genital area, groin, infra-mammary region, peri-anal region and buttocks) become blocked and inflamed resulting in painful recurrent deep-seated boils or nodules. Boils and nodules may progress to abscesses, sinus tracts and scarring. In most patients, disease flares occur at varying intervals, often pre-menstrually in women. Disease flares are characterised by increased pain and suppuration with a foul smelling discharge which stains clothing.[1-3]

Several risk factors probably contribute to HS, including smoking, obesity genetic predisposition and endocrine influences. HS affects young adults, with disease onset typically between the second and fourth decades of life.[4, 5] It is likely that HS is a progressive disease, with some patients reporting a progression from Hurley Stage I to II to III over time; the risk factors that predispose patients to progression include smoking and obesity.[6, 7]

The prevalence of HS is not precisely known, but a number of estimates are available in the literature. A prevalence of 1% in the adult Eurpoean population has been reported in several studies,[2] although actual rates are likely to be higher due to problems of under-recognition.[1, 3] There are no published data on prevalence rates in the UK, although it has been suggested that this might be in the region of 1 in 600.[4] HS has higher prevalence in women than men and around one-third of patients have a disease in first-degree relatives.[2] The other important known risk factors for HS are obesity and cigarette smoking.[1-3]

The pathogenesis of HS is largely unknown and it is defined by its clinical features and its chronicity.[2] Diagnosis relies on the presence of: (1): typical lesions, i.e. deep-seated painful nodules: 'blind boils' in early lesions, abscesses, draining sinus, and bridged scars; (2) typical topography, i.e. axillae, groin, perineal and perianal region, buttocks, infra- and inter-mammary folds, and; (3) chronicity and recurrences. These three criteria must be met to establish a diagnosis of HS. The population referred to in the final NICE scope[8] relates to patients with active moderate to severe HS who have failed prior systemic therapy. The CS[9] provides a description of HS in accordance with the terminology used in the NICE scope.

HS is classified according to the Hurley staging system, as shown in Table 1.

Table 1: Hurley's classification[10]

Stage	Clinical features
Grade I	Abscess formation, single or multiple without sinus tracts and cicatrisation
Grade II	Recurrent abscesses with tract formation and cicatrisation. Single or multiple, widely separated lesions
Grade III	Diffuse or near-diffuse involvement, or multiple interconnected tracts and abscesses across entire area

Hurley's grades are used to classify each disease location in an individual, such as armpit, groin etc. in a disease severity category.[2] It has been suggested that the Hurley classification is a useful guide for baseline severity and for helping to select appropriate treatment options, but that the MSS scoring system offers a more precise means of detailing severity of disease in the context of evaluating improvement.[2] This score has not been formally validated but is frequently used and has been shown to be highly correlated with Hurley's classification, as well as degree of suppuration, which are good markers of inflammation and burden of the disease.[2]

HS is associated with malodorous discharge that stains clothing and is therefore accompanied by embarrassment, disabling social stigma, low self-worth and impacts on interpersonal relationships. Studies have found that active disease can have a substantial impairment on a patient's HRQoL, exceeding that of other skin diseases that are generally perceived to have a high burden and substantial disability, for example, alopecia, acne, mild to moderate psoriasis, vascular anomalies of the face and atopic dermatitis. Given the debilitating impact of HS on a person's life, measures of pain and quality of life, especially the DLQI, are recognised as being useful in the clinical management of HS.[2,3,11]

2.2 Critique of manufacturer's overview of current service provision

The ERG and clinical advisors considered the company's description of current service provision for the treatment of populations with HS to be appropriate and relevant to the decision problem (see CS,[12] pages 29-31 and pages 39-43) and that the recommendations of relevant clinical guidelines have been taken into account.[13]

The CS, literature, guidelines and clinical advice received by the ERG, all indicate that there is no current standard of care for HS in the UK, but that treatment is determined by the specifics of the disease in the individual patient, together with clinical and patient experience. The aim of treatment is usually to control the disease and to reduce the number of outbreaks. Total cure is generally not expected. In addition to lifestyle changes (smoking cessation and weight loss), therapeutic options include topical antiseptics and antibiotics, systemic antibiotics (e.g. oral tetracyclines, clindamycin and rifampicin), antiandrogens, systemic retinoids, immunomodulatory agents, laser treatment,

surgery and anti-TNF-α therapies.[13-15] The choice of therapy typically depends on frequency, severity and spread of lesions and also gender in the case of the retinoid acitretin.

Topical antimicrobials are recommended for Hurley Stage I local disease, whilst systemic antibiotics are typically used for widespread or severe disease. Medical therapy is generally recommended for multiple, widely-spread lesions, and surgery for stable, locally-recurring lesions or severe and advanced disease. There is currently no known effective monotherapy, as confirmed by recent reviews,[13, 14, 16] hence a combination of different treatment modalities is often applied. Clinical advice received by the ERG suggests that surgery is usually an option after the failure of medical treatments and might involve simple local incision and drainage (usually as a response to acute flares, rather than to control the disease or reduce recurrence); narrow margin excision (which might see recurrence at the edge of the excised area) and wide margin excision for patients with advanced disease.[15] All of these interventions are mentioned as possible therapies or potential comparators in the CS.[9]

A survey of current UK practice among dermatologists confirmed that, after topical treatments, oral antibiotics, such as lymecycline or doxycycline, represent the first-line medical treatment of choice, followed by clindamycin and rifampicin, dapsone, acitretin, ciclosporin, depending on response and gender.[15] TNF-α inhibitors, such as etanercept, infliximab and adalimumab are already being used in the treatment of patients with moderate to severe HS, especially infliximab as the dose can be adjusted according to patient weight.[15] The CS states that adalimumab would typically be used after the failure of antibiotic therapy and before other therapies such as dapsone (antibiotic), retinoids and immunomodulators (ciclosporin) or surgery (see CS,[9] page 42). However, in response to a request for clarification from the ERG, the company's initial proposed positioning of adalimumab was amended to a position *"after all effective conventional systemic HS treatments have been exhausted"* and *"before or after surgery"* (see clarification response,[17] question C1). Clinical advisors to the ERG agreed that this was an appropriate place in the pathway. The number of patients who are likely to be suitable or eligible for treatment with adalimumab is unclear (see clarification response,[17] question C2). Adalimumab would only be prescribed in secondary care (see CS,[9] page 28). It is administered via subcutaneous (s.c.) injection, but clinical advice received by the ERG confirmed that initial and ongoing patient training would not be required from secondary care services because this support was to be provided by AbbVie Care (see CS,[9] page 27).

3. CRITIQUE OF THE COMPANY'S DEFINITION OF THE DECISION PROBLEM

This chapter presents a summary and critique of the decision problem addressed by the CS.[9] A summary of the decision problem as outlined in the final NICE scope[8] and addressed in the CS[9] is presented in Table 2.

Table 2: Company's statement of the decision problem (adapted from CS[9] page 14)

Element	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope
Population	Adults with active moderate to severe HS which has not responded to conventional therapy	Adults with active moderate to severe HS which has not responded to conventional therapy	As specified in the scope
Intervention	Adalimumab	Adalimumab	As specified in the scope
Comparator(s)	Established clinical management without adalimumab	Where the data allows AbbVie has performed comparisons in line with the licence	As per scope where data allows
Outcomes	The outcome measures to be considered include: • Disease severity • Clinical response • Inflammation and fibrosis • Discomfort and pain • AEs of treatment • HRQOL	The outcome measures to be considered include: • Disease severity • Clinical response • Inflammation and fibrosis • Discomfort and pain • AEs of treatment • HRQOL	As specified in the scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year (QALY). The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services (PSS) perspective.	• Cost-effectiveness will be presented as incremental cost per QALY. • The time horizon for the modelling is a lifetime. • Costs will be considered from an NHS and PSS perspective.	As specified in the scope
Subgroups to be considered	None stated	None stated	As specified in the scope
Special considerations including issues related to equity or equality	None stated		

3.1 Population

The population defined in the final NICE scope[8] relates to *"adults with active moderate to severe HS which has not responded to conventional therapy."* This is in line with the marketing authorisation for adalimumab and reflects the populations of the PIONEER I/II studies[18, 19] which form the main basis of the clinical evidence presented within the CS.[9] The health economic model submitted by the company is largely based on evidence relating to the relative efficacy and safety of adalimumab versus placebo within the PIONEER I/II trials. The model also employs additional data on long-term responders to adalimumab 40mg EW who were initially enrolled into the PIONEER I/II trials[18, 19] who were subsequently enrolled into the M12-555 OLE study.[20]

3.2 Intervention

The intervention defined in the CS is adalimumab 40mg EW administered via subcutaneous (s.c.) injection. Adalimumab is available as either as an auto-injection pen or pre-filled syringe (40mg/0.8ml solution).

Adalimumab is a recombinant human IgG1 monoclonal antibody expressed in Chinese Hamster Ovary cells.[12] Adalimumab inhibits the activity of the pro-inflammatory cytokine TNF-α, a key component in the inflammatory process. Adalimumab binds specifically to TNF-α and blocks its interaction with TNF receptors 1 and 2.

Adalimumab has a marketing authorisation for the treatment of active moderate to severe HS in adult patients with an inadequate response to conventional systemic HS therapy.[12] Adalimumab also holds a European marketing authorisation for the treatment of juvenile idiopathic arthritis, rheumatoid arthritis, axial spondyloarthritis, psoriatic arthritis, psoriasis, paediatric plaque psoriasis, Crohn's disease and ulcerative colitis.

As of December 2015, the NHS indicative price for adalimumab is £704.28. Each pack contains two syringes.

The Summary of Product Characteristics (SmPC)[12] states that the recommended adalimumab dose regimen for adult patients with HS is 160mg initially at Day 1 (four 40mg injections in one day or two 40mg injections per day for two consecutive days), followed by 80mg two weeks later at Day 15 (two 40mg injections on the same day). From Day 29 onwards, the recommended dose regimen is 40mg EW. The SmPC notes that antibiotics may be continued during treatment with adalimumab if necessary and that patients should use a topical antiseptic wash on their HS lesions on a daily basis during treatment with adalimumab. The SmPC advises that continued therapy beyond 12 weeks should be carefully reconsidered in a patient with no improvement within this time period. Should

treatment be interrupted, adalimumab 40mg EW may be re-introduced. The SmPC also notes that the benefits and risks of continued long-term treatment should be periodically evaluated.[12]

According to the European Medicines Agency (EMA), treatment with adalimumab should be initiated and supervised by specialist physicians experienced in the diagnosis and treatment of conditions for which adalimumab is indicated and patients treated with adalimumab should be given the special alert card. The SmPC notes that patients require training in injecting after which time patients might self-inject with adalimumab if their physician determines that it is appropriate and with medical follow-up as necessary. The CS[9] states that adalimumab will be administered in the home setting via AbbVie Care (the company's home care service). During treatment with adalimumab, other concomitant therapies should be optimised.

The SmPC[12] notes that the safety and efficacy of adalimumab in children aged 12-17 years with HS have not yet been established and that no data are available. There is no relevant use of adalimumab in children aged below 12 years in this indication.

Contraindications to adalimumab treatment include hypersensitivity to the active substance, the presence of active tuberculosis (TB) or other severe infections such as sepsis, and opportunistic infections, and moderate to severe heart failure (NYHA class III/IV). The administration of adalimumab during pregnancy is not recommended.

3.3 Comparators

Within the clinical section of the company's review, all RCT evidence for adalimumab is drawn from trials which included a placebo control. Within the company's model, the comparator is defined as "standard care"; this is assumed to include surgery and non-surgery related hospital visits and A&E attendances. Whilst the company considered the feasibility and appropriateness of undertaking NMAs for various outcomes, these were not performed for any outcome and the CS does not include any comparison of adalimumab against any specific pharmacological or surgical comparator.

With reference to the decision problem, the CS states that, *"where the data allows AbbVie has performed comparisons in line with the licence."* The relevance of this statement is unclear however as the licence relates to adalimumab rather than any selected comparator. Further, whilst the company argues that there is no effective licensed or NICE-recommended treatment for HS, this does not preclude the consideration of such therapies as potentially relevant comparators to adalimumab.[21]

With respect to currently used therapies for HS, the CS notes that within a survey of 142 patients from 10 UK hospitals funded by the company, patients took an average of 10 medications within the 5-year

retrospective period (range 1-43 medications). The CS[9] also highlights that there are no licensed therapies for the treatment of HS in the UK and that various pharmacological therapies are used off-label (including antiseptics, non-steroidal anti-inflammatory drugs [NSAIDs], immunosuppressants, corticosteroids, anti-androgens, retinoids and TNF-α inhibitors). The CS also notes that there is limited robust evidence to demonstrate the efficacy of any of these therapies in the management of HS.

The CS argues that neither surgery nor antibiotics represent relevant comparators for adalimumab. Surgery is argued to be an inappropriate comparator since adalimumab and surgery are not alternative or exclusive treatment choices and because, within the PIONEER I/II trials,[18,19] patients were allowed to undergo surgery such as incision and drainage to control symptoms (although it is unclear whether this was actually the case – see Section 4.2.1 and Section 5.3). Antibiotics are argued to be inappropriate comparators because they are used alongside adalimumab and because the use of oral antibiotics was allowed in both the intervention and control arms of the PIONEER II trial[19] and as rescue therapy in the PIONEER I trial.[18] The CS also argues that dapsone, retinoids and immunomodulators are not relevant comparators for adalimumab because these are prescribed before adalimumab. According to the company's NMA feasibility assessment,[9] a comparison with infliximab would not be feasible due to evidence limitations and heterogeneity between studies. As such, the CS argues that the main comparator for the analysis is standard care, as represented by the placebo arms in the PIONEER I/II trials.[18,19] Issues surrounding the implementation of this economic comparison is discussed in further detail in Section 5.3.

Clinical advisors to the ERG agree that there are few obvious comparators for adalimumab and that standard care, as defined within the company's model, represents a reasonable comparator. One clinical advisor to the ERG did however note that infliximab and adalimumab are typically used interchangeably, with the choice of treatment often being guided mainly by cost concerns. Whilst the ERG agrees that an indirect comparison based on the HiSCR measure would not be possible for adalimumab versus infliximab, it may have been possible to compare the two treatments using an alternative clinical outcome measure, such as pain. This would however have required a very different model structure to that presented within the CS.

3.4 Outcomes

The company's clinical review includes evidence relating to the following outcomes:

- Clinical response, measured by the HS-PGA or the HiSCR measures, which assess clinical improvement following pre-specified thresholds for reduction or maintenance in the number of a patient's lesions, abscesses, inflammatory nodules and draining fistulae;

- Disease severity and inflammation and fibrosis, which are also assessed by counts of lesions, abscesses, inflammatory nodules and draining fistulae using measures such as the HiSCR, and the MSS and Hurley scores;
- Discomfort and pain, which are measured by specific dermatology and generic pain scores;
- Any AE of treatment, including serious AEs, in particular those which led to discontinuation of the study drug, as well as common AEs such as headache and nasophryngitis, or serious infections associated with adalimumab, such as TB;
- HRQOL, assessed by specific dermatology quality of life measures (e.g. Hidradenitis Suppurativa Quality of Life [HSQOL] and the DLQI) as well as more general measures, such as the SF-36 and the EQ-5D.

3.5 Economic analysis

The CS includes the methods and results of a *de novo* model-based health economic analysis to assess the incremental cost-utility of adalimumab versus standard care for the treatment of adults with active moderate to severe HS which has not responded to conventional therapy. The company's model is detailed and critiqued in Chapter 5. The ERG notes that whilst the efficacy and safety data used within the model are based on the PIONEER I/II trials[18,19] and the OLE study,[20] the resource costs associated with the comparator are instead drawn from a survey on UK physicians funded by the company.[22]

3.6 Subgroups

Within the company's review of clinical effectiveness evidence (see CS,[9] Chapter 4), pre-specified subgroup analyses were undertaken for all three adalimumab RCTs in order to assess the consistency of the primary efficacy endpoint by demographic and baseline characteristics. *Post hoc* analyses were also undertaken in the dose-finding trial (M10-467) in order to compare the clinical response for patients in the adalimumab 40mg EW group compared with those in the placebo group. No specific subgroups are considered within the company's health economic analysis.

3.7 Special considerations

The CS[9] notes that currently no therapies have been approved for the treatment of HS in England and that various therapeutic options are used off-label in clinical practice. The CS highlights that there is little research around the treatment of HS, hence the evidence base supporting existing treatment options is limited. The CS states that the use of unlicensed treatments exposes patients to potential safety risks and also results in variations in clinical practice and inequities in the access to effective HS therapies.

A confidential PAS has not been submitted by the company. End-of-Life criteria are not applicable to this appraisal.

4. CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the reviews submitted by the company on the efficacy and safety of adalimumab in adults with moderate to severe HS. The critique was performed following the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and checklist.[23]

4.1 Critique of the methods of review(s)

The CS[9] reports the methods and results of three separate reviews:

(i) A review of the efficacy evidence from RCTs (see CS, Sections 4.1-4.10);

(ii) A review of the efficacy and safety evidence from non-randomised and non-controlled studies (see CS, Section 4.11), and;

(iii) A review of safety evidence from RCTs and a non-randomised study (see CS, Section 4.12).

Each review applies different inclusion criteria depending on the intended analysis and the included study designs.

The main review of efficacy evidence from RCTs was a poorly-reported systematic review. Following a request for clarification from the ERG regarding certain process elements adopted by the company, the ERG considered the review to be generally sound (see clarification response,[17] questions A1-A7).

The review of the efficacy evidence from non-randomised and non-controlled studies was limited to a single open-label, non-controlled extension study (M12-555 OLE). This review was not considered to be a systematic review because it was unclear how the evidence was identified, selected and extracted, no inclusion or exclusion criteria were provided, and a list of excluded studies was not reported. Quality assessment of the OLE study was performed by the company using a checklist, but the choice and origin of this was neither justified nor specified. This was clarified by the company in response to a request by the ERG (see clarification response,[17] questions A22).

The review of the safety evidence was also not considered by the ERG to be a systematic review because it was unclear from the original submission how the included non-RCT evidence was identified and selected, no detailed inclusion or exclusion criteria or details of data extraction were provided, and a list of potentially relevant excluded studies was not reported.

4.1.1 Searches

The company conducted a systematic literature review search for evidence on the comparative efficacy and safety of interventions in HS. The ERG notes that, since the searches focussed on treatment of the condition (HS) rather than the specific intervention under review (adalimumab),

studies describing AEs where the drug was used for other conditions would not have been retrieved. Studies were identified by a literature search of MEDLINE, EMBASE and the Cochrane CENTRAL register of clinical trials. Whilst these are the key sources identified by the Cochrane Handbook,[24] many STAs go beyond this and include additional sources in order to increase the coverage of the search and to ensure that all potentially relevant evidence has been taken into account. The CS also reports an additional search of the US National Institutes of Health Ongoing Trials Register (clinicaltrials.gov), but no searches of the equivalent WHO or EU registers (http://www.who.int/ictrp/search/en/ and https://www.clinicaltrialsregister.eu/ respectively) were reported. In addition, standard supplementary methods such as reference tracking were not used.

The ERG queried the interface on which the MEDLINE and EMBASE searches were conducted as there were significant logical errors in the searches as reported (for example, the omission of brackets, without which the search would not function or produce the number of results reported). For example, Line 17 of the EMBASE search strategy (see CS,[9] Appendix 2, Table A1) was written as *"observational adj3 study or studies or design or analysis or analyses.mp"* should read *"observational adj3 (study or studies or design or analysis or analyses).mp"*. Without these brackets, the query would be interpreted as: "*(observational adj3 study) or studies or design or analysis or analyses.mp."* This search would additionally find many instances of the "study" terms occurring without "observational" in proximity. There are also similar problems in lines 17-19 of the EMBASE search and the combinations at lines 22 and 24. The MEDLINE search, which follows a similar structure, also contains the same errors.

When attempting to reproduce and verify the company's searches, the ERG found that, after correcting the syntax, the numbers of results retrieved suggested that the errors had not been made in the company's live search, rather they were present only in the reported version. Whilst this raises some concerns about the accuracy of the reporting, it appears that correct syntax was used in the search itself.

In addition to the syntax errors, the ERG noticed several typographical errors and/or spelling mistakes in the searches, which appear from the very first line of the EMBASE search: *"exp hidradenitis suppurative/"* (the correct heading is *"exp hidradenitis suppurativa/"*). Line 2 also contains a spelling error (*"hidradenitis supportiva"*) and a variant spelling of the archaic term *"pyodermia fistulans significa"* is sometimes found: *"pyodermia fistulans sinifica"*. Another search term which has been omitted is the reversed form of the name for the condition, *"suppurative hidradenitis"*, which the ERG found to be in relatively common use.

Line 19 of the company's EMBASE search includes the phrase: *not "randomized controlled trial".pt.* Unfortunately, "randomized controlled trial" is not a valid publication type in EMBASE, and is present only in MEDLINE, so this clause of the search string will not have any effect.

The MEDLINE search shares many of the above errors and omissions.

The ERG found that, despite these errors, the numbers of results retrieved by the company were in accordance with the results obtained when all terms were entered correctly by the ERG. It would appear that the search strategies have been re-typed for the CS rather than providing a screenshot or a copy-pasted version of the search (as is the convention), and that errors were made during this transcription process. The ERG notes that it is difficult accurately to assess the robustness of searches if they are not presented in a reproducible form. However, after correcting the various errors described above, the ERG found no additional studies over and above those identified by the company.

4.1.2 Inclusion criteria

The inclusion criteria for the reviews are described in Section 4.1.1 of the CS (pages 46-47, see Table 3). These criteria describe RCTs measuring the efficacy and safety of a number of biologic (including adalimumab), antibiotic, steroid, retinoid and surgical interventions (limited to laser only) compared with any of these interventions or placebo in adult patients with HS. These are the inclusion criteria for the potential performance of an indirect comparison, which is discussed in Section 4.10 of the CS[9] (page 99). However, Sections 4.2 to 4.9 of the CS (pages 54-99) report a clinical efficacy review of a subset of studies satisfying the following inclusion criteria: RCTs measuring the efficacy and safety of adalimumab compared with other interventions or placebo in adult HS patients. The four RCTs satisfying these criteria are: M10-467,[25] PIONEER I,[20] PIONEER II,[19] and Miller *et al.*[26] The RCTs include two doses of adalimumab, 40mg EW and 40mg EOW. One study (Miller *et al*[26]) evaluated only an unlicensed dose of adalimumab in the HS population and was therefore correctly excluded using additional inclusion/exclusion criteria described later in Section 4.2 of the CS. It is unclear why the definition of the surgical comparator in the review was restricted to laser treatment only.

The review of the efficacy evidence from non-randomised and non-controlled studies did not specify any inclusion criteria (see Section 4.11). This review reported a single open-label, non-controlled extension study, M12-555 OLE,[20] whose participants were recruited from the PIONEER I and II trials. According to the inclusion criteria outlined in Section 4.1.1 of the CS and the searches described in Appendix 2 of the CS, non-randomised studies were explicitly excluded (an RCT study filter is applied in the reported searches). It is therefore unclear how the included, unpublished non-RCT was identified or whether additional, relevant evidence might have been excluded.

The inclusion criteria for the review of safety evidence from RCTs and non-randomised studies were not specified. The safety review included the three RCTs from the main clinical efficacy review (M10-467, PIONEER I, PIONEER II), as well as the M12-555 OLE trial from the review of non-randomised and non-controlled studies. However, as noted above, the methods by which the non-randomised study was identified and the criteria by which it was selected, and others were excluded, are not clear.

Table 3: Inclusion and exclusion criteria for the broad clinical efficacy/safety systematic literature review (reproduced from CS,[9] Section 4.1.1)

Inclusion criteria	
Population	Adult patients with moderate to severe HS were included. The inclusion criteria were not limited by the definition of HS severity, and hence, severity, as defined by HS severity index (HSSI), HS-Physician's Global Assessment score (HS-PGA) or Hurley score
Intervention	▪ Biologics: adalimumab, etanercept and infliximab ▪ Antibiotics: erythromycin, metronidazole, minocycline, clindamycin, cephalosporins, penicillins, long-term antibiotics (erythromycin, tetracycline etc.) ▪ Steroids: high-dose oral steroids, prednisolone, intralesional corticosteroid injection, oestrogens and dapsone ▪ Retinoids (acitretin) ▪ Surgery: laser
Comparators	The comparators of interest included placebo, any of the interventions of interest mentioned above or standard of care. The choice of comparators matches the commonly used comparators in the trials of HS.
Outcomes	At least one of the following efficacy measures should be reported in the relevant studies identified: ▪ Clinical response as assessed by HiSCR, HS-Physician's global assessment (HS-PGA) or HS severity index (HSSI) ▪ Hurley score ▪ HS-lesion, activity and severity (HS-LASI) score ▪ Patient skin pain assessment ▪ MSS ▪ DLQI ▪ Major Depression Inventory (MDI)
Study design	The study selection was restricted to RCTs conducted in more than 10 patients. Data reported at the end of the first period of randomised crossover studies were considered.
Language	English only

4.1.3 Critique of study selection and data extraction

No information was given in any of the reviews regarding the data extraction process (for example, the number of reviewers involved, or actions taken to minimise error). This was addressed however in response to clarification requests from the ERG, in which the company detailed standard processes for data extraction in systematic review (see clarification response,[17] question A3). Following standard systematic review good practice, trials were independently selected for inclusion by two reviewers, with any discrepancies between reviewers resolved through discussion or the intervention of a third reviewer. Data extraction was also performed by one reviewer and independently checked for errors against the original trial report by a second reviewer. Any discrepancies were resolved through discussion or through the intervention of a third reviewer. This is standard good practice for conducting systematic reviews. During the clarification stage, discrepancies and inadequacies in some of the numbers reported in the PRISMA flowchart were acknowledged and addressed by the company, and an updated PRISMA flowchart was provided (see clarification response,[17] question A4).

4.1.4 Quality assessment

For the review of clinical efficacy evidence, the company conducted a critical appraisal of the three adalimumab 40mg EW trials using the NICE risk of bias assessment tool (see CS,[9] Section 4.6) and a critical appraisal of all four adalimumab and relevant comparator studies using the Cochrane risk of bias assessment tool (see CS,[9] Appendix 4). This summary focusses only on the three EW adalimumab trials: M10-467 and PIONEER I and II.

The CS reports that the M10-467 trial was at low risk of bias across all domains using both tools (see CS,[9] Section 4.6, Table 12, page 76). The assessment in Appendix 4 correctly made separate risk of bias assessments for Period 1 (a triple-arm, randomised, blinded study period) and Period 2 (a single arm, open-label extension period). The data from Period 2 are not relevant to this appraisal because all participants received the unlicensed EOW dose of the study drug. The ERG accepts that the data from Period 1 of M10-467 are likely to be subject only to a low risk of bias.

For PIONEER I and II, the CS reports that, *"The results for PIONEER I and PIONEER II are published only as two abstracts. Therefore, most of the details required for quality assessment are not reported for these two studies"* (CS,[9] Section 4.6, page 76). As a result, the company judged the trials to be at "intermediate" or "low risk of bias" across all domains using the NICE tool (see CS,[9] Section 4.6, Table 12, page 76). Given the acknowledged limitations in performing critical appraisal of study design and conduct using the very limited information available in published abstracts, a more

accurate assessment using the NICE tool might have been to categorise the risk of bias as "unclear" across all domains.

Using the Cochrane risk of bias tool, the company then assessed PIONEER I and II to be at low risk of bias across all domains both in Period A and Period B, except for an assessment of "unclear risk of bias" concerning attrition in Period B, and "unclear risk of bias" regarding other, unspecified potential sources of bias (see CS,[9] Appendix 4). There are a number of issues with this assessment. First, different tools are used and the findings are different (the submission used the NICE tool to judge the PIONEER trials to be at "intermediate" risk of bias across most domains, and the Cochrane tool to judge the PIONEER trials to be at "low" risk of bias across most domains). Following a request for clarification on this issue, the company explained that the NICE tool was used for the adalimumab trials in the main efficacy review and the Cochrane tool was used for the trials included in both the efficacy review and the potential indirect comparison (clarification response,[17] question A22). The company also explained that the NICE tool was used for an assessment based on the published M10-467 paper and the published abstracts relating to the PIONEER trials, whilst the Cochrane risk of bias assessment was based on the clinical study reports (CSRs) only (see clarification response,[17] questions A20, A21). There was no reported rationale for this distinction. Second, the two PIONEER trial periods (A and B) were not formally assessed separately, even though there are differences in study design and conduct between these periods (specifically relating to randomisation, attrition and discontinuation). In response to a request for clarification from the ERG on this matter, the company reiterated the findings reported in the CS, in which PIONEER I and II were judged to be at low risk of bias across all domains both in Period A and Period B, except for an assessment of "unclear risk of bias" concerning attrition in Period B (see clarification response,[17] questions A20, A21 and A23). The ERG disagrees with some of the company's risk of bias assessments relating to the PIONEER I and II trials. The differences between the company's assessments and those made by the ERG are detailed in Tables 4 and 5 using the Cochrane risk of bias criteria only, as this is the accepted standard tool for conducting assessments of risk of bias in RCTs. The assessment has had to be made for the PIONEER trials using the CSRs alone because the trials are currently unpublished and have not been subjected to peer review.

Table 4: Risk of bias assessment - PIONEER I

Risk of bias	Period A		Period B	
	CS	**ERG**	**CS**	**ERG**
Selection bias	LOW	LOW	LOW	MODERATE *All are re-randomised to maintain blind, but randomisation is false for some who can only be assigned to placebo for Period B*
Performance bias	LOW	LOW	LOW	LOW-MODERATE *There is no evaluation of blinding to determine whether it was effective*
Detection bias	LOW	LOW	LOW	LOW
Attrition bias	LOW	LOW-MODERATE *NRI for some primary and LOCF for some secondary outcomes due to up to 6% attrition (CS, p.71); imputation might over-estimate effect*	UNCLEAR	MODERATE *NRI for some primary and LOCF for some secondary outcomes due to 45% attrition from 12-week baseline across arms (CS, p.71); imputation might over-estimate effect*
Reporting bias	LOW	MODERATE *The protocol lists original and "current" outcomes, which are different; DLQI, TSQM, HADS, SF-36, CRP, fistulas, AEs are all reported in CS but are not listed in the protocol. However, clinical advice did not specify any other outcomes that were not included. Outcomes are listed in the protocol for 12 weeks only, but a text description of the trial makes mention of the Period B and a study duration of 36 weeks*	LOW	MODERATE *The protocol lists original and "current" outcomes, which are different; DLQI, TSQM, HADS, SF-36, CRP, fistulas, AEs are all reported in CS but are not listed in the protocol. However, clinical advice did not specify any other outcomes that were not included. Outcomes are listed in the protocol for 12 weeks only, but a text description of the trial makes mention of the Period B and a study duration of 36 weeks*
Other bias	UNCLEAR	MODERATE *Manufacturer-funded, some issues with selective reporting*	UNCLEAR	MODERATE *Manufacturer-funded, some issues with selective reporting*

NRI - non-responder imputation; LOCF - last observation carried forward; DLQI - Dermatology Life Quality Index; TSQM - Treatment Satisfaction Questionnaire for Medicine; HADS: Hospital Anxiety and Depression Score; SF-36: Short-Form 36, CRP - C-Reactive Protein; AE - adverse event; CSR - clinical study report

Table 5: Risk of bias assessment - PIONEER II

Risk of bias	Period A		Period B	
	CS	**ERG**	**CS**	**ERG**
Selection bias	LOW	LOW	LOW	MODERATE *All are re-randomised to maintain blind, but randomisation is false for some who can only be assigned to placebo for Period B*
Performance bias	LOW	LOW	LOW	LOW-MODERATE *There is no evaluation of blinding to determine whether it was effective*
Detection bias	LOW	LOW	LOW	LOW
Attrition bias	LOW	LOW-MODERATE *NRI for some primary and LOCF for some secondary outcomes due to 6% attrition (CS, p.71); imputation might over-estimate effect*	UNCLEAR	MODERATE *NRI for some primary and LOCF for some secondary outcomes due to more than 50% attrition across arms (CS, p.71); imputation might over-estimate effect*
Reporting bias	LOW	MODERATE *The protocol lists original and "current" outcomes, which are different; DLQI, TSQM, HADS, SF-36, CRP, fistulas, AEs are all reported in CS but are not listed in the protocol. However, clinical advice did not specify any other outcomes that were not included. Outcomes are listed in the protocol for 12 weeks only, but a text description of the trial makes mention of the Period B and a study duration of 36 weeks*	LOW	MODERATE *The protocol lists original and "current" outcomes, which are different; DLQI, TSQM, HADS, SF-36, CRP, fistulas, AEs are all reported in CS but are not listed in the protocol. However, clinical advice did not specify any other outcomes that were not included. Outcomes are listed in the protocol for 12 weeks only, but a text description of the trial makes mention of the Period B and a study duration of 36 weeks*
Other bias	UNCLEAR	MODERATE *Manufacturer-funded, some issues with selective reporting*	UNCLEAR	MODERATE *Manufacturer-funded, some issues with selective reporting*

NRI - non-responder imputation; LOCF - last observation carried forward; DLQI - Dermatology Life Quality Index; TSQM - Treatment Satisfaction Measure; HADS - Hospital Anxiety and Depression Score; SF-36 - Short-Form 36, CRP - C-reactive protein; AE - adverse events; CSR - clinical study report

With respect to Period A of both trials, the ERG agrees with the company's judgement that the overall risk of bias is low, albeit with the exception of possible low-to-moderate level bias in terms of attrition and reporting. However, the ERG considers there to also be a moderate or unclear risk of selection and attrition bias for the results of Period B, especially given the absence of any evaluation of the blinding, and the high level of attrition. LOCF imputation was used for secondary outcomes to manage missing data; the ERG notes that it has been shown that using LOCF can overestimate efficacy in certain diseases.[27] However, the disease trajectory is difficult to determine for HS, so there is some uncertainty concerning the results based on this method of imputation.

For the non-randomised evidence, a single additional, non-RCT study (M12-555 OLE[20]) was identified and its findings were presented within the CS. A quality assessment was performed for this study using an unspecified tool and no rationale was provided for its selection. In response to a request for clarification from the ERG, the tool was later specified by the company as the Centre for Reviews and Dissemination (CRD) non-RCT tool (see clarification response,[17] question A22). Given that only simple "Yes", "No" or "Not relevant" responses are presented by the company, it is difficult to establish how these judgements were reached. The ERG disagrees with some of the company's risk of bias assessments relating to the M12-555 OLE study (Table 5). The differences between the company's assessments and those made by the ERG are detailed in Table 6.

Table 6: Company's critical appraisal of M12-555 OLE using CRD non-RCT tool (reproduced from CS,[9] Table 27, page 105)

Criterion	Assessment	Response
Bias in results?	Was there significant potential for bias? List the reasons that have led to this conclusion.	No Clear inclusion and exclusion criteria
Study question	Does the study clearly address a specific question? Has the study question been specifically stated?	Yes Yes
Methodology	Were the methods clearly described, with enough detail that you could repeat the study exactly? Were appropriate methods used to answer the specified research question? Were the outcome measures used appropriate? Are the methods sufficiently flawed as to make the results unreliable?	Yes Yes Yes No
Population and data collection	Was the population under study described adequately? Were the inclusion/exclusion criteria sufficiently described? Was the population under study selected/ recruited in an appropriate way? Was the collection of data complete enough (in terms of size of population and follow-up period)?	Yes Yes Yes, OLE Interim results only
Results and confounding factors	Were the results presented in a clear and useful manner? Were the tables/graphs clearly labeled, easily interpretable, and discussed sufficiently to enable understanding of the meaning of the results? Could the results be due to chance or bias (as highlighted by the authors and/or by your own judgment)? Have the authors identified possible confounding factors that may have influenced the results (such as age, gender, ethnicity, socioeconomic status, occupation, etc.)? Have these factors been incorporated into the analysis (i.e. have the results been presented as crude and adjusted ratios)?	Yes Yes No Not relevant Not relevant No
Statistical methods	Were the statistical methods clearly described? Was any rationale given for the methodology of analysis used? Were the factors used to adjust a model (if any) detailed clearly, with reasoning given for their selection? Were any unusual methods used?	Yes Yes Not relevant No
Conclusions	Do the authors provide a clear discussion of the results that leads to a single, specified conclusion in answer to the specified study question? Do the authors relate their results to any previous literature in the field? Is there consistency between the conclusions and the results presented?	Yes, but interim results Yes Yes

Owing to difficulties in qualifying the company's judgements regarding the risk of bias in the OLE study, the ERG conducted its own critical appraisal using the Critical Appraisal Skills Programme (CASP) tool for cohort studies,[28] as this is an accepted standard tool for conducting assessments of risk of bias in studies with this type of design (see Table 7). The ERG identified the following issues: the study was not blinded so there is potential for detection bias; regression analyses have not yet been conducted to control for potentially confounding variables, and; LOCF is used to account for a large amount of missing data. There is therefore a great deal of uncertainty regarding the findings of this single arm, non-controlled, unblinded, unpublished OLE study.

Table 7: ERG critical appraisal of M12-555 OLE using CASP cohort study checklist

Question	Assessment
1. Did the study address a clearly focused issue? HINT: A question can be 'focused' In terms of • The population studied • The risk factors studied • The outcomes considered • Is it clear whether the study tried to detect a beneficial or harmful effect?	Yes
2. Was the cohort recruited in an acceptable way? HINT: Look for selection bias which might compromise the generalisibility of the findings: • Was the cohort representative of a defined population? • Was there something special about the cohort? • Was everybody included who should have been included?	Yes, an extension study of all responders, partial responders and non-responders from two relevant, placebo-controlled RCTs in the same trial
3. Was the exposure accurately measured to minimise bias? HINT: Look for measurement or classification bias: • Did they use subjective or objective measurements? • Do the measurements truly reflect what you want them to (have they been validated)? • Were all the subjects classified into exposure groups using the same procedure	Yes. All subjects were classified into asingle group. Compliance was measured and monitoring conducted at 4-8 week time-points to determine outcomes or discontinuation
4. Was the outcome accurately measured to minimise bias? HINT: Look for measurement or classification bias: • Did they use subjective or objective measurements? Principally PROMs and some investigator assessments (all subjective), plus some objective measures e.g. CRP • Do the measures truly reflect what you want them to (have they been validated)? • Has a reliable system been established for detecting all the cases (for measuring disease occurrence)? • Were the measurement methods similar in the different groups? • Were the subjects and/or the outcome assessor blinded to exposure (does this matter)?	Overall, Yes, clinical and patient-reported outcome measures. Most measures were validated, though the primary outcome measure, HiSCR, has some known correlation and inter-rater reliability issues.[29] Further, there are some concerns about the "partial response" outcome measure, which is *post hoc* and non-validated. Yes, frequent visits; efforts to make sure the same investigator is making judgments each time: CSR, section 9.5.1.1 Yes No, this was an un-blinded, open-label study: there is potential for detection bias
5. (a) Have the authors identified all important confounding factors? **List the ones you think might be important, that the author missed.** **(b) Have they taken account of confounding factors in the design and/or analysis? List:** HINT: Look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors	Yes. Severity of disease; gender, BMI, antibiotic use, disease duration, CRP, concomitant interventions, smoking status etc. No. Details of subgroups and confounding factors at baseline are given, but the reported results are simply proportions of patients exposed to "continuous" adalimumab who achieved a response: there were no regression or sensitivity analyses (CS, pp.102, 106)
6. (a) Was the follow up of subjects complete enough?	Most of those subjects without data are

Question	Assessment
(b) Was the follow up of subjects long enough? HINT: Consider □ The good or bad effects should have had long enough to reveal themselves □ The persons that are lost to follow-up may have different outcomes than those available for assessment □ In an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort?	patients who simply have not reported data yet - so LOCF is used - which introduces greater uncertainty into the results
7. What are the results of this study? HINT: Consider □ What are the bottom line results? □ Have they reported the rate or the proportion between the exposed /unexposed, the ratio/the rate difference? □ How strong is the association between exposure and outcome (RR)? □ What is the absolute risk reduction (ARR)?	Results consist of basic proportions of patients in the different groups achieving a response: only results for patients who have experienced "continuous" adalimumab exposure are presented, not all groups
8. How precise are the results? HINT: Look for the range of the confidence intervals, if given.	Basic frequencies, based on LOCF to manage missing data - therefore some uncertainty
9. Do you believe the results? HINT: Consider □ Big effect is hard to ignore! □ Can it be due to bias, chance or confounding? □ Are the design and methods of this study sufficiently flawed to make the results unreliable? □ Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency)	Proportions with response - and trends of response - are similar and consistent across groups. However, large numbers of missing patients and data, and the extensive use of LOCF after week 24, renders these findings more uncertain
10. Can the results be applied to the local population? HINT: Consider whether □ A cohort study was the appropriate method to answer this question □ The subjects covered in this study could be sufficiently different from your population to cause concern □ Your local setting is likely to differ much from that of the study □ You can quantify the local benefits and harms	No UK centres, but clinical advice to the ERG suggests that results for the trial patients are generaliseable
11. Do the results of this study fit with other available evidence?	Yes, similar to the main findings of the original two RCTs
12. What are the implications of this study for practice? HINT: Consider □ One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making □ For certain questions observational studies provide the only evidence □ Recommendations from observational studies are always stronger when supported by other evidence	More longer-term RCT evidence with improved follow-up and fewer missing data is needed, with larger numbers to manage any attrition, and more complete sensitivity analyses of confounding factors to address uncertainties. Otherwise, this study offers some limited but useful data on efficacy and useful data on medium-to-long-term safety

For the review of the safety evidence (see CS,[9] Section 4.12), data from four studies were presented: M10-467, PIONEER I/II, and M12-555 OLE. Quality assessment of these studies was performed within the CS. The ERG accepts the overall low risk of bias affecting the safety data from M10-467,

PIONEER I and II, but has identified a number of issues with the conduct and reporting of the M12-555 OLE study (see Table 7).

4.1.5 Evidence synthesis

The synthesis for the review of clinical efficacy was a basic descriptive summary of the evidence from the M10-467, PIONEER I and PIONEER II trials. The selected approach to evidence synthesis was neither described nor justified in the CS, but was described in response to a clarification question from the ERG, as *"evidence extracted … was summarised and then reported in tabulated form"* (see clarification response,[17] question A2). A meta-analysis was not performed. In response to a request for clarification from the ERG, the company stated that a separate meta-analysis was unnecessary because data from the PIONEER trials had been pooled in the efficacy results section (see clarification response,[17] question A7).

An NMA comparing effects across all treatments was not performed by the company. The CS notes that there were substantial differences in trial characteristics in trials comparing different pairs of treatments. The CS argues that trial charcteristics such as smoking status, CRP status, disease severity, and prior and concomitant medication were potential treatment effect modifiers. Therefore, the company argues that because there were insufficient trials to adjust for trial characteristics it was not possible to produce unbiased estimates of treatment effects. In addition, the company argues that trials did not provide data on all outcome measures so that the number of trials with usable data varied with the outcome measure.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Review of clinical efficacy (relevant RCT evidence)

The CS provides a very detailed, extensive description of three trials identified by the company as satisfying the requirements of the final NICE scope,[8] i.e. adalimumab compared with alternative treatments (see Table 8). Three RCTs compared adalimumab 40mg EW with placebo: a Phase II trial, M10-467,[25] and two Phase III trials: PIONEER I,[18] and PIONEER II.[19]

It should be noted that only one of the trials (M10-467) has been published in full in a journal article.[25] Whilst some details of the study design and some of the results of the PIONEER trials have been published as conference abstracts,[19, 30-32] these have not been fully published as journal articles. As a result, these two trials and their results have not been subjected to rigorous peer review. The ERG has therefore conducted its critique principally based on information contained within the CSRs and the data presented in the main text of the CS.

All three included trials were international and multicentre. The inclusion criteria in all three trials were adult patients with moderate or severe HS. Moderate to severe HS requires lesions to be present in at least two distinct anatomical areas, one of which has to be Hurley Stage II or III. Patients had to have an AN count of >3 at the baseline visit. Patients who were unresponsive or intolerant to oral antibiotics were eligible for enrolment, although antibiotics were permitted as concomitant therapy for some or all participants in all trials.

Table 8: Characteristics of included RCTs (reproduced in part from CS,[9] Table 6, pages 52-53)

Study	Interventions	Study duration	Study design	Inclusion criteria	Exclusion criteria
M10-467	ADA 40mg EW vs. ADA 40mg EOW vs. placebo	52 weeks	International, multicentre, 16 week double-blind randomised controlled phase followed by a 36 week open label phase in which all patients received ADA	≥18 years, moderate to severe HS (HS-PGA score of moderate or worse) in at least 2 distinct anatomical areas and were unresponsive or intolerant to oral antibiotics as assessed by the investigator were eligible for enrolment*	Prior treatment with ADA or any other TNF antagonist therapy (e.g., infliximab or etanercept) or had received any systemic nonbiologic therapy within 4 weeks of baseline. Patients were allowed stable doses of oral (tetracycline, doxycycline, or minocycline) or topical (clindamycin) antibiotic treatment for HS
PIONEER I	ADA 40mg EW vs. placebo	36 weeks	International, multicentre, 12 week double-blind randomised controlled phase (Period A) followed by a 24 week double-blind phase (Period B) in which patients treated with ADA EW in Period A were re-randomised to ADA EW or EOW or placebo. Patients who were on placebo in Period A were assigned (using re-randomisation numbers) to receive ADA40 mg EW	Men or women ≥18 years; HS diagnosis >1 year, HS lesions in at least two distinct anatomical areas, one of which must be at least Hurley Stage II or Hurley Stage III, stable HS for at least 60 days prior to screening visit, inadequate response to at least a 90 day treatment of oral antibiotics for treatment of HS, and a count of ≥3 at baseline	Previously treated with ADA or another anti-TNF therapy (e.g., infliximab or etanercept); not on a stable dose of antibiotic (for at least 28 days prior to entry; received oral concomitant analgesics (including opioids) for HS-related pain, on opioid analgesics, not on a stable dose of non-opioid oral analgesics, within 14 days prior to entry
PIONEER II	ADA 40mg EW vs. placebo	36 weeks	International, multicentre, 12 week double-blind randomised controlled phase (Period A) followed by a 24 week double-blind phase (Period B) in which patients treated with ADA EW in Period A were re-randomised to ADA EW or EOW or placebo. Patients who were on placebo in Period A continued on placebo in Period B	Men or women ≥ 18 years; HS diagnosis >1 year, HS lesions in at least two distinct anatomical areas, one of which must be at least Hurley Stage II or Hurley Stage III, stable HS for at least 60 days prior to screening visit, inadequate response to at least a 90 day treatment of oral antibiotics for treatment of HS, and a count of ≥3 at baseline	Previously treated with ADA or another anti-TNF therapy (e.g., infliximab or etanercept); not on a stable dose of antibiotic for at least 28 days prior to the baseline visit; received oral concomitant analgesics (including opioids) for HS-related pain, on opioid analgesics, not on a stable dose of non-opioid oral analgesics, within 14 days prior to baseline visit

ADA – adalimumab; EW - every week; EOW - every other week; HS - hidradenitis suppurativa

More than 600 participants received the licensed 40mg EW dose during the three RCTs and the non-controlled OLE study. The three RCTs were also the only trials to evaluate the licensed 40mg EW dose of adalimumab. The final selection of the three included trials for the main clinical efficacy review was therefore considered to be appropriate by the ERG.

The M10-467 Phase II "dosing" trial recruited adults with moderate to severe HS, according to the HS Physician's Global Assessment [HS-PGA] score, who were "unresponsive or intolerant to oral antibiotics" as assessed by the investigator, using the following definition:

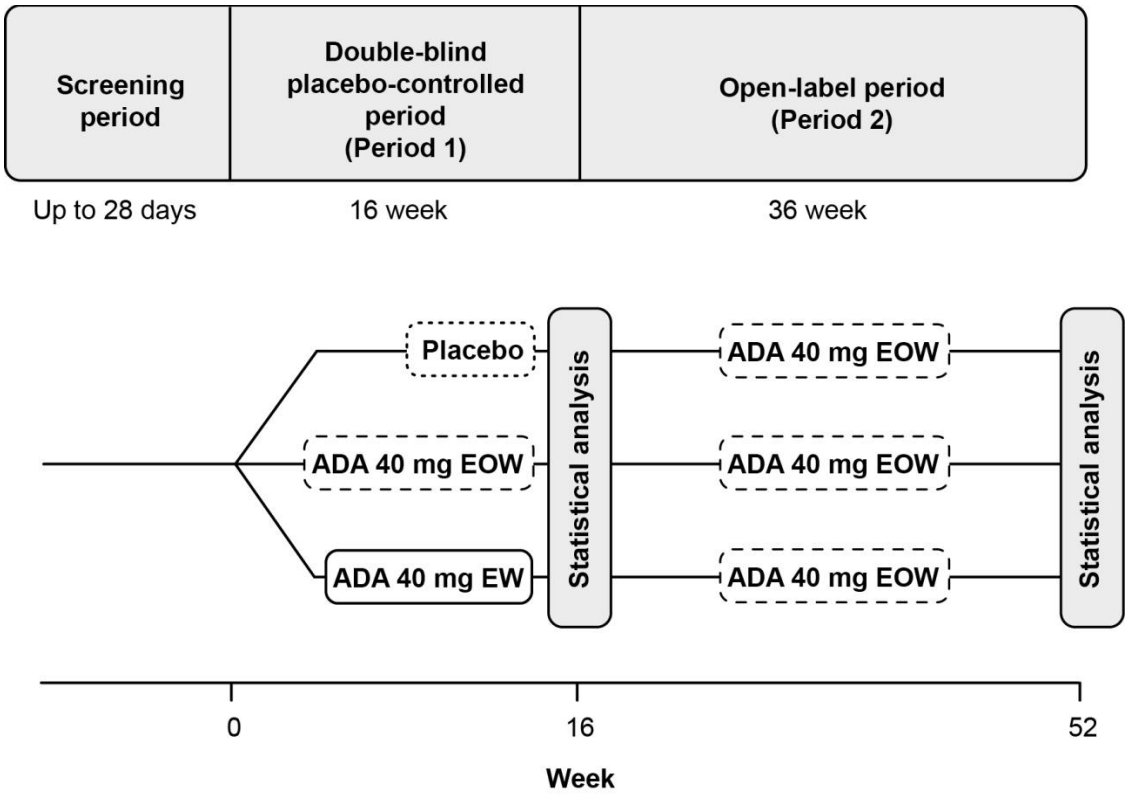
If, after at least 90 days of oral antibiotic therapy, any of the following had occurred, the patient was deemed to have experienced an inadequate response, or loss of response to oral antibiotics:

- Progression of Hurley Stage (i.e., the Hurley Stage of at least one affected anatomic region has progressed from I→II, II→III, or I→III).
- Requirement for at least 1 intervention (e.g., incision and drainage or intra-lesional injection of corticosteroid).
- Pain interfering with activities of daily living, with unsatisfactory relief from over-the-counter analgesics (e.g., ibuprofen or paracetamol).
- Pain requiring opioids, including tramadol.
- Drainage interfering with activities of daily living (e.g., requires multiple dressing changes and/or changes of clothes daily)
- An increase in the number of anatomic regions affected by HS
- At least one new abscess or one new draining fistula (CS,[9] page 62).

Patients were ineligible if they had previously received treatment with adalimumab or any other anti-TNF agent or if they had received any systemic non-biologic therapy within 4 weeks of baseline. In Study M10-467, patients were allowed oral (tetracycline, doxycycline, or minocycline) or topical (clindamycin) antibiotic treatment for HS if they had received a stable dose for at least 4 weeks before the baseline visit and were willing to maintain stable dosing during the study.

The trial design and patient flow is represented in the CS[9] (see Figure 1).

Figure 1: Design of study M10-467 (reproduced from CS,[9] page 58)



ADA - adalimumab; EW - every week; EOW - every other week

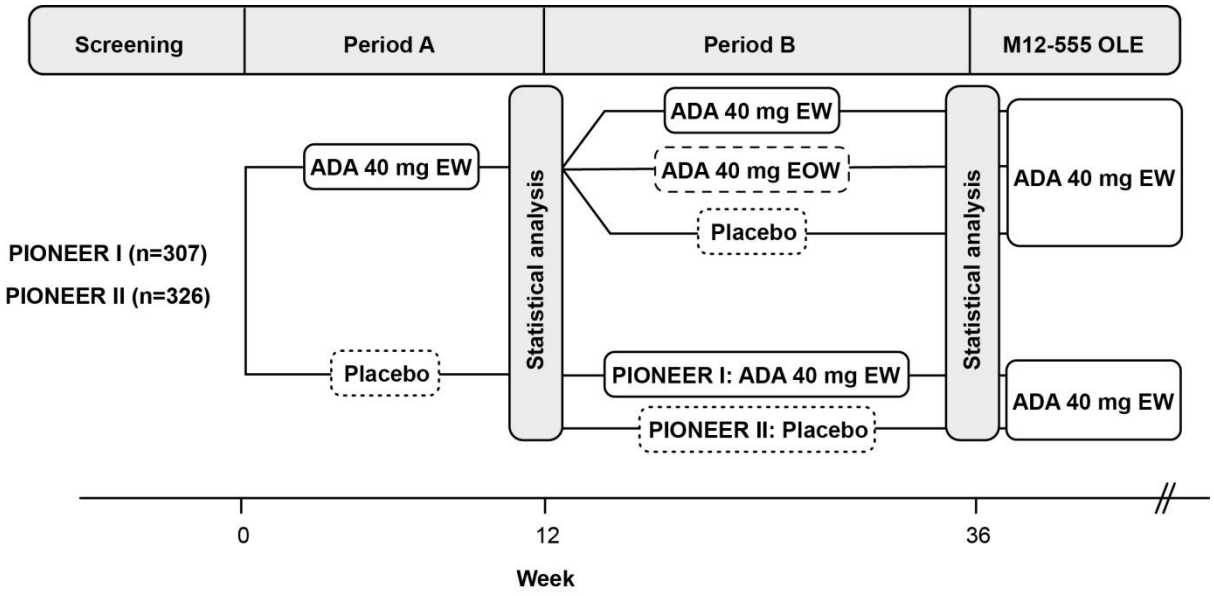
The M10-467 trial directly compared adalimumab in its licensed dose of 40mg EW with placebo in adults with moderate to severe HS. However, only the adalimumab 40mg EW and placebo data from Period 1 (baseline to week 16) are relevant to this efficacy appraisal because the EOW dose received by some participants in Period 1, and by all participants in Period 2, is not licensed for use in HS in the UK.

The PIONEER I and II trials recruited adults with moderate to severe HS, according to the HS-PGA score of moderate or worse, in at least two distinct anatomical areas, who were unresponsive or intolerant to oral antibiotics as assessed by the investigator, as defined above. Patients were ineligible if they had previously received treatment with adalimumab or any other anti-TNF agent or had received any systemic non-biologic therapy within 4 weeks of baseline, but were allowed oral (tetracycline, doxycycline, or minocycline) or topical (clindamycin) antibiotic treatment for HS if they had received a stable dose for at least 4 weeks before the baseline visit and were willing to maintain stable dosing during the study.

In PIONEER I, doxycycline or minocycline were permitted as "rescue therapy" if required, and in PIONEER II these two antibiotics were permitted during the trial if participants had received a stable

dose for at least 4 weeks before the baseline visit and were willing to maintain stable dosing during the study (see CS,[9] page 62). The differing use of antibiotics in the two trials is attributed to the principal location of the trial centres: the US Food and Drug Administration (FDA) requested that no antibiotics be used in PIONEER I, whilst the EMA advised that patients should be able to continue on antibiotics in PIONEER II (see clarification response,[17] question A14). The ERG notes that it is unclear why patients who were "unresponsive or intolerant to oral antibiotics" might still receive these treatments as either a background or rescue therapy. The ERG submitted a clarification request on this point and was informed by the company that it depended on whether "the treating physician believed there was some benefit associated with this" (clarification response[17] question A15). The list of other permitted co-interventions for PIONEER trials included: antiseptic wash (chlorhexidine gluconate, triclosan, benzoyl peroxide, or dilute bleach in bathwater) to their HS affected body regions; injection with intralesional triamcinolone acetonide suspension, and; incision and drainage (see clarification response,[17] question A12). The trial design and patient flow is represented in Figure 2 (see CS[9] page 61).

Figure 2: Design of the PIONEER I and II trials (reproduced from CS,[9] page 61)



ADA - adalimumab; EW - every week; EOW - every other week

The PIONEER trials directly compared adalimumab in its licensed dose of 40mg EW with placebo in adults with moderate to severe HS in both Period A and Period B. The adalimumab 40mg EOW data from Period B are not relevant to this appraisal of efficacy because the EOW dose received by some participants in Period B is not licensed for use in HS in the UK.

A list of excluded studies, with reasons, was provided by the company (see CS,[9] Section 4.1 and Appendix 3).

The ERG noted that the three RCTs included in the main clinical efficacy review (CS, Sections 4.2-4.9) compared adalimumab with placebo, which was not a designated comparator in the final NICE scope,[8] which only listed "established clinical management without adalimumab." This was justified by the company, with the CS stating: *"Given that there is no standard of care for moderate to severe HS, placebo is an appropriate comparator for ADA 40 mg"* (see CS,[9] page 56). The ERG accepts that there is no published head-to-head RCT evidence comparing adalimumab with other biologics, steroids, retinoids or surgical intervention for HS, hence a comparison with placebo provides the best available, relevant trial evidence. Antibiotics were available as a possible background therapy in all arms of the PIONEER II trial, whilst incision and drainage of lesions was permitted as required in all three trials[18,19,25] and was reported as being performed in the M10-467 trial on 7%-10% of patients during the trial period.[25] However, it should be noted one of the company's clarification responses suggests the opposite, stating that, *"Surgery was not permitted in the PIONEER I and II studies per protocol"* (Clarification response,[17] question B5), whilst a second included incision and drainage in a list of permitted co-interventions (see clarification response,[17] question A12). The definition and role of surgery in the trials is therefore unclear.

The three trials collected data on several outcomes. The outcome measures for each trial and their relationship to the final NICE scope are summarised in Table 9. This information was compiled by the ERG, with supplementary details provided by the company in response to a request for clarification (see clarification response,[17] question A9).

Table 9: Final scope outcomes and trial outcome measures

NICE final scope outcomes	M10-467	PIONEER I	PIONEER II
Primary outcome			
Clinical response	HS-PGA, HiSCR*, MSS, AN counts/lesion counts	HiSCR, MSS, AN counts/lesion counts	HiSCR, MSS, AN counts/lesion counts
Secondary outcomes			
Disease severity	Hurley, MSS, AN counts/lesion counts, representative lesions	Hurley, MSS, AN counts/lesion counts, representative lesions	Hurley, MSS, AN counts/lesion counts, representative lesions
Inflammation and fibrosis	Hurley, HiSCR, AN counts/lesion counts, representative lesions, erythema lesions	Hurley, HiSCR, AN counts/lesion counts, representative lesions, erythema assessments	Hurley, HiSCR, AN counts/lesion counts, representative lesions, erythema assessments
Discomfort / pain	VAS	PGA- Skin Pain (NRS30)	PGA-Skin Pain (NRS30)
HRQoL	DLQI	DLQI, HSQOL, SF-36	DLQI, HSQOL, EQ-5D
Additional outcomes	WPAI-SHP	WPAI-SHP	WPAI-SHP
	PHQ-9	HADS	

**As a secondary outcome*

Details of the full list of outcomes are given below.

Primary outcomes

- HS-PGA[2,10]
- HiSCR: at least a 50% reduction in the total abscesses and inflammatory nodule (AN) count with no increase in abscess count and no increase in draining fistula count relative to baseline[29]

Secondary outcomes

- MSS score: a clinical scoring system that assesses the number of involved anatomical regions, the number and type of lesions, the extent of involvement and the Hurley stage, was used to assess disease activity;
- Pain Visual Analogue Scale (VAS): Pain assessed using a questionnaire with a VAS ranging from 0 mm (no pain) to 100 mm (maximum pain);
- PGA-Skin Pain: Patient Global Assessment of Skin Pain (NRS30: Numeric Rating Scale 0-30);
- Dermatology Life Quality Index questionnaire (DLQI): a questionnaire which measures dermatology specific HRQoL and ranges from 0 to 30, with 0 being no impairment;
- HS Quality of Life (HSQOL);
- Short Form-36 (SF-36) Health Status Survey;

- Euroqol EQ-5D;
- Work Productivity and Activity Impairment-Specific Health Problem (WPAI-SHP) questionnaire (which ranges from 0 to 100, with 0 being no impairment);
- Patient Health Questionnaire-9 (PHQ-9): self-assessment for depression ranging from 0 to 27, with 0 being no depressive symptoms;
- Hospital Anxiety and Depression Scale (HADS).

The primary efficacy outcome in all three trials was clinical response. In the M10-467 trial, this was measured using the standard HS-PGA scale (see Table 10). Response was defined as achieving a HS-PGA score of clear, minimal or mild, with at least a 2 grade improvement relative to baseline, at week 16.

Table 10: HS-PGA scale[1,2] (reproduced in part from CS,[9] Table 8, page 59)

Rating	Description
Clear	0 abscesses, 0 draining fistulas, 0 inflammatory nodules and 0 non-inflammatory nodules
Minimal	0 abscesses, 0 draining fistulas, 0 inflammatory nodules and presence of non-inflammatory nodules
Mild	0 abscesses, 0 draining fistulas, and 1–4 inflammatory nodules or 1 abscess or draining fistula and 0 inflammatory nodules
Moderate	0 abscesses, 0 draining fistulas, and $\geq$ 5 inflammatory nodules or 1 abscess or draining fistula and $\geq$ 1 inflammatory nodule or 2–5 abscesses or draining fistulas and <10 inflammatory nodules
Severe	2–5 abscesses or draining fistulas and $\geq$10 inflammatory nodules
Very severe	>5 abscesses or draining fistulas

In the PIONEER trials, clinical response was measured using the HiSCR measure. HiSCR is defined as at least a 50% reduction in the total abscesses and inflammatory nodule (AN) count with no increase in abscess count and no increase in draining fistula count relative to baseline.[9] This measure was validated using data from the M10-467 but has only been used as a primary endpoint in the PIONEER trials and is untested in other published studies evaluating therapies for HS. The validation study found moderate to strong correlations between HiSCR and MSS, Hurley stage and HS-PGA.[29] Clinical advice received by the ERG suggests that the HiSCR measure is appropriate, but has some limitations: principally that clinical response alone is inadequate for decision-making and that clinically efficacious treatment must also take account of patient reported outcome measures relating to pain and HRQoL. A published clinical commentary also noted that the HiSCR did not achieve moderate levels of correlation (Spearman's rho >0.4) with measures of skin pain or quality of life as measured by the DLQI in the validation study, and that treatment effect must include a separate

assessment of pain and quality of life, as well as HiSCR.[33] The validation study[29] and the commentary[34] both acknowledge that the inter-rater reliability of the measure has not been demonstrated.

The ERG also notes that the CS[9] includes "partial response" as an efficacy outcome in Section 4.7.2.3, page 89; this is defined by the CS as a ≥25% reduction in the total abscesses and AN count with no increase in abscess count and no increase in draining fistula count relative to baseline. However, this was not a pre-specified response category in the PIONEER I/II trials, nor is it explained or justified in the CS, and its clinical validity as a response category has not been demonstrated. Rather, this represents a *post hoc* analysis; this was acknowledged as such in the company's clarification response[17] (question A10). The ERG notes that it is unclear whether a 25% reduction in AN count represents a clinically meaningful difference.

In terms of secondary outcomes, all three trials used the MSS to measure disease activity by scoring the number of involved anatomical regions, the number and type of lesions, the extent of involvement and the Hurley stage. However, clinical advice received by the ERG acknowledges that the application of the MSS is both complex and time consuming.

For pain and quality of life, the trials used a variety of patient-reported outcomes measures (PROMs). Study M10-467 used a standard VAS for pain, whilst the PIONEER I/II trials used the NRS30 specific skin pain tool. Quality of life was assessed across all three trials using the DLQI tool and, in the PIONEER trials, the condition-specific HSQOL tool was also used. In addition, PIONEER I also used the SF-36 and PIONEER II used the EQ-5D. The ERG asked the company to clarify why each PIONEER trial had not used one or both measures. The company responded that the decision was made to include only one instrument in each study, despite measuring different aspects of quality of life, because of the unacceptable patient burden involved in the large number of questions across both measures (see clarification response,[17] question A17). Whilst, the PIONEER trials used different quality of life instruments, the ERG considered the use of all of these measures to be appropriate. The trials also collected data on depression and productivity outcomes, but these were not listed in the final NICE scope and were therefore not considered relevant to this appraisal. Clinical advice received by the ERG also suggests that depression is multifactorial, therefore it is difficult to attribute any improvement in depression scores to changes in the severity of HS. It should also be noted that the CS inclusion criteria specified the HSSI and the HS-LASI score as outcomes (see CS,[9] page 47), neither of which were reported for any of the included trials, although clinical advice to the ERG suggests these are similar measures to the MSS score.

4.2.2 Results

Participants' baseline characteristics

The patients in each of the trials were generally similar in terms of age, gender and disease duration, and were similar between the PIONEER trials in terms of Hurley Stage (see Table 11). The M10-467 trial had a smaller proportion of patients with Hurley Stage III disease (29.4% in M10-467 vs 46.6% in PIONEER I and 46.3% in PIONEER II).

There were some notable differences in patient characteristics between the trials in terms of potential treatment effect modifiers, such as AN count and MSS score, especially between the PIONEER trials. PIONEER I participants appear to have had more severe disease based on these criteria. In the PIONEER trials there were higher proportions of participants with prior surgery in the adalimumab arms compared with the placebo arms (13.7% vs 8.4% in PIONEER I and 16.6% vs 11.0% in PIONEER II), which might be suggestive of more severe disease. PIONEER II included a higher proportion of smokers than the other trials. However, clinical advice received by the ERG suggested that the trial patients were broadly representative of the patients that are encountered in usual clinical practice.

Table 11: Participants' baseline characteristics in M10-467, PIONEER I and PIONEER II (reproduced from CS,[9] Table 11, pages 74-75)

Baseline characteristic	M10-467		PIONEER I			PIONEER II		
	Placebo (n=51)	ADA EW (n=51)	Placebo (n=154)	ADA EW (n=153)	Total (n=307)	Placebo (n=163)	ADA EW (n=163)	Total (n=326)
Female, n (%)	36 (70.6)	36 (70.6)	105 (68.2)	91 (59.5)	196 (63.8)	113 (69.3)	108 (66.3)	221 (67.8)
White, n (%)	37 (72.5)	37 (72.5)	118 (76.6)	116 (75.8)	234 (76.2)	130 (79.8)	143 (87.7)	273 (83.7)
Black, n (%)	8 (15.7)	9 (17.6)	29 (18.8)	33 (21.6)	62 (20.2)	20 (12.3)	9 (5.5)	29 (8.9)
Other	6 (11.7)	5 (9.8)	7 (4.5)	4 (2.6)	11 (3.6)	13 (7.9)	11 (6.7)	24 (7.3)
Age, years; mean [SD]	37.8 [12.1]	35.1 [10.7]	37.8 [11.33]	36.2 [10.83]	37.0 [11.10]	36.1 [12.18]	34.9 [9.96]	35.5 [11.13]
Hurley stage I, n (%)	36 (70.6)	36 (70.6)		-				
Hurley stage II, n (%)			81 (52.6)	80 (52.3)	161 (52.4)	89 (54.6)	86 (52.8)	175 (53.7)
Hurley stage III, n (%)	15 (29.4)	15 (29.4)	73 (47.4)	73 (47.7)	146 (46.6)	74 (45.5)	77 (47.2)	151 (46.3)
Disease duration, years; mean [SD]	13.4 [10.4]	11.3 [9.1]	11.6 [8.86]	11.3 [9.00]	11.5 [8.92]	11.8 [9.41]	11.3 [8.66]	11.5 [9.03]
AN count; mean [SD]			14.4 [14.80]	14.3 [11.92]	14.3 [13.42]	11.9 [11.02]	10.7 [8.10]	11.3 [9.68]
MSS; mean [SD]			147.3 [97.16]	151.0 [131.17]	149.1 [115.19]	122.6 [88.00]	107.5 [80.03]	115 [84.32]
NRS skin pain at worst; mean [SD]			(n=146) 4.8 [2.68]	(n=151) 5.1 [2.51]	(n=297) 5.0 [2.60]	(n=155) 4.8 [2.73]	(n=159) 4.3 [2.62]	(n=314) 4.5 [2.69]
BMI, kg/m^2; mean [SD]			(n=154) 34.5 [7.94]	(n=152) 33.0 [7.62]	(n=306) 33.8 [7.80]	32.9 [7.94]	31.3 [7.41]	32.1 [7.71]
Body weight, kg, mean [SD]	96.5 [24.8]	95.4 [22.9]						
Prior surgery for HS, n (%)			13 (8.4)	21 (13.7)	34 (11.1)	18 (11)	27 (16.6)	45 (13.8)
HS-CRP (C-reactive protein), mg/L; mean [SD]	13.3 [15.0]	21.5 [33.1]	17.4 [20.2]	20.3 [25]	18.9 [22.75]	18.3 [30.72]	13.3 [17.96]	15.8 [25.25]
Current smokers, n (%)	29 (56.9)	30.0 (58.8)	92 (59.7)	81 (52.9)	173 (56.4)	109 (67.3)	105 (64.4)	214 (65.8)

ADA - adalimumab; EW - every week; HS - hidradenitis suppurativa; BMI - body mass index; CRP - C-reactive protein; SD - standard deviation; MSS score - Modified Sartorius Severity score; NRS - numerical rating scale

Participant flow and numbers

The trials all experienced substantial loss of patients to follow-up (see Table 12). Clinical advice received by the ERG suggests that this is expected in trials of HS because patients who do not experience a response are unlikely to be motivated to continue on the trial. The loss to follow-up in the three trials was reported in the participant flow figures in the CS (pages 70-72), although the company had to provide, at the request of the ERG, the correct flowchart for the PIONEER II trial because this was erroneously a duplicate of the PIONEER I flowchart in the original submission (see clarification response,[17] question A24). Patient loss to follow-up in Period B was produced in part by protocol-driven discontinuation. This was based on either LOR, defined as a loss of 50% or more of the improvement gained during Period A among patients who achieved response according to HiSCR at week 12, or WOAI, defined as the second incidence of two consecutive visits with AN count higher than the baseline AN count in patients randomised to adalimumab 40mg EW in Period A who were week-12 HiSCR non-responders.[9]

Table 12: Patient loss to follow-up in trials in the adalimumab 40mg EW and placebo arms

Time endpoint (weeks)	**M10-467 n (%)**		**PIONEER I n (%)**		**PIONEER II n (%)**	
	ADA	PBO	ADA	PBO	ADA	PBO
Baseline total	51 (100)	51 (100)	153 (100)	154 (100)	163 (100)	163 (100)
12			145 (95)	145 (94)	155 (95)	151 (93)
16	45 (88)	46 (90)				
36			170 (55)*		116 (40)*	
52	31 (69)	34 (74)				

ADA - adalimumab; EW - every week; PBO – placebo
**Pooled numbers because of crossover between periods A and B*

According to the CS, clinical response data for the first period in each study (12 or 16 weeks) were analysed according to the intention-to-treat (ITT) principle, so that all patients randomised at week 0 were included (see CS,[9] pages 68 and 69). The primary approach for managing missing values was non-responder imputation (NRI). However, many of the results for the secondary endpoints, as presented in the CS, were based on LOCF imputation, which has particular implications for the results beyond weeks 12 or 16 as the level of attrition was more than 40% (see Table 12). Consequently, when this approach has been used, it was specified in CS and is also specified in this ERG report. In other instances, when the imputation approach has not been specified in the CS, it is assumed that NRI was used for binary outcomes.

4.2.2.1 Primary outcome: Clinical response

Results for the primary outcome for all three trials were reported in the CS. The M10-467 dosing study measured this outcome using both HS-PGA (see Table 13) and HiSCR, whilst PIONEER I and

II both used the HiSCR (Table 14). Response using the HS-PGA scale was defined as a HS-PGA score of clear, minimal or mild, with at least a 2-grade improvement relative to baseline.

The trials each had two separate periods of treatment. Period 1 (M10-467) and Period A (PIONEER I, II) evaluated whether adalimumab induces clinical response in patients with moderate or severe HS. The duration of this period was 16 weeks in Study M10-467, and 12 weeks in PIONEER I and II. M10-467 had a Period 2, for weeks 16-52, but this period only assessed the unlicensed 40mg EOW dose and so these data are not relevant to this appraisal. The PIONEER trials also included a Period B, covering weeks 12 to 36.

Weeks 12 and 16 (Period A in the PIONEER I/II trials and Period 1 in Study M10-467)

In Study M10-467, using the HS-PGA outcome measure, significantly more patients in the adalimumab 40mg EW group achieved clinical response compared with placebo at week 16 (17.6% vs 3.9%, $p<0.025$).

Table 13: Percentage of patients achieving clinical response measured by HS-PGA relative to baseline at 16 weeks (data reproduced from CS,[9] pages 76-77)

Trial	n	Follow-up (weeks)	Adalimumab EW	Placebo	Percentage difference relative to placebo (95% CI)	*p*-value
M10-467	102	16	17.6	3.9	13.7% (1.7 to 25.7)	<0.025

ADA - adalimumab; EW - every week

Across all three trials, the percentages of patients experiencing clinical response using HiSCR, defined as at least a 50% reduction in the total AN count with no increase in abscess count and no increase in draining fistula count relative to baseline, are reported in Table 14. Across all three RCTs, the percentage of patients achieving clinical response according to the HiSCR measure at week 12 or week 16 was significantly higher for patients receiving adalimumab 40mg EW compared with placebo ($p<0.01$).

Table 14: Percentage of patients achieving clinical response by HiSCR relative to baseline at 12 or 16 weeks (data reproduced from CS,[9] pages 78-79, and Table 14, page 80)

Trial	n	Follow-up (weeks)	ADA EW	Placebo	Percentage difference relative to placebo (95% CI)	*p*-value
M10-467*	102	16	54.5	25.6	38.9 (NR)	<0.007
M10-467*	102	12	59.1	16.3	42.8 (NR)	<0.001
PIONEER I†	307	12	41.8	26.0	15.9 (5.3, 26.5)	0.003
PIONEER II†	326	12	58.9	27.6	31.5 (20.7, 42.2)	0.001

ADA - adalimumab; EW - every week; NR - not reported.
Note: the figures here are reproduced from the CS; the percentage differences in PIONEER I and II are inaccurate, and should be 15.8% and 31.3% respectively
**Secondary outcome; † ITT population (using NRI);*

The rate of absolute clinical response using the HiSCR was similar across the placebo arms of the three trials at 12 weeks (25.6%, 26% and 27.6% for M10-467, PIONEER I and PIONEER II, respectively). The rate of absolute clinical response using the HiSCR was numerically different across the adalimumab arms of the three trials at 12 weeks (42.8%, 15.9% and 31.5% for M10-467, PIONEER I and PIONEER II, respectively, see Table 15).

Table 15: Percentage of patients achieving clinical response measured by HiSCR relative to baseline at week 12 in PIONEER I and II† (reproduced from CS,[9] Table 14, page 80)

	Adalimumab EW	Placebo	Difference (95% CI)	p-value
PIONEER I	**64/153 (41.8%)**	**40/154 (26.0%)**	**15.9 (5,3, 26.5)**	**0.003**
Hurley stage II	37/83 (44.6%)	25/84 (29.8%)	14.8 (0.3, 29.3)	0.048
Hurley stage III	27/70 (38.6%)	15/70 (21.4%)	17.1 (22, 32.1)	0.027
PIONEER II	**96/163 (58.9%)**	**45/163 (27.6%)**	**31.5 (20.7, 42.2)**	**<0.001**
Antibiotic use	20/31 (64.5%)	7/32 (21.9%)	42.6 (17.8, 67.5)	<0.001
No antibiotic use	76/132 (57.6%)	38/131 (29.0%)	28.6 (16.9, 40.6)	<0.001
Hurley stage II	53/85 (62.4%)	32/87 (36.8%)	25.5 (10.5, 40.5)	< 0.001
-Antibiotic use	7/11 (63.6%)	3/12 (25.0%)	38.6 (1.1, 76.2)	0.004
-No antibiotic use	46/74 (62.2%)	29/75 (38.7%)	23.5 (7.9, 39.1)	<0.001
Hurley stage III	43/78 (55.1%)	13/76 (17.1%)	38.1 (22.8, 53.3)	<0.001
-Antibiotic use	13/20 (65.0%)	4/20 (20.0%)	45.0 (17.7, 72.3)	0.004
-No antibiotic use	30/58 (51.7%)	9/56 (16.1%)	35.7 (19.6, 51.7)	<0.001

ADA - adalimumab; EW - every week
† ITT population (using NRI)

Differences in the rate of absolute clinical response between the adalimumab groups across the PIONEER trials might be explained by potential treatment effect modifiers such as differences in patient characteristics at baseline: PIONEER II participants appear to have had less severe disease based on AN count (11.3 for PIONEER II vs 14.3 for PIONEER I) and MSS score (115 for PIONEER II vs 149.1 for PIONEER I, see Table 11), as well as a higher BMI and a higher draining

fistula count (see clarification response,[17] question A33). It might also be explained in part by study design differences between the two PIONEER trials: concomitant antibiotic use – permitted according to differences in the inclusion criteria between PIONEER I and II - was also substantially higher among responders in the adalimumab 40mg EW arm in PIONEER II compared with the placebo arm (64.5% vs 21.9%) with a higher percentage difference compared with placebo for those responders continuing to take antibiotics (42.6% vs 28.6%, see Table 15). The extent to which patients' baseline characteristics and co-interventions modify the treatment effect remains unclear and are the subject of further analyses by the company (see CS,[9] page 118).

Weeks 12-36 (Period B in the PIONEER trials)

The outcomes for clinical response in Period B in PIONEER I and II (Weeks 12-36), after re-randomisation, are reported below. Participants are categorised according to exposure to the licensed dose of adalimumab of 40mg EW (e.g. Period A and Period B exposure is categorised as EW/EW; Period A placebo, and Period B adalimumab exposure is categorised as PBO/EW). The data for the unlicensed adalimumab 40mg EOW dose are not reported here.

The results for clinical response in Period B are reported in the CS and in a poster presentation.[9] Numerical data on clinical response were not provided separately for the two trials (only graphs were provided, see CS,[9] page 89). An arm-based integrated summary, which breaks randomisation, was conducted for the two PIONEER trials to tabulate Period B response (HiSCR "full response" according to the definition of response used elsewhere) for, first, all patients and, second, for a subgroup of responders and "partial responders" from week 12 at the end of Period A. Separate numerical data were therefore not provided for the Period A week 12 responders: this group was combined with week 12 "partial responders". This "partial responder" group (defined as HiSCR responders with ≥25% rather than ≥50% AN reduction) represents a *post hoc* analysis group. This group was not defined in the trial protocols or published or unpublished descriptions of study design or pre-specified analysis methods for the PIONEER trials. It was also not considered in the published validation study for the HiSCR measure.[29] It was neither justified nor explained in the CS, but was confirmed by the company, in response to a request for clarification, as a trial outcome group defined *post hoc* in the PIONEER trials through analysis of response (see clarification response,[17] question A10). The categories of "response" and "partial response" are both included in the company's model; this is discussed in further detail in Section 5.3. Issues relating to the arm-based integrated summary are also discussed in Section 5.3.

Participants were stratified by response and Hurley Stage across the adalimumab 40mg EW and the placebo arms at week 12 re-randomisation for Period B. By week 36, the percentage of patients experiencing clinical response had reduced over time in both trial arms, but the reduction was greatest

in the placebo arm compared with the adalimumab 40mg EW arm (from 53% to 28% for placebo vs 53.5% to 43.4% for adalimumab, *p*-value not reported, see Table 16).

By week 36, the percentage of patients experiencing clinical response, who were categorised as responders or partial responders at week 12, reduced over time in both trial arms, but the reduction was greatest in the placebo arm compared with the adalimumab 40mg EW arm (from 72.6% to 30.1% for placebo vs 75.7% to 55.7% for adalimumab, *p*-value not reported, see Table 16).

Table 16: Proportion of patients in PIONEER I and II (amalgamated data) achieving HiSCR during Period B (reproduced from CS,[9] Table 20, page 90)

PIONEER I and II	Period B intervention	n	HiSCR rate at week 12 n (%)	HiSCR rate at week 24 n (%)	HiSCR rate at week 36 n (%)
All patients*	Placebo	100	53 (53%)	30 (30%)	28 (28%)
	ADA 40mg EW	99	53 (53.5%)	44 (44.4%)	43 (43.4%)
Week 12 responders and partial responders†	Placebo	73	53 (72.6%)	24 (32.9%)	22 (30.1%)
	ADA 40mg EW	70	53 (75.7%)	40 (57.1%)	39 (55.7%)

HiSCR – hidradenitis suppurativa complete response; ADA – adalimumab; EW – every week
**ITT analysis; †ITT_B_R (Period B Responders) analysis*

The CS states that the reduction in HiSCR rate over time in Period B might be explained by the study design, according to which any patient who experienced protocol-defined LOR, during Period B relative to week 12 at the end of Period A (which may have been explained by temporary exacerbation of disease), was discontinued from the study and imputed as a non-responder for this period. LOR was defined as a loss of 50% or more of the improvement gained during Period A among patients who achieved response according to HiSCR at week 12.

[redacted] However, data were not provided to support this statement. [redacted] Again, data were not provided by the company to support this statement.

4.2.2.2 Secondary outcomes

Results for the secondary outcomes for all three trials were reported in the CS[9] (pages 77-79 and pages 81-88). The trials had two separate periods of treatment, but secondary outcomes were only reported for the first period of each trial.

Weeks 12 and 16 (Period A in the PIONEER trials and Period 1 in M10-467)

Abscesses and inflammatory nodule counts

In Study M10-467, patients receiving adalimumab 40mg EW demonstrated a significant improvement in inflammatory nodules (p=0.019) and draining fistulae (p=0.05), but not in abscesses (p=0.22, see Table 17).

Table 17: M10-467: Improvement from baseline in individual symptoms in Period 1 (LOCF) (reproduced in part from CS,[9] Table 13, page 78)

M10-467	n	Follow-up (weeks)	Percentage difference ADA EW versus placebo (95% CI)	*p*-value
Inflammatory nodules	[redacted]102	16	37.0 (6.2 , 67.8)	0.019
Abscesses			26.8 (-16.0, 69.5)	0.22
Draining fistulae			36.9 (0.1, 73.7)	0.050

LOCF - last observation carried forward; ADA – adalimumab; EW – every week; CI – confidence interval
** per protocol analysis*

In PIONEER I, at week 12, there was no significant difference between placebo and adalimumab 40mg EW in the proportion of patients achieving an AN count of 0, 1, or 2, either in patients with Hurley Stage II at baseline or in all patients (see Table 18). In PIONEER II, there was a significant difference between placebo and adalimumab 40mg EW in the proportion of patients achieving an AN count of 0, 1, or 2 at week 12, both in patients with Hurley Stage II at baseline (p=0.01) and in all patients (p<0.001, see Table 18). It is not clear why separate data are presented on Hurley Stage II patients alone and there are no separate data on Hurley Stage III patients. It is also not specified in the CS whether these results are from observed or imputed data and, if the latter, whether the imputation was based on LOCF.

Table 18: Percentage of patients who achieved AN count of 0, 1, or 2 at week 12 in PIONEER I and PIONEER II (reproduced from CS,[9] Table 15, page 82)

Trial and patients	n	Follow-up (weeks)	ADA EW	Placebo	Percentage difference relative to placebo (95% CI)	*p*-value
[redacted]	167	12	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	307		[redacted]	[redacted]	[redacted]	[redacted]
PIONEER II (Hurley II only)	172		44/85 (51.8%)	28/87 (32.2%)	19.6 (4.7, 34.2)	0.01
[redacted]	325		[redacted]	[redacted]	[redacted]	[redacted]

ADA – adalimumab; EW – every week; CI – confidence interval

Modified Sartorius Severity Score (MSS score)

Adalimumab 40mg EW was only associated with a statistically significant improvement in MSS score compared with placebo at week 12 in the PIONEER II trial ($p<0.001$). The change from baseline was not significantly different between the adalimumab and placebo groups in the M10-467[25] or PIONEER I trials (see Table 19). It has been argued that this might be because the MSS score includes elements that are not expected to change with adalimumab therapy, such as the number of fistulas.[25]

Table 19: Improvement from baseline in MSS score at weeks 12 and 16 (LOCF) (reproduced in part from CS,[9] Table 13, page 78, and Figure 14, page 84)

Trial	n	Follow-up (weeks)	ADA EW	Placebo	Percentage difference relative to placebo (95% CI)	*p*-value
Change from baseline: mean (±SE)						
M10-467*	102	16	-40.2 (9.8)	17.2 (9.8)	-22.0 (-50.1, 4.2)	0.097
PIONEER I	304	12	24.4	15.7	8.7 (NR)	NS
PIONEER II	325	12	28.9	9.5	19.4 (NR)	<0.001

LOCF - last observation carried forward; NR - not reported; NS - not significant; ADA – adalimumab; EW – every week; CI – confidence interval; SE – standard error
**Data from Kimball 2012 (not the CS)*

Pain

The PIONEER trials used the Patient's Global Assessment of Skin Pain (NRS30) score. The Patient's Global Assessment of Skin Pain is a numerical rating scale ranging from 0 (no skin pain) to 10 (skin pain as bad as you can imagine). The PIONEER trials measured the mean change in skin pain in all patients and in those with a baseline of NRS≥3. In both PIONEER trials, it was reported that patients with a baseline of NRS≥3 taking adalimumab 40mg EW had statistically significant improvements in

this pain score compared with such patients taking placebo (p=0.016 for PIONEER I, and p<0.001 for PIONEER II, see Table 20).

Table 20: Mean change in NRS30 skin pain relative to baseline at Week 12 in patients with baseline of NRS ≥ 3 (LOCF) (reproduced from CS,[9] Table 18, page 87)

Trial	Within group change (LS mean ± SE)		Between group change	*p*-value
	Placebo	ADA EW	LS mean difference (95% CI)	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

LOCF - last observation carried forward; LS - least squares ADA – adalimumab; EW – every week; CI – confidence interval; SE – standard error

Data reported in the CS enabled an assessment as to whether these improvements in pain were clinically meaningful. The M10-467 trial used a VAS measure. According to this measure, a clinically relevant response requires at least a 30% reduction and 10mm reduction in pain relative to baseline (see CS,[9] page 78). A statistically significant percentage of patients achieved a clinically relevant reduction in pain at week 16 (47.9% on adalimumab versus 27.1% on placebo, p=0.037). According to the Patient's Global Assessment of Skin Pain (NRS30), as used in the PIONEER trials, a clinically meaningful response requires at least a 30% reduction and at least a 1 unit reduction from baseline pain score among patients with baseline NRS≥3 (see CS,[9] page 82). In PIONEER I, there was a non-significant numerical improvement in pain using this measure in Period A. However, there was a statistically significant difference at earlier timepoints (see Table 21). In PIONEER II, there was a statistically significant difference in the percentage of patients achieving the clinically relevant endpoint at all timepoints up to and including week 12. It is not specified in the CS whether these results are from observed or imputed data and, if the latter, whether the imputation was based on LOCF (see CS,[9] pages 78 and 82).

Table 21: Percentages of all patients with clinically relevant improvement in pain relative to baseline at weeks 12 and 16 (reproduced from CS,[9] Table 13, page 78, and Figure 13, page 83)

Trial	n	Follow-up (weeks)	Adalimumab EW	Placebo	*p*-value relative to placebo
VAS					
M10-467	103	16	47.9	27.1	0.037
NRS30					
PIONEER I†	231	12	27.9	24.8	NS
PIONEER II†	216	12	45.7	20.7	<0.001

ADA – adalimumab; EW – every week; VAS – visual analogue scale; NRS – numerical rating scale; NS – not significant
† ITT population; NS: Not significant

Quality of life

Several measures were used across the three trials, but the principal recognised measure is the DLQI. DLQI scores range from 0 to 30, with higher scores indicating a more impaired quality of life (see Table 22). Across all three RCTs, adalimumab 40mg EW was associated with a statistically significant improvement in DLQI compared with placebo at week 12 and week 16 (p<0.001).

Table 22: Quality of Life measured by DLQI scores relative to baseline in Weeks 12 and 16 (LOCF) (reproduced from CS,[9] Table 13, page 78, and Table 17, page 86)

Trial	Within group change (LS mean ± SE)		Between group change	*p*-value
	ADA EW	Placebo	LS mean difference (95% CI)	
M10-467	-6.0 ± 0.9	-1.9 ± 0.9	-4.2 (-6.6, 1.8*)	<0.001
PIONEER I	-5.4 ± 0.5	-2.9 ± 0.5	-2.5 (-3.0,-1.8)	<0.001
PIONEER II	-5.1 ± 0.53	-2.3 ± 0.53	-2.8 (-4.1,-1.5)	<0.001

LOCF - last observation carried forward; ADA - adalimumab; EW - every week; LS - least squares; SE – standard error; CI – confidence interval
**This figure from CS, Table 13, page 78*

The CS states that, in all trials, the within arm mean change from baseline in DLQI at week 12 (Period A) or week 16 (Period 1) for patients in the adalimumab 40mg EW group exceeded the minimum clinically important difference (MCID) of 5 (see CS,[9] page 86). It also exceeded the MCID of 4 established by Basra *et al* 2015.[34] However, the ERG notes that the between arm mean change from baseline for the adalimumab arm compared with the placebo arm did not meet this MCID threshold in either PIONEER I or II. In PIONEER I, the CS states that the percentage of patients with a clinically relevant change in DLQI at week 12 was [redacted] in the adalimumab 40mg EW group compared with [redacted] in the placebo group [redacted] and 49% versus 34% (p=0.011) in PIONEER II.

The condition-specific HSQOL scale was also used. Clinical advice received by the ERG suggests that this is a new measure which has not been published. Ratings range from 0 (worst possible) to 10 (best possible). In PIONEER I and PIONEER II, patients receiving adalimumab 40mg EW had significantly improved HSQOL scores compared with placebo patients [redacted] (see Table 23).

[redacted]Table 23: Quality of life measured by HSQOL scores relative to baseline at week 12 (LOCF) (reproduced from CS,[9] Table 17, page 86)

Trial	Within group change (LS mean ± SE)		Between group change	*p*-value
	Placebo	ADA EW	LS mean difference (95% CI)	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

LOCF - last observation carried forward; ADA - adalimumab; EW - every week; LS - least squares; SE – standard error; CI – confidence interval

For the HSQOL, the MCID is defined as an increase in HSQOL of 50% or greater than the standard deviation of HSQOL for all patients at baseline (see CS,[9] page 86). The CS states that, in PIONEER I, numerically more patients in the adalimumab 40mg EW arm than the placebo arm achieved the MCID [redacted] however this was not statistically significant (*p*-value and 95% CIs were not reported in the submission). In PIONEER II, the difference was statistically significant [redacted]

PIONEER I assessed overall quality of life using the SF-36, and demonstrated a significant benefit in the overall physical component with adalimumab 40mg EW compared with placebo ($p<0.05$, see Table 24). The CS states that significantly more patients receiving adalimumab 40mg EW achieved a MCID in SF-36 than patients receiving placebo: [redacted] respectively [redacted] In terms of specific components, patients on adalimumab 40mg EW reported clinically relevant statistically significant improvements in general health compared with placebo, and significant but not clinically meaningful improvements in physical functioning compared with placebo, but reported no significant difference compared with placebo in physical functioning or bodily pain (Table 24). It should be noted that 95% confidence intervals were not reported in the CS. The differences in the mental component of the SF-36 were also not significantly different between the two groups (actual data were not reported in the submission). It was not stated in the CS whether these results were based on LOCF, as the other quality of life results, or data as observed.

[redacted]Table 24: PIONEER I: Change in SF-36 physical component score relative to baseline at week 12 (reproduced from CS,[9] Figure 15, page 85)

SF-36	n	Follow-up (weeks)	Placebo	ADA EW	MCID	*p*-value
Physical component summary	325	12	1.5	4.2	>2.5	<0.05
Physical functioning			1.6	3.2	>2.5	NS
Role physical			2.2	4.5	>2.5	<0.05
Bodily pain			2.4	4.9	>5	NS
General health			-0.4	3	>2.5	<0.001

SF-36 – Short Form 36; ADA - adalimumab; EW - every week; MCID - minimum clinically important difference

PIONEER II assessed overall quality of life using the EQ-5D (using both the health state questionnaire and the VAS). There was a baseline difference in the mean EQ-5D health state scores between the adalimumab and placebo arms (0.6 [SE=0.33] and 0.5 [SE=0.36] respectively), but an apparent significant benefit in both health state and VAS for adalimumab 40mg EW compared with placebo ($p<0.001$, see Table 25).

Table 25: PIONEER II: Mean change from baseline in EQ-5D at week 12 (LOCF) (reproduced in part from CS,[9] Table 16, page 84)

Instrument	Within group change (LS mean ± SE)		Between group change	*p*-value
	Placebo	ADA EW	LS mean difference (95% CI)	
Health state	0 ± 0.02	0.1 ± 0.02	[redacted]	<0.001
VAS	0.5 ± 1.87	9.2 ± 1.88	[redacted]	<0.001

ADA - adalimumab; EW - every week; VAS – visual analogue scale; LOCF - last observation carried forward; LS - least squares

Weeks 12-36 (Period B in the PIONEER trials)

A small number of secondary outcomes were reported for PIONEER I and II only for participants who had achieved a clinical response at week 12 (see Table 26). By week 36 of Period B, there were higher percentages of responders with an AN count of 0, 1 or 2 and a clinically relevant NRS30 score in the adalimumab group compared with the placebo group, and improved MSS scores in the adalimumab group compared with the placebo group. However, it should be noted that some of these differences were not large and the results were based on some very small numbers of patients (range of 15 to 22 patients across all outcomes for both PIONEER trials).

Table 26: PIONEER I and PIONEER II secondary outcomes at week 36 relative to week 12 (reproduced from CS,[9] Table 21, page 91)

Outcome	PIONEER I		PIONEER II	
	EW/EW (n=21)	EW/placebo (n=22)	EW/EW (n=20)	EW/placebo (n=20)
AN count of 0/1/2	9 (42.9%)	5 (22.7%)	10 (32.3%)	9 (29%)
NRS30*	(n=16) 5 (31.3%)	(n=15) 1 (6.7%)	(n=19) 3 (15.8%)	(n=20) 1 (5%)
MSS (LS change from baseline ± SE)	-47.7 ± 9.99	-41.9 ± 9.76	-37.1 ± 11.8	-33.8 ± 13.19

EW – every week; NRS – numerical rating score; MSS – Modified Sartorius Score; SE – standard error; LS – least squares; AN - abscess and inflammatory nodule

** Proportion of patients who achieved at least 30% reduction and at least 1 unit reduction from baseline in Patient's Global Assessment of Skin Pain (NRS30) – at worst at week 12 among patients with baseline NRS ≥ 3*

The CS does not specify whether these results are from observed or imputed data and, if the latter, whether the imputation method was appropriate to generate unbiased estimates of treatment effect or whether sensitivity analyses were used to assess the robustness of the results (see CS,[9] page 91). It is only noted that this was an ITT analysis conducted on Period B data for responders (ITT_B_R).

Other measures

The CS also reports results for depression, treatment satisfaction and WPAI; these data are not reported here as they were not included in the final NICE scope.[8]

Pre-specified subgroups

Pre-planned analyses in the three studies are shown in Table 27. The variables considered within these analyses were chosen to assess the consistency of the primary efficacy endpoint by demographic and baseline characteristics.

The CS concludes that the distribution of patients within each subgroup was similar across treatments, with the exception of baseline AN count (≤ 5, 6-10, 11+) which was significantly different in PIONEER I between the adalimumab 40mg EW and placebo arms: more patients were in the <5 and >11 bands in the placebo group than in the adalimumab 40mg EW group (p=0.018). However, there was no significant difference between treatments in AN count by median (see CS,[9] page 95). The CS found that AN count by median was a treatment effect modifier using an interaction test. However, the ERG notes that AN count by median is data-dependent and is not based on clinical relevance; consequently, this finding may have occurred by chance. Nevertheless, the CS highlights that baseline balance does not mean that a variable is not prognostic of outcome nor a modifier of treatment effect. Similarly, an imbalance in a baseline characteristic does not mean that it affects outcome or treatment effect. Ideally, relevant covariates should be pre-specified.

Table 27: Primary endpoint analysis subgroups in MI0-467, PIONEER I and PIONEER II (reproduced from CS,[9] Table 22, page 93)

Subgroup	M10-467	PIONEER I	PIONEER II
Baseline concomitant use of oral antibiotics (yes/no)	✓		✓
Age group (< 40; 40-64; ≥ 65, if less than 10% of patients were in the ≥ 65 group, that group was combined with the 40-64 group)		✓	✓
Sex (male, female)		✓	✓
Race (white, non-white)		✓	✓
Duration of HS (by median)		✓	✓
Weight (by median)		✓	✓
BMI category: normal (< 25), overweight (25 – < 30), obese (30 – < 40), morbid obesity(≥ 40)		✓	✓
BMI (by median)	✓		
Current smoking status (Y/N)	✓	✓	✓
Baseline hs-CRP level (by median)		✓	✓
Baseline AN count (≤ 5, 6-10, 11+)	✓	✓	✓
Baseline AN count (< median, ≥ median)		✓	✓
Hurley stage (I or II, III)	✓		
Prior HS surgery history (yes, no)		✓	✓
Smoking habit change (increase, decrease)*.		✓	✓
Time from prior HS surgery to the first dose of study drug (by median)		✓	✓

HS – hidradenitis suppurativa; BMI – body mass index; AN - abscess and inflammatory nodule; CRP – C-reactive protein
**Increase in smoking habit was defined as an at least 25% increase from baseline in both the urine cotinine and the urine nicotine level. Decrease in smoking habit was defined as patients with at least 25% decrease from baseline in both the urine cotinine and the urine nicotine level. A change from ND (not detectable) to detectable (< 2 ng/ml or any value ≥ 2 mg/ml) was considered as an increase in smoking habit; and a change from detectable to not detectable was considered as a decrease in smoking habit. Note: This is a reproduction of CS, Table 22, page 93.*

The CS states that response according to the HiSCR criteria was generally not affected by baseline characteristics (patients were stratified by response and Hurley Stage when re-randomised at week 12), but notes also that, [redacted] (see CS,[9] page 97). [redacted]

[redacted]

Post hoc subgroups

A *post hoc* analysis of subgroups based on clinical response measured by the HS-PGA tool was conducted for the M10-467 trial (see Table 28). This analysis found that larger proportions of patients in the adalimumab group, compared with placebo, achieved clinical response in the following subgroups: those with more mild disease (based on Hurley stage) compared with those with more severe disease; those who were current smokers; those who were taking concomitant antibiotics, and; those who had a BMI greater than or equal to the median. The CS notes that these results should be treated with caution because some of the subgroups contained few people (see CS,[9] page 94).

Table 28: M10-467 clinical response at week 16 (proportion of patients achieving HS-PGA of clear minimal or mild with at least a 2-grade improvement relative to baseline at week 16) (reproduced from CS,[9] Table 23, page 95)

Variable	Adalimumab EW (n=51)	Placebo (n=51)	Difference (95% CI) EW vs. placebo
Hurley Stage			
I or II, n/N (%)	8/36 (22.2%)	2/36 (5.6%)	16.7 (1.2, 32.2)
III, n/N (%)	1/15 (6.7%)	0/15 (0)	6.7 (-6.0,19.3)
Current smokers			
Yes, n/N (%)	7/30 (23.3%)	1/29 (3.4%)	18.4 (0.7, 36.1)
No, n/N (%)	2/21 (9.5%)	1/22 (4.5%)	7.2 (-8.8, 23.1)
Received concomitant oral antibiotics for HS			
Yes, n/N (%)	4/9 (44.4%)	0/4 (0)	39.4 (-2.2, 81.0)
No, n/N (%)	5/42 (11.9%)	2/47 (4.3%)	8.0 (-3.1, 19.1)
BMI			
>median, n/N (%)	5/22 (22.7%)	0/25 (0)	26.2 (8.5, 44.0)
<median, n/N (%)	4/29 (13.8%)	2/26 (7.7%)	5.4 (-11.5, 22.4)
CRP level			
>median, n/N (%)	3/18 (16.7%)	1/21 (4.8%)	13.1 (-5.6, 31.8)
<median, n/N (%)	4/20 (20.0%)	1/18 (5.6%)	14.3 (-8.0, 36.6)

ADA - adalimumab; EW - every week; BMI - body mass index; CRP - C-reactive protein

4.2.3 Review of clinical efficacy (non-randomised and non-controlled evidence)

The CS presents findings from a single ongoing, non-randomised, non-controlled, open-label, unpublished study: M12-555 OLE.[20] The ERG notes that the company did not perform a systematic review of the non-randomised and non-controlled evidence. The inclusion criteria for this review were not specified, the methods by which this study was identified, and the criteria by which it was selected, and any others excluded, are not reported. The only searches reported for the clinical effectiveness section of the CS contain an RCT filter (see CS,[9] Appendix 2). Furthermore, no information was given about the data extraction process, as required by standard systematic review guidelines.

The M12-555 OLE study is a continuation study for patients enrolled in PIONEER I and PIONEER II in which all participants receive adalimumab 40mg EW. The aim of the study was to generate longer-term safety, tolerability and efficacy data on adalimumab 40mg EW in patients with moderate or severe HS. The CS justifies the inclusion of this study on the basis that it provided long-term data (see CS,[9] page 100).

Approximately 600 patients from PIONEER I and PIONEER II were eligible to enrol in M12-555 OLE, of which 497 were enrolled. Patients were evaluated for entry into the OLE at the final study visit of the PIONEER trial in which they participated. Starting at baseline, all patients received open-label adalimumab 40mg EW, regardless of treatment assignment in the PIONEER I and II studies. The dose could be reduced to adalimumab 40mg EOW at any time on or after week 24 of the OLE if patients achieved clinical response according to HiSCR criteria during M12-555 OLE, relative to the baseline at week 12 of the initial PIONEER trials, and achieved an AN count of 0 or 1 on at least two consecutive study visits, and the clinician and patient decided that the risk/benefit of reducing the dose of adalimumab was favourable. This reduced dose is currently not licensed for use in patients with moderate or severe HS in the UK.[12] The dose could be increased back up to adalimumab 40mg EW if required by the clinician or patient, although the dose could only be increased once. Study visits occurred at baseline, week 4, week 8, week 12, week 18, week 24, week 36 and every 12 weeks thereafter, at least to week 60. If after week 24, there was no clinically relevant response, then the clinician and the patient explored the risk/benefit of remaining on treatment.

The following concomitant drugs were not allowed: use of oral antibiotics for HS within 28 days of baseline (except those used in prior PIONEER studies), use of prescription topical therapies for HS within 14 days of baseline, use of systemic non-biologic therapies for HS <28 days before baseline,

use of oral concomitant analgesia (including opioids) for HS-related pain within 14 days of baseline or received any other investigational drug for HS within 30 days or five half-lives of baseline.

Patients who prematurely discontinued from the trial, or who completed the trial and did not initiate adalimumab therapy outside the context of the clinical trial, had study visits 4 and 8 weeks after the last administration of study drug to collect blood samples for the measurement of serum adalimumab concentrations and anti-adalimumab antibody.

The results presented in the CS are from an interim data cut, as of 29 April 2014, for 497 patients who received at least one dose of the study drug. The study is ongoing and there were missing data for a total of 368 subjects (74.0%) at the data cut. In other words, only data on 129 (26%) of enrolled patients are reported.

Efficacy results

In terms of efficacy, the primary outcome was the proportion of subjects achieving HiSCR. The unpublished results for those participants who received adalimumab in at least one period (A or B, or A and B) in PIONEER I and II, and who continued into the OLE, are presented in Table 29. The CS reported that [redacted] The numbers listed in Table 29 are the baseline number of patients in each of the groups providing some data on "continuous" exposure to adalimumab 40mg EW, however [redacted] Consequently, these data have been imputed using LOCF, which might overestimate the true level of HiSCR for these later timepoints. Details of the results for secondary outcomes such as MSS and NRS30 were not reported (see CS,[9] page 106).

Table 29: Proportion of patients achieving HiSCR over time from the first dose of adalimumab (LOCF) (reproduced from CS,[9] Table 28, page 106)

Weeks of adalimumab treatment (relative to the first dose in the PIONEER studies)	EW/EW/EW (n=84)	EW/EOW/EW (n=90)	EW/PBO/EW (n=91)
[redacted]	[redacted]	[redacted]	[redacted]

ADA - adalimumab; EW - every week; EOW - every other week; LOCF - last observation carried forward

Summary

The CS states that the OLE study was of "a good standard" following appraisal using a quality assessment tool (see CS,[9] page 105). However, the source of the tool used was not given within the CS, and its selection was not justified, although this was addressed in a clarification from the company (see clarification response,[17] question A22). The ERG conducted a separate appraisal using the CASP tool[28] – a recognised critical appraisal tool for single-arm, cohort studies such as this. The ERG's critical appraisal identified the following issues: the study was not blinded so there is potential for detection bias; regression analyses have not yet been conducted to control for potentially confounding variables, and; sensitivity analyses have not been performed to assess the robustness of the results to different methods for accounting for the large amount of missing data.

The ERG therefore considers the efficacy results in the follow-up phase to be subject to a high degree of uncertainty because they are the result of interim analyses of unpublished study data from a single-arm, non-controlled, un-blinded study with the risk of bias inherent to that design, as well as other methodological issues such as the methods used to account for missing data. The study also only potentially offers efficacy data for up to 72 weeks for a drug that might be taken for many years by patients with moderate to severe HS.

4.2.4 Review of safety (randomised and non-randomised trial evidence)

The submitted review of the safety evidence for adalimumab was extensive; all key AEs are included and particular events are addressed in detail. This review of the safety evidence from the RCTs included the same studies as the main efficacy reviews, thus the ERG has assumed that the description of the inclusion criteria and methods employed for the adalimumab RCT efficacy review also applied to the identification and extraction of the safety data (see CS,[9] Section 4.1.1-4.1.5, pages 46-48).

M10-467, PIONEER I and II (Weeks 12 and 16)

There were no deaths during the studies in any patients who received adalimumab. The rates of AEs leading to discontinuation, any AE, SAEs, and any infectious AEs were all higher in the adalimumab 40mg EW arm in the M10-467 trial than the placebo arm as well as the adalimumab arms in the PIONEER trials. The CS (page 107) states that most of the excess AEs in the adalimumab arms were attributable to headache, which was typically described as being mild or moderate in severity; that the majority of AEs were grade 1 or 2 or similar across all treatment groups, and that AEs were consistent with those seen with adalimumab in previous studies in other indications (see Tables 30 and 31).

However, across the PIONEER trials, the rates of these categories of AE were generally either comparable between the adalimumab and placebo arms, or the rates were actually lower in the adalimumab treatment arms compared with the placebo arms. For example, for the outcome of any AEs, in PIONEER I, the rate was 52.9% for adalimumab 40mg EW versus 61.8% for placebo, whilst in PIONEER II, the rate was 57.7% for adalimumab 40mg EW versus 66.9% for placebo. The ERG suggests that this is because exacerbations of HS were questionably classified as an AE, when such exacerbations might equally reflect the absence or failure of treatment to control HS: rates of exacerbations were higher in the placebo group, leading to higher AE rates in that group. This is supported by the data on specific AEs, which list the most common AEs across the three trials as exacerbation of HS, nasopharyngitis and headache. AEs which occurred in >5% of patients are shown in Table 30.

For the first 12-week period, therefore, the pattern of AEs was generally consistent throughout the three studies and similar tolerability was reported for both PIONEER trials. With the exception of some higher rates in the M10-467 trial, the AEs for the groups treated with adalimumab 40mg EW were comparable with placebo and were reported as being consistent with the known adalimumab safety profile.

Table 30: Key AE rates in in M10-467, PIONEER I and PIONEER II (weeks 16 and 12) (reproduced from CS,[9] Table 29, page 108)

AE, n(%)	M10-467			PIONEER I			PIONEER II		
	ADA EW (n=51)	Placebo (n=51)	Difference EW versus placebo Relative risk (95% CI)	ADA EW (n=153)	Placebo (n=152)	Difference EW versus placebo Relative risk (95% CI)	ADA EW (n=163)	Placebo (n=163)	Difference EW versus placebo Relative risk (95% CI)
Death	0	0		0	0		0	0	
Any AE leading to discontinuation of study drug	2 (3.9%)	0		1 (0.7%)	3 (2%)	0.33 (0.03, 3.15)	4 (2.5%)	7 (4.3%)	0.57 (0.17, 1.91)
Any AE	36 (70.6%)	30 (58.8%)	1.2 (0.9, 1.6)	81 (52.9%)	94 (61.8%)	0.86 (0.7, 1.04)	94 (57.7%)	109 (66.9%)	0.86 (0.73, 1.02)
SAE	4 (7.8%)	2 (3.9%)	2 (0.38. 10.44)	3 (2%)	5 (3.3%)	0.6 (0.14, 2.45)	3 (1.8%)	6 (3.7%)	0.5 (0.13, 1.97)
Any infectious AE	17 (33.3%)	18 (35.3%)	0.94 (0.55, 1.62)	38 (24.8%)	43 (28.3%)	0.88 (0.6, 1.28)	41 (25.2%)	53 (32.5%)	0.77 (0.55, 1.09)
Infectious SAE	1 (2.0%)	0		1 (0.7%)	0		1 (0.6%)	2 (1.2%)	0.5 (0.005, 5.46)
Cancer	0	0		0	1 (0.7%)		0	0	

ADA – adalimumab; EW – every week; AE – adverse event; SAE – serious adverse event; CI – confidence interval

Table 31: Specific AEs occurring in >5% of patients in M10-467, PIONEER I and PIONEER II (weeks 16 and 12) (reproduced from CS,[9] Table 30, page 109 and Table 34, page 114)

AE	M10-467			PIONEER I			PIONEER II		
	ADA EW (n=51)	Placebo (n=51)	Difference EW versus placebo Relative risk (95% CI)	ADA EW (n=153)	Placebo (n=152)	Difference EW versus placebo Relative risk (95% CI)	ADA EW (n=163)	Placebo (n=163)	Difference EW versus placebo Relative risk (95% CI)
Exacerbation of HS	4 (7.8%)	6 (11.8%)	0.67 (0.2, 2.22.)	14 (9.2%)	20 (13.2%)	0.7 (0.36, 1.33)	7 (4.3%)	21 (12.9%)	0.33 (0.15, 0.76)
Nasopharyngitis	6 (11.8%)	6 (11.8%)	1 (0.35, 2.89)	9 (5.9%)	16 (10.5%)	0.56 (0.25, 1.23)	9 (5.5%)	10 (6.1%)	0.9 (0.38, 2.16)
Headache	8 (15.7%)	2 (3.9%)	4 (0.89, 17.93)	14 (9.2%)	15 (9.9%)	0.93 (0.46, 1.86)	21 (12.9%)	21 (12.9%)	1 (0.57, 1.76)
Upper respiratory tract infection	4 (7.8%)	2 (3.9%)	2 (0.38,10.44)				8 (4.9%)	9 (5.5%)	0.89 (0.35, 2.25)
Diarrhoea	0	2 (3.9%)					9 (5.5%)	4 (2.5%)	2.25 (0.71, 7.16)

ADA – adalimumab; EW – every week; HS – hidradenitis suppurativa; CI – confidence interval

Other specific AEs occurring in >5% of patients in the M10-467 trial are listed in Table 32.

Table 32: AEs occurring in >5% of patients in study M10-467 (week 16) (reproduced from CS,[9] Table 30, page 109)

AE	ADA EW (n=51)	Placebo (n=51)	Difference EW versus placebo Relative risk (95% CI)
Arthralgia	3 (5.9%)	1 (2.0%)	3 (0.2, 27.89)
Back pain	1 (2.0%)	1 (2.0%)	1 (0.06, 15.56)
Cellulitis	0	1 (2.0%)	
Cough	3 (5.9%)	0	
Fatigue	2 (3.9%)	2 (3.9%)	1 (0.15, 6.83)
Folliculitis	0	3 (5.9%)	
Gastroenteritis	0	0	
Gastroesophageal reflux disease	3 (5.9%)	0	
Influenza	2 (3.9%)	0	
Nausea	4 (7.8%)	1 (2.0%)	4 (0.46, 35.57)
Oropharyngeal pain	1 (2.0%)	1 (2.0%)	1 (0.06, 15.56)
Pruritus	1 (2.0%)	0	
Sinusitis	2 (3.9%)	1 (2.0%)	2 (0.19, 21.37)
Vomiting	1 (2.0%)	3 (5.9%)	0.33 (0.04, 3.10)

ADA – adalimumab; EW – every week; CI – confidence interval

PIONEER 1 and II (Weeks 12-36)

AE rates were similar for weeks 12-36 (Period B) to those seen at week 12 (Period A) (see Table 33 and Table 34).

[redacted]

Table 33: AEs in PIONEER I and PIONEER II during Period B[34, 35](reproduced from CS,[9] Table 33, page 113)

AE	PIONEER I				PIONEER II			
	Placebo/EW (n=145)	EW/placebo (n=49)	EW/EOW (n=48)	EW/EW (n=48)	Placebo/placebo (n=151)	EW/placebo (n=51)	EW/EOW (n=53)	EW/EW (n=51)
[redacted]								

EW – every week; EOW – every other week; AE – adverse event; SAE – serious adverse event

Table 34: AEs occurring in >5% of patients PIONEER I and PIONEER II during Period B (reproduced from CS,[9] Table 34, page 113)

AE	PIONEER I				PIONEER II			
	Placebo/EW (n=145)	EW/placebo (n=49)	EW/EOW (n=48)	EW/EW (n=48)	Placebo/placebo (n=151)	EW/placebo (n=51)	EW/EOW (n=53)	EW/EW (n=51)
[redacted]								

EW – every week; EOW – every other week; HS – hidradenitis suppurativa

M12-555 OLE

[redacted] from an interim data cut (29th April 2014) [redacted]

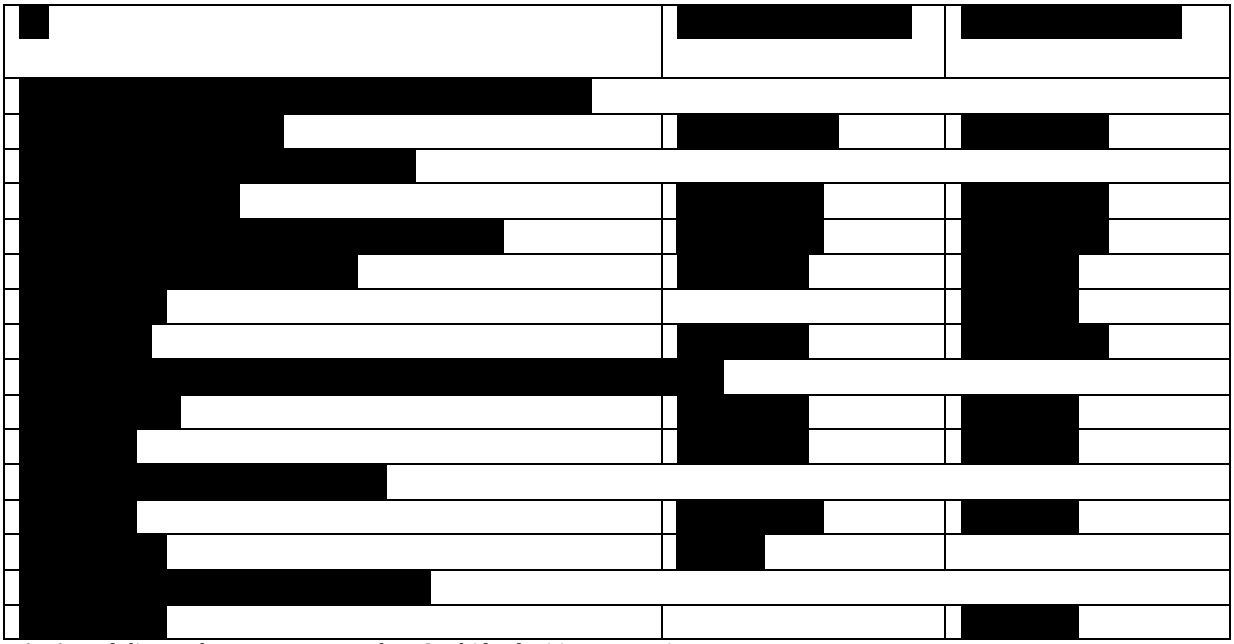
The ERG also notes that the OLE study only potentially offers safety data for up to 72 weeks for some participants (given the high levels of attrition) for a drug that might be taken for many years by patients with moderate or severe HS.

Table 35: AEs in M12-555 OLE in all adalimumab groups and in the EW/EW/EW trial population (reproduced from CS,[9] Table 35, page 115)

[redacted]

ADA – adalimumab; EW – every week; AE – adverse event; SAE – serious adverse event

Table 36: AEs occurring in >5% of in patients in M12-555 OLE in all adalimumab groups and in the EW/EW/EW trial population (reproduced from CS,[9] Table 36, page 115)



ADA – adalimumab; EW – every week; HS – hidradenitis suppurativa

Overall summary

On the basis of the evidence presented withn the CS, adalimumab appears to have a generally acceptable safety profile. However, longer-term data are required to determine whether AE rates are maintained for patients on long-term maintenance doses of adalimumab 40mg EW, whether or not certain subgroups of patients are at higher risk of certain AEs, and to confirm whether or not there are any differences between the interrupted and uninterrupted regimens.

4.2.5 Ongoing studies

The CS indicates that the M12-555 OLE[20] is the only ongoing study of adalimumab for HS. Final data from this study are expected to be available in 2016.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

An NMA comparing effects across all treatments was not performed. The company noted that there were substantial differences in baseline characteristics in trials comparing different pairs of treatments. The company argues that trial charcteristics such as smoking status, CRP status, disease severity, and prior and concomitant medication were potential treatment effect modifiers. Therefore, they argue that because there were insufficient trials to adjust for trial characteristics it was not possible to produce unbiased estimates of treatment effects. However, the company did not find any evidence that the specified subgroups were treatment effect modifiers in PIONEER I and that AN count by median was the only potential treatment effect modifier in PIONEER II. Neverthless, there

may be other trial characteritstics that are treatment effect modifiers. The CS does not discuss the potential to perform a simulated treatment comparison or a matching-adjusted treatment comparison.

In addition, the CS argues that trials did not provide data on all outcome measures, hence the number of trials with usable data varies between outcome measures. In principle, it might be possible to perform a multivariate NMA and borrow strength about treatment effects on different outcome measures across trials, although the CS does not consider this option. The ERG notes that even if an indirect comparison had been deemed suitable using one or more secondary outcomes measured in the PIONEER trials, such evidence could not have been used within the response-based health economic model structure developed by the company (see Chapter 5).

4.4 Critique of the indirect comparison and/or multiple treatment comparison

As discussed in Section 4.3, an NMA comparing effects across all treatments was not performed.

4.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG did not undertake any additional analyses for the clinical effectiveness review.

4.6 Conclusions of the clinical efficacy section

The CS consisted of three separate reviews: (1) a review of the clinical efficacy evidence from RCTs of treatments for HS, specifically RCTs comparing adalimumab with placebo; (2) a review of the evidence from a non-controlled, OLE study; and (3) a review of safety evidence from the RCTs of adalimumab versus placebo and the non-controlled, OLE study.

The principal clinical efficacy review is a poorly-reported systematic review of three relevant RCTs comparing adalimumab with placebo in adults with moderate to severe HS: these were a Phase II "dosing" trial, M10-467, and two Phase III trials, PIONEER I and II. The three trials all have two periods: an initial period (weeks 0-12 in the PIONEER I/II trials and weeks 0-16 in the M10-467 trial) comparing adalimumab 40mg EW with placebo, and a second period (weeks 12-36 in the PIONEER trials), initiated by re-randomisation of patients at Week 12 to arms of adalimumab 40mg EW, placebo or adalimumab 40mg EOW (PIONEER trials only). The included trials are generally consistent with the final NICE scope. The primary efficacy outcome was clinical response to treatment using two measures: the HS-Physicians' Global Assessment (HS-PGA) and the company's own HiSCR. Clinical advice received by the ERG confirms that the HiSCR measure has been validated but, in terms of clinical decision-making, its findings must be viewed alongside the results of patient-reported outcome measures, in particular quality of life assessed by the DLQI and a pain measure. Secondary outcomes in the trials included assessments of disease severity and symptoms, using the MSS score and AN counts, pain and quality of life (various measures).

The ERG considers the M10-467 trial to be at low risk of bias across all domains for the relevant Period 1 (up to week 16). The ERG also considers the results from Period A (i.e. up to week 12) in PIONEER I and II to be generally at low risk of bias. However, the ERG considers there to be a moderate or unclear risk of selection and attrition bias for the results of Period B in the PIONEER trials. There is also a low-to-moderate risk of reporting bias in Period B in the two trials. It should also be noted that whilst M10-467 has been published, the PIONEER trials have not.

In PIONEER I and II, significantly more patients in the adalimumab 40mg EW group achieved a clinical response (defined as achieving HiSCR [at least a 50% reduction in the total AN count with no increase in abscess count and no increase in draining fistula count relative to baseline] at week 12) than patients receiving placebo: 41.8% for adalimumab vs 26.0% for placebo, p=0.003 in PIONEER I, and 58.9% for adalimumab vs 27.6% for placebo, p<0.001 in PIONEER II. Subgroup analyses indicated that patients achieved benefit with adalimumab 40mg EW regardless of their baseline characteristics, although some subgroups had small patient numbers. Significant or clinically relevant differences in favour of adalimumab 40mg EW that were reported for secondary outcomes in PIONEER II were not always found in PIONEER I, especially for AN count, MSS score, pain and some components of quality of life measured by the SF-36. The treatment effect varied between the trials. This might be explained by differences in patient demographics and study design between trials. The company is conducting ongoing analyses of the data from the PIONEER trials and the OLE study to understand these differences. An NMA was not considered feasible.

An arm-based integrated summary, which breaks randomisation, was conducted for the two PIONEER trials to tabulate Period B response (12-36 weeks) for all patients and for a group of HiSCR "responders" and "partial responders." According to this analysis, improvements in response were maintained or reduced in this second period. However, the "partial responder" group (defined as HiSCR responders with ≥25% reduction rather than ≥50% reduction) are a *post hoc* analysis group. This group was not defined in protocols or published descriptions of study design or pre-specified analysis methods for the PIONEER trials. It was also not considered in the published validation study for the HiSCR measure, nor was it justified or explained in the company's clinical review. A small number of secondary outcomes were reported for PIONEER I and II for weeks 12-36, but only for patients who had had clinical response at week 12. However the results were based on analyses with small sample sizes (range of 15 to 22 patients across all outcomes for both PIONEER trials).

These trials were supplemented by a single, unpublished, non-randomised, non-controlled, un-blinded cohort study, which was an OLE study of the PIONEER trials (M12-555 OLE). In terms of efficacy, the results suggested

[redacted]

[redacted].

Details of the results for secondary outcomes such as MSS and NRS30 were not reported. The ERG considers these efficacy results to be affected by a large degree of uncertainty because they are drawn from interim analyses of unpublished study data, and a sizeable amount of data was missing. The study also only potentially offers efficacy data for up to 72 weeks for a drug that might be taken for many years by patients with moderate or severe HS.

The submission of safety evidence was a review of the three RCTs, supplemented by the single arm cohort study. There were no obvious safety concerns, with most AEs being balanced across adalimumab 40mg EW and placebo trial arms, and small numbers of SAEs. Longer-term data are required to determine whether reported AE rates are maintained for patients on long-term maintenance doses of adalimumab 40mg EW; whether or not certain subgroups of patients are at higher risk of certain events; and to confirm whether or not there are any differences between the interrupted and uninterrupted regimens. The submission states that the M12-555 OLE is the only ongoing study of adalimumab in this indication. Final data from this study are expected to be available in 2016.

5. COST EFFECTIVENESS

This chapter presents a summary and critical appraisal of the methods and results of the company's review of published economic evaluations and the *de novo* health economic analysis presented within the CS.

5.1 ERG comment on the company's systematic review of cost-effectiveness evidence

5.1.1 Description of company's systematic review of cost-effectiveness evidence

The CS[9] presents the methods and results of systematic reviews of existing health economic evaluations of treatments for patients with moderate to severe HS, HS cost and resource use studies and HRQoL studies in patients with HS. The searches for the economic evaluation review and the cost and resource use review were run together in order to avoid potential duplicates, whilst the HRQoL search was run separately. According to the CS, the purpose of the combined search was *"to identify healthcare resource use, costs, cost drivers, previous economic evaluations and health technology assessment (HTA) economic models of treatments for patients with moderate to severe HS"* (CS[9] page 127).

Search strategy

All searches were undertaken across the following electronic databases:

- MEDLINE
- MEDLINE In-Process
- EMBASE (using EMBASE.com)
- Econlit (using EBSCO.com)
- The Cochrane Library including the following:
    - The Cochrane Database of Systematic Reviews
    - The Database of Abstracts of Reviews of Effectiveness (DARE)
    - The Cochrane Central Register of Controlled Trials (CCRCT)
    - The Health Technology Assessment (HTA) Database.

Both the combined search and the HRQoL search were restricted to studies which were published in English in the last 15 years (up to 30th July 2015).

In addition to the searches of electronic databases, key HTA websites (NICE, the Scottish Medicines Consortium [SMC] and the All Wales Medicines Strategy Group [AWMSG]) were searched for

relevant HTA evaluations and models from the last two years (precise dates of the searches are not specified in the CS). Conference posters and abstracts were also searched within the following conferences over the last two years:

- American Academy of Dermatology (AAD)
- European Society for Dermatological Research (ESDR)
- World Congress of Dermatology (WCD)
- International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Precise dates for these searches are not specified in the CS.

Inclusion and exclusion criteria

The company's inclusion and exclusion criteria for the three searches are summarised in Box 1. Inclusion and exclusion criteria relating to outcomes and study design differed between each of the three reviews, although all other criteria were the same.

Box 1: Inclusion and exclusion criteria for company's review of cost and resource use / utility / economic evaluation studies (adapted from CS[9] Tables 40, 41 and 44)

Inclusion criteria

- **Population:** Treated and/or untreated adult patients with moderate to severe HS.
- **Interventions:** Any treatment for HS, including antibiotics, retinoids, dapsone, ciclosporin, biologics, surgical options, oral prednisolone, intralesional triamcinolone injections, oral contraceptive pills, incision/drainage and analgesia.
- **Outcomes:** Cost and resource use studies: Any study quantifying costs or resource use requirements of HS and its management or quantifying the costs or resource use associated with disease- or treatment-related AEs / Utility studies: Utility values produced using generic preference-based measures of patient utility, disease-specific measures or vignettes / Economic evaluation studies: Studies with a comparison of costs between the intervention and comparator arms with results reported in terms of cost per QALY gained, cost per life year gained or just cost if accompanied by a cost-minimisation argument.
- **Study design:** Cost and resource use studies: Cost studies or resource use studies or economic evaluations reporting costs or resource use / Utility studies: QoL studies, economic studies or observational studies reporting utility values / Economic evaluation studies: Full economic evaluations, comparing at least two interventions.
- **Other criteria:** Studies providing sufficient detail regarding methods used and studies which provided extractable results, studies must present cost and resource use data (preferably for the UK).

Exclusion criteria

- **Population:** Patients with any skin disease other than HS, healthy volunteers, children only.
- **Outcomes:** Cost and resource use studies: Studies that do not report either cost or resource use / Utility studies: Disease-specific and non-preference-based measures not converted to utilities / Economic evaluation studies: Cost-only outcomes without a cost-minimisation argument.
- **Study design:** Reviews, letters, comments / Economic evaluation studies: burden of illness studies.
- **Other criteria:** Studies that failed to present sufficient methodological detail or failed to present extractable results.

Study selection

The CS[9] (page 132) presents a PRISMA flow diagram for all three searches combined. Results of all three searches were screened together. The PRISMA diagram indicates that after studies were identified through the searches, they were subject to primary screening followed by secondary

screening of those studies that had not been excluded thus far, based on the inclusion/exclusion criteria presented in Box 1. No details were provided regarding the difference between primary and secondary screening or the number of researchers involved in the screening process.

Results of the company's review of cost-effectiveness evidence

The company's electronic searches yielded a total of 400 potentially relevant unique citations. Of these, 143 articles were economic and cost/resource use studies, whilst 212 articles were HRQoL studies. Forty-five citations were also identified through "published ahead of print" searches. Of the 400 potentially relevant citations, 276 publications were excluded at the primary screening stage following a review of titles and abstracts. Following secondary screening, 92 of the remaining 124 studies were excluded.

A total of 21 studies reported across 32 publications were included in the company's reviews of economic evaluations, resource use and cost studies and HRQoL studies. Of these, five studies reported resource use data for patients with moderate to severe HS and 20 reported HRQoL data. The CS[9] notes that no relevant HTAs were identified through searching the NICE, SMC or AWMSG websites and no economic evaluations or modelling studies were identified for HS patients from any of the company's searches.

The CS notes that a Cochrane review[35] was published in October 2015; owing to the time of its publication, this study was not included in the company's review. The Cochrane review did not identify any additional studies that were not identified by the company's systematic review.

The results of the reviews are summarised in the CS in Table 45 (study characteristics) and Table 46 (relevant outcomes reported) for utility studies, and Table 48 (study characteristics) and Table 49 (resource use outcomes reported) for cost and resource use studies.[9] The included studies of the HRQoL review are summarised on pages 164 to 167 of the CS and the included cost and resource use studies are summarised on pages 172 to 173 of the CS.[9]

5.1.2 ERG critique of company's systematic review of cost-effectiveness evidence

The searches undertaken by the company were of a reasonable quality; however, there were some errors in the search strategies which, if executed as reported in the CS, may have led to results being missed or irrelevant results being retrieved. The economic evaluation and utility searches were designed for the Embase.com platform, whereas the ERG uses Ovid to access this database. On attempting to translate the searches to the Ovid interface, the ERG was unable to reproduce the searches exactly as presented within the CS.[9] For example, the same number of results was not retrieved with the string in line 48 of the economic evaluations search of Medline/EMBASE (*"patient*

acceptance of health care"); this was the case both when the search string was defined as a subject heading search and when it was defined as a phrase search. Some lines were confusing and appeared to include redundant repetition or duplication; for example, in line 1 of the EMBASE search: "*hidradenitis suppurativa'/exp OR 'hidradenitis suppurativa' OR 'hidradenitis' OR 'hidradenitis'/exp OR hidradenitis OR suppurativa OR hidradeniti* NEAR/1 suppurativ**". This appears to include exploded and unexploded forms of the same subject headings as well as searches for individual terms "suppurativa" and "hidradenitis" irrespective of whether they occur in proximity.

The ERG queried the source of the filters used to identify relevant studies (see clarification response,[17] question A5). In response, the company clarified that the cost-effectiveness searches were based on filters published by the Scottish Intercollegiate Guidelines Network (SIGN). It should be noted that SIGN filters are not always validated prior to publication, but the ERG acknowledges that this is a reputable source and one which is endorsed by NICE. The company did make some modifications to the SIGN filters (in the form of additional terms) but as they did not remove any of the existing terms this would not have had any detrimental effect on the search results.

The utility searches used terms taken from a paper by Papaioannou *et al*,[36] however, this contained several typographical errors (for example, in line 20 of the utility search on EMBASE, "*shorform thirtysix"* should have been written as *"shortform"* or *"short form"*; in line 23 of the same search, *"shortfrom sixteen"* should read *"shortform sixteen"*). When a corrected version of the utility search was run by the ERG, the number of results was not markedly different from that reported in the CS.[9]

The single search for "ahead of print" studies in PubMed was deemed by the ERG to be a valuable complement to the searches. However, the CS failed to include "PreMEDLINE" (MEDLINE-in-Process) citations in PubMed as well (i.e. those studies which have been printed and will be added to MEDLINE, but are not yet fully indexed). The ERG notes that there can be a backlog of several months between the print publication of an article and its appearance in MEDLINE.

The ERG also notes that imposing a restriction to include only studies published in English introduces a risk of missing relevant foreign language studies.

The ERG notes that the Cochrane systematic review[35] did not identify any cost-effectiveness studies. However, it should be noted that the Cochrane search had been designed to identify RCTs whereas this restriction on study design would not normally be applied for a cost-effectiveness search.

Despite the limitations described above, the ERG is satisfied that the company's searches are unlikely to have missed any relevant economic evaluation studies.

5.2 Description of the company's model

5.2.1 Health economic evaluation scope

As part of their submission to NICE,[9] the company submitted a fully executable health economic model programmed in Microsoft Excel. The scope of the company's economic analysis is summarised in Table 37. The company's model assesses the cost-effectiveness of adalimumab versus standard care for the treatment of patients with active moderate to severe HS who have had an inadequate response to conventional systemic therapy. The incremental health gains, costs and cost-effectiveness of adalimumab are evaluated over a lifetime horizon from the perspective of the UK NHS and PSS, although the ERG notes that no relevant PSS costs are included in the company's model. All costs and health outcomes are discounted at a rate of 3.5% per annum. Unit costs are valued at 2013/14 prices.

Table 37: Scope of company's health economic analysis

Population	Patients with active moderate to severe HS who have had an inadequate response to conventional systemic therapy
Intervention	Induction: Adalimumab 160mg at week 0, 80mg at week 2 and 40mg EW from week 4 onwards. Maintenance (week 12 responders* only): Adalimumab 40mg EW
Comparator	Standard care
Primary health economic outcome	Incremental cost per QALY gained
Perspective	NHS and PSS
Time horizon	Lifetime
Discount rate	3.5% per year

HS – hidradenitis suppurativa; EW – every week; NHS – Naitonal Health Service; PSS – Personal Social Services; QALY – quality-adjusted life year
**Including high response, response and partial response*

Population

All clinical efficacy data used in the company's model are based on the PIONEER I/II trials[18, 19] and the M12-555 OLE study,[20] hence the population represented within the model reflects the populations of patients recruited into these studies. At model entry, the population is assumed to be 35 years of age; this is broadly reflective of the mean age of patients in the PIONEER I/II trials. 65.9% of patients are assumed to be female. 44.7% are assumed to have moderate disease, whilst the remainder are assumed to have severe disease; this variable is used only to determine expected resource use in each health state. The economic analysis presented in the CS does not include any subgroup analyses.

Interventions

The intervention under consideration is adalimumab administered subcutaneously via an auto injection pen or pre-filled syringe.[12] Adalimumab is given at a dose of 160mg at week 0, 80mg at week 2 and 40mg EW from week 4 onwards. The company's model assumes that at 12 weeks, patients who achieve high response, response or partial response, based on the HiSCR measure, will

continue to receive adalimumab maintenance therapy. Patients who do not achieve at least a partial HiSCR response at 12-weeks are assumed to discontinue adalimumab treatment and subsequently receive standard care. During weeks 12-36 of the maintenance phase, patients are assumed to discontinue adalimumab at a constant rate irrespective of response status, based on the PIONEER I/II studies;[18, 19] thereafter differential withdrawal rates are applied to patients achieving at least a partial response and non-responders based on the OLE study.[20] This approach to handling adalimumab discontinuation is not fully justified in the CS. It is also noteworthy that according to the CS, the model assumes that from week 36 onwards, patients who are non-responders will continue to receive adalimumab and will discontinue if a further 12 weeks of adalimumab treatment fails to achieve at least a partial response (i.e. from week 48 onwards). The implementation of this continuation rule within the company's model is discussed in detail in Section 5.3.

Comparators

The comparator in the company's economic analysis is defined as "standard care." According to the CS[9] (page 139), surgery was not considered to be an appropriate comparator as surgery and adalimumab are not alternative or exclusive treatment choices. The CS also states that patients in the PIONEER trials were allowed surgery for symptom control and that an online survey of members of the UK Dermatology Trials Network and British Association of Dermatologists revealed that extensive surgery was generally used later in the treatment pathway.[9] However, the ERG notes that in response to a request for clarification (see clarification response,[17] questions A31 and B5), the company later stated that patients were not permitted to undergo either planned or unplanned surgery in the PIONEER I/II trials (see Section 4.2.1). The CS states that antibiotics were not considered to represent a relevant comparator, as antibiotics are typically used throughout the treatment pathway and these may be used concomitantly with adalimumab. The CS further notes that a comparison of adalimumab versus dapsone, retinoids and immunomodulators was not performed since UK clinical experts consulted in the preparation of the CS suggested that these therapies would currently be prescribed before adalimumab, noting also that there is currently a lack of efficacy evidence for these therapies in HS.[9] The company also considered that a comparison of adalimumab versus infliximab was not appropriate as infliximab is used in very specific subgroups of patients (for example, those who are very overweight) and such a comparison was not possible given the limited evidence base and heterogeneity between the infliximab and adalimumab trials. Clinical advisors to the ERG disagree that infliximab is only used in specific subgroups and a 2015 survey of UK clinicians suggests that that despite funding constraints, infliximab is currently used more widely in HS than adalimumab.[15]

Given the arguments presented by the company, the CS states that the relevant comparator is standard care, based on the placebo groups within the PIONEER I/II trials.[18, 19] The ERG notes that whilst the

placebo group data from PIONEER I/II are used to characterise the efficacy of standard care, resource use associated with standard care is instead based on expert opinion and is predominantly driven by the usage and costs of surgical inpatient admissions in both the adalimumab and placebo groups.

5.2.2 Description of the company's health economic model structure and logic

The general structure of the company's model is presented diagrammatically in Figure 3. The model adopts a simple Markov approach based on depth of HiSCR response. The model structure is comprised of five discrete mutually exclusive health states: (i) high response; (ii) response; (iii) partial response; (iv) non-response, and; (v) death. Table 38 presents the definitions for the categorisation of these different levels of response applied within the model. Within the adalimumab group, separate health states are used to define whether the patient is currently receiving adalimumab or whether they have discontinued and are currently receiving standard care. Within the standard care group, fewer discrete health states are required as all patients remain on standard care from the point of model entry. The model adopts a 2-weekly cycle length for the first 2 cycles and a 4-weekly cycle length thereafter; the CS[9] states that these cycle durations were selected to reflect the dosing schedule of adalimumab[12] and the timing of assessments within the PIONEER I/II trials.[18, 19] The model is evaluated over a total of 859 cycles (equating to a total time horizon of 66 years), although all surviving patients are forced into the death state at age 100 (at cycle 846, that is, after 65 years). A half-cycle correction is partially applied to account for the timing of events (see Section 5.3).

Figure 3 Company's health economic model structure (reproduced from CS[9] page 134)

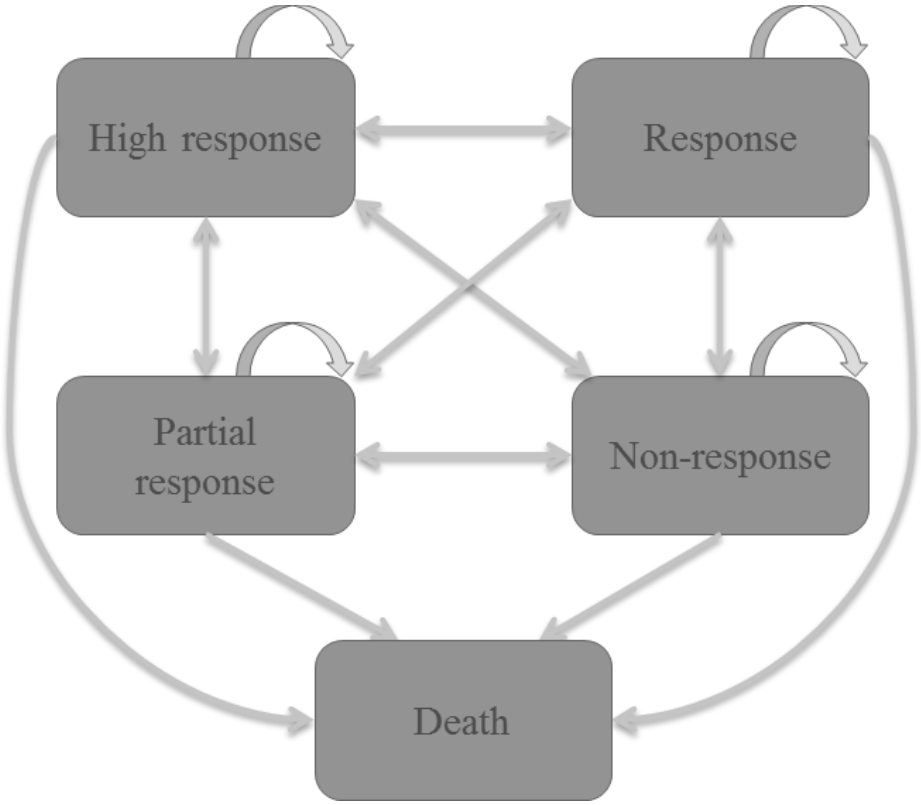


Table 38: Categorisation of depth of response in the company's model

HiSCR-based state definition	HiSCR-based state description
High response	At least 75% total AN count reduction, with no increase in abscesses or draining fistulas from baseline
Response	At least 50% but less than 75% AN reduction, with no increase in abscesses or draining fistulas from baseline
Partial response	At least 25% but less than 50% AN reduction, with no increase in abscesses or draining fistulas from baseline; or at least 25% AN reductions, with an increase in abscesses and/or draining fistulas
No response	Defined as less than 25% AN reduction

HiSCR - Hidradenitis Suppurativa Clinical Response; AN - abscess and inflammatory nodule

Model logic - adalimumab group

All patients enter the model in the no response state. Cycle-specific time-variant transition matrices based on cross-tabs of observed HiSCR outcomes pooled from the PIONEER I/II trials[18, 19] are used to determine health state populations for each cycle up to week 12 using simple matrix multiplication. All surviving patients remain on adalimumab induction therapy up to week 12. At 12 weeks, the model cohort separates into four discrete Markov submodels, each of which is defined according to the patients' 12-week HiSCR response (achievement of high response, response, partial response and no response at the end of induction); this distribution of patients is then used to define the initial vector for each of the maintenance submodels. Patients who do not achieve at least a partial response

during the first 12 weeks of adalimumab induction therapy are assumed to discontinue and subsequently receive standard care. Up to week 36, patients who have previously achieved an adalimumab induction response transit through the model health states based on cycle-specific time-variant probabilities based on cross-tabs of observed HiSCR outcomes for patients initially randomised to the adalimumab 40mg EW group who were subsequently re-randomised to receive adalimumab 40mg EW in the PIONEER I/II trials.[18, 19] Beyond week 36, a single time-invariant transition matrix is used to extrapolate the trajectory of patients through the health states, based on a GLM fitted to data from the M12-555 OLE study[20] (note that from week 48, the adalimumab non-responder discontinuation rate is increased, although the underlying HiSCR transition data remain the same). During each maintenance cycle, patients have a probability of discontinuing adalimumab treatment; these patients discontinue adalimumab, transit to the equivalent response state and subsequently receive standard care.

During weeks 12-36, cycle-specific time-variant transition matrices for patients discontinuing adalimumab therapy are based on cross-tabs of observed HiSCR outcomes for patients who were initially randomised to adalimumab 40mg EW induction therapy in PIONEER I/II[18, 19] who subsequently switched to placebo during the maintenance period (Period B). Beyond week 36, transition probabilities for adalimumab discontinuers are based on a single time-invariant transition matrix derived from a GLM fitted to the week 12-36 data for this population.

During each Markov cycle, a cycle-specific age-dependent probability of death is modelled to reflect the risk of all-cause mortality.

Model logic - standard care group

Within the standard care group, the model logic is similar to that for the adalimumab arm, albeit without the possibility of treatment discontinuation. Rather, patients continue treatment with standard care during induction and maintenance irrespective of their level of response. Patients enter the model in the no response state. Cycle-specific time-variant transition matrices based on cross-tabs of observed HiSCR response outcomes pooled from the PIONEER I/II trials[18, 19] are used to determine health state populations for each cycle up to week 12 using simple matrix multiplication. From week 12 to week 36, cycle-specific time-variant transition probabilities are based on cross-tabs of observed HiSCR outcomes for patients initially randomised to placebo in PIONEER II[19] who subsequently continued on placebo during Period B (note that data from PIONEER I[18] are not used in this portion of the model as the design of the trial did not allow for placebo group patients to be re-randomised to placebo during Period B). Beyond week 36, a single time-invariant transition matrix is applied for all subsequent Markov cycles based on a GLM fitted to the week 12-36 data from PIONEER II[19] described above.

During each Markov cycle, a cycle-specific age-dependent probability of death is modelled to reflect the risk of all-cause mortality.

Model logic – modelling health gains, costs and cost-effectiveness

Health utility is differentiated according to the patient's level of HiSCR outcome, with higher values applied to better response states. State-specific utilities are assumed to be the same for adalimumab and standard care. The model does not include disutilities to account for the impact of AEs in either treatment group.

The model includes the acquisition costs associated with adalimumab treatment (taking into account reductions in costs incurred due to imperfect compliance), inpatient admissions for HS surgeries, outpatient visits due to HS surgery, outpatient visits for wound care due to HS surgery, non-surgery related hospitalisations, routine outpatient visits, outpatient visits for wound care which are not due to HS surgery, A&E visits, and costs associated with managing AEs. The costs of adalimumab administration are assumed to be zero; the CS[9] states that adalimumab will be administered in the patient's home via AbbVie Care (homecare). The costs of concomitant pharmacological therapies used to manage HS are not included in the model.

The application of different transition matrices for the adalimumab and standard care groups leads to different trajectories of patients through the model's health states, thereby producing different profiles of costs and health outcomes for the two groups. Incremental cost-effectiveness is calculated as the difference in costs divided by the difference in QALYs for the two options.

Key structural assumptions employed within the company's model

The company's model makes the following structural assumptions:

- All patients enter the model in the no response health state.
- HRQoL and health care resource use are assumed to differ according to depth of response, defined according to the HiSCR measure.
- Patients who are non-responders at 12 weeks discontinue treatment with adalimumab. Patients who achieve at least a partial response continue to receive adalimumab as a maintenance therapy.
- During weeks 12-36, the probability of discontinuing adalimumab is assumed to be independent of the patient's current state of response. The company's model applies the discontinuation rates observed within Period B of the PIONEER I/II trials[18, 19] to patients irrespective of their level of HiSCR response.

- During weeks 40-48, the probability of discontinuing adalimumab is applied based on the discontinuation rate observed in the OLE study;[20] the same discontinuation rate is applied to patients in the high response, response and partial response states, with a higher discontinuation rate applied to non-responders.
- According to the CS,[9] the company's model assumes that patients who lose response to adalimumab will continue to receive a further 12-weeks of therapy before discontinuing adalimumab treatment. The ERG notes problems in the implementation of this continuation rule (see Section 5.3).
- Up to week 36, transition probabilities are assumed to be time-variant.
- Beyond week 36, transition probabilities are assumed to be time-invariant. Separate matrices are applied to (a) patients who are currently receiving adalimumab maintenance therapy; (b) patients who have received adalimumab but have subsequently discontinued, and (c) patients who are receiving standard care.
- The impact of AEs is reflected in the HiSCR-based utility values used in the model.
- Neither the disease itself, nor its treatment, is assumed to impact upon the life expectancy of patients.
- All adalimumab administration costs will be borne by the AbbVie homecare service.

5.2.4 Evidence used to inform the company's model parameters

Table 39 summarises the evidence sources used to inform the company's model parameters. These are discussed in more detail in the following sections.

Table 39: Summary of evidence sources used to inform the company's model parameters

Parameter/group	Source(s)
Patient characteristics (start age, proportion cohort who are female, percent of patients with moderate disease)	PIONEER I/II[18, 19]
HiSCR response transition probabilities	PIONEER I/II[18, 19] and M12-555 OLE study[20]
Probability of discontinuation	PIONEER I/II[18, 19] and M12-555 OLE study[20]
HiSCR state utilities	PIONEER II[18, 19]
Adverse event rates	PIONEER I/II[18, 19]
Health state resource use (inpatient/outpatient visits related/unrelated to HS surgery and A&E)	UK Physician Survey[22]
Adalimumab compliance (induction and maintenance periods)	PIONEER I/II[18, 19]
Adalimumab acquisition costs	BNF[37]
Resource use costs	NHS Reference Costs 2013/14[38]
Adverse event costs	NHS Reference Costs 2013/14,[38] and Curtis *et al*[39]

HS – hidradenitis suppurativa; HiSCR - Hidradenitis Suppurativa Clinical Response; OLE – open-label extension

5.2.4.1 Patient characteristics

The model includes three patient characteristics: patient age, gender and the percentage of the cohort that has moderate disease. Patients are assumed to enter the model aged 35 years; this is broadly in line with the mean age of patients in the PIONEER I/II trials[18, 19] (weighted mean=36.2 years). 65.9% of patients are assumed to be female; this is based directly on the characteristics of the populations recruited into PIONEER I/II.[18, 19] 44.7% patients are assumed to have moderate disease.

5.2.4.2 Transition probabilities

Transition probabilities – patients receiving standard care

Transition probabilities in the model were sourced from the PIONEER I/II trials[18, 19] and the M12-555 OLE study,[20] either using cross-tabs of observed trial outcomes relating to the specific time period in which the matrix is applied, or using GLMs fitted to outcomes from assessments at multiple timepoints. In the company's base case, a non-responder imputation rule was applied to all data from the PIONEER trials, whilst an LOCF imputation rule was used for the OLE data. The sources of the HiSCR transition matrices are summarised in Table 40; the full set of matrices employed within the company's base case analysis is presented in Appendix 1.

Within the standard care group, transition probabilities during weeks 0-12 are based directly on cross-tabs of observed HiSCR outcomes for patients who were initially randomised to the placebo groups during Period A within the PIONEER I/II trials.[18, 19] Cycle-specific time-variant transition matrices are used for the periods 0-2 weeks, 2-4 weeks, 4-8 weeks and 8-12 weeks.

For the period 12-36 weeks, transition probabilities for the standard care group are based on cross-tabs of observed HiSCR outcomes for patients who were initially randomised to the placebo group during Period A within the PIONEER II trial[19] who subsequently continued to receive placebo during Period B of the trial. Cycle-specific time-variant transition matrices are used for the periods 12-16 weeks, 16-20 weeks, 20-24 weeks, 24-28 weeks, 28-32 weeks and 32-36 weeks.

Beyond week 36, a single time-invariant transition matrix is used to model HiSCR outcomes for the standard care group. This matrix was based on a GLM of all transitions observed for patients initially randomised to placebo during Period A within the PIONEER II who subsequently continued on placebo during maintenance during Period B of the trial (i.e. the week 12-36 maintenance data described above).

Transition probabilities – patients receiving adalimumab

Within the adalimumab group, transition probabilities during the induction phase of the model are based on cross-tabs of observed HiSCR outcomes for patients initially randomised to the adalimumab

40mg EW groups during Period A within the PIONEER I/II trials.[18, 19] Cycle-specific time-variant transition matrices are used for the periods 0-2 weeks, 2-4 weeks, 4-8 weeks and 8-12 weeks.

For the period 12-36 weeks, transition probabilities for patients remaining on adalimumab are based on cross-tabs of HiSCR outcomes for patients initially randomised to adalimumab 40mg EW during Period A within the PIONEER I/II trials[18, 19] who were 12-week responders and were subsequently re-randomised to receive adalimumab 40mg EW during Period B of the trial. Cycle-specific time-variant transition matrices are used for the periods 12-16 weeks, 16-20 weeks, 20-24 weeks, 24-28 weeks, 28-32 weeks and 32-36 weeks.

Transitions during the period from week 36-48 within the company's model are based on a GLM of HiSCR outcomes for weeks 0-24 within the M12-555 OLE study.[20] According to the CS,[9] LOCF imputation was used as fewer than 50% of patients had 24-weeks of follow-up data. A single cycle-specific time-invariant matrix is used for cycles beginning at week 40 and week 44. Transitions for all cycles from week 48 onwards are based on the same GLM described above; however, the transition from the adalimumab no response state to the standard care no response state is raised to the power of three (i.e. cubed) and is assumed to reflect a situation whereby patients continue to receive a further 12-weeks of adalimumab treatment despite remaining non-responsive for this period. All other transitions out of the non-response state are adjusted accordingly. The justification for this mathematical approach is unclear from the CS (see Section 5.3).

Transition probabilities – patients who have discontinued adalimumab

For patients who previously received adalimumab and either failed to achieve a 12-week induction response, or who achieved an induction response but subsequently lost that response whilst receiving adalimumab maintenance therapy, transition probabilities during weeks 12-36 are based on cross-tabs of HiSCR outcomes for patients randomised to adalimumab 40mg EW in PIONEER I/II[18, 19] who switched to placebo during Period B (irrespective of whether they had previously achieved an induction response on adalimumab). Cycle-specific time-variant transition matrices are applied for the periods 12-16 weeks, 16-20 weeks, 20-24 weeks, 24-28 weeks, 28-32 weeks and 32-36 weeks.

Beyond week 36, a single time-invariant transition matrix is applied based on a GLM of all transitions observed for patients initially randomised to adalimumab in PIONEER I/II[18, 19] who subsequently switched to placebo during Period B (i.e. the 12-36 week data described above).

Table 40: Sources of transition probabilities used in the company's model

Matrix description	Source
Standard care – induction phase	
Week 0-12	Cross-tabs of outcomes based on pooling of patients initially randomised to the placebo groups within PIONEER I/II
Standard care – maintenance phase	
Week 12-36	Cross-tabs of outcomes for patients initially randomised to the placebo group in PIONEER II who subsequently continued on placebo during maintenance.
Week 36+	GLM based on 12-36 week data described above
Adalimumab – induction phase	
Week 0-12	Cross-tabs of outcomes based on pooling of patients initially randomised to adalimumab 40mg EW groups within PIONEER I/II.
Maintenance phase – adalimumab 12-week responders	
Week 12-36	Cross-tabs of adalimumab 40mg EW patients re-randomised to adalimumab 40mg EW after responding at 12-weeks in PIONEER I/II.
Week 36-48	GLM based on weeks 0-24 of M12-555 OLE study (including LOCF imputation as <50% patients had 24-weeks follow-up data).
Week 48+	Same as above except the probability of transiting from adalimumab no response state to standard care no response state is cubed.
Maintenance phase – adalimumab 12-week non-responders and subsequent discontinuers	
Week 12-36	Cross-tabs of patients randomised to adalimumab 40mg EW in PIONEER I/II who switched to placebo in the maintenance period (irrespective of whether they achieved an induction response on adalimumab).
Week 36+	GLM based on week 12-36 data described above

OLE – open-label extension; EQ- every week; LOCF – last observation carried forward

The ERG's concerns regarding the use of these pooled data from the PIONEER trials, particularly with respect to breaking randomisation, are detailed in Section 5.3.

5.2.4.3 Probability of discontinuing adalimumab treatment

The company's model applies two types of discontinuation for patients receiving adalimumab: (i) discontinuation due to a lack of induction response, and (ii) general discontinuation during the maintenance phase (this presumably reflects discontinuation due to loss of treatment response, although given that the discontinuation rate is greater than zero for the response states, this may include other reasons such as the incidence of treatment-related AEs).

In line with the wording of the marketing authorisation for adalimumab,[12] the model assumes that patients who do not achieve treatment benefit 12 weeks after initiating therapy will discontinue adalimumab. This discontinuation rule is applied in the model only to those patients who are in the no response state at the end of the induction phase (after 12 weeks). Patients who achieve full response,

response, or partial response according to the HiSCR measure are assumed to continue adalimumab as maintenance therapy.

During weeks 12-36 within the company's model, patients receiving adalimumab are assumed to discontinue adalimumab according to a constant discontinuation rate based on the rate observed in the PIONEER I/II trials.[18, 19] The company's model applies a constant 4-week probability of discontinuation during each cycle; this discontinuation rate is assumed to be independent of HiSCR state (see Table 41).

Beyond week 36, adalimumab discontinuation rates are based on data from the M12-555 OLE study,[20] calculated using the person-year approach. Patients were assumed to remain in their prior HiSCR response state until a change occurred and patients' health states at the time of discontinuation were used to derive response-specific discontinuation events.[9] According to the CS,[9] the estimation of discontinuation rates from the M12-555 OLE study was based on all adalimumab-treated patients who were week 12 responders, who received adalimumab during the maintenance periods of PIONEER I/II[18, 19] and who were subsequently enrolled into the OLE study.[20] Within the model, this is applied as the same discontinuation rate for the states of high response, response and partial response, with a higher discontinuation rate applied to patients in the no response state (see Table 41). The ERG notes that the company's assumption regarding the continued use of adalimumab for a further 12-weeks for non-responding patients is applied in the model as a 4-weekly probability of discontinuation of 0.56.

Table 41: Probability of discontinuing adalimumab maintenance therapy

Treatment period / states	Annual rate	4-weekly rate	Source
Maintenance period (weeks 12-36)			
All states (full response, response, partial response and no response)	20.48%	1.75%	PIONEER I/II[18, 19]
Maintenance period (week 36+)			
High response, response and partial response	7.47%	0.60%	M12-55 OLE study[20]
No response	44.99%	4.49%	

5.2.4.4 Health-related quality of life

Health utilities were based on EQ-5D valuations obtained from the PIONEER II trial.[19] The PIONEER I trial[18] did not include the use of a preference-based measure of HRQoL. The company assumes differential HRQoL according to depth of HiSCR outcome, based on data from week 12 and week 36 data from the trial (see Table 42). The model does not consider the separate impact of disutilities due to AEs or surgical intervention over and above those already reflected in the HiSCR-based health utility scores.

Table 42: Health utilities used in the company's model

Response state	Mean value	Upper 95% CI	Lower 95% CI
High response	0.782	0.816	0.746
Response	0.718	0.766	0.667
Partial response	0.576	0.639	0.512
Non-response	0.472	0.542	0.402

CI – confidence interval

5.2.4.5 Resource use and costs

The company's model includes the following resource components:

- Drug acquisition
- Health state resource use
    - Inpatient admissions for HS surgeries
    - Outpatient visits due to HS surgery
    - Visits to wound-care due to HS surgery (which are assumed by the company to take place in the outpatient setting)
    - Hospitalisations which are non-surgery related
    - Routine outpatient visits
    - Visits to wound-care not due to HS surgery (which are assumed by the company to take place in the outpatient setting)
    - A&E visits
- Management of AEs

Drug acquisition

The acquisition cost of 40mg adalimumab was assumed to be £352.14 per dose, based on the BNF list price.[37] The model also includes small reductions in the costs of adalimumab due to imperfect compliance. Patient compliance with the dosing schedule was estimated from the PIONEER I/II trials.[18, 19] During the induction period (weeks 0-12), adalimumab compliance was estimated to be 98.8%. Subsequently, adalimumab compliance during the maintenance phase was estimated to be slightly lower at 97.4%.

Health state resource use

The company's model includes health state costs associated with hospital admissions for HS surgery, hospital admissions for non-surgical reasons, outpatients visits related to surgery, outpatient visits unrelated to surgery, visits for wound care related to surgery, visits for wound care unrelated to surgery and visits to A&E departments. Estimates of resource use per cycle are assumed to be dependent on HiSCR outcome and are not directly related to whether the patient receives adalimumab

or standard care, although the ERG notes that the model assumes that spending more time in better response states results in lower estimates of total resource use.

The company's model does not use data from the PIONEER I/II studies to inform estimates of resource use; instead, the company undertook a survey of 40 physicians who actively treat patients with moderate to severe HS in the UK.[22] The survey involved the separate elicitation of estimates of resource use for patients with moderate disease and for those with severe disease. Estimates elicited from the respondents were aggregated and the mean of the responses provided were used as inputs into the model;[17] aggregate estimates of resource use per year were weighted according to the proportion of patients with moderate or severe disease in the PIONEER I/II trials[18, 19] and then adjusted to reflect the 2- or 4- week cycle length used in the model. All resource estimates were valued using NHS Reference Costs 2013/14.[38] Annual health state resource use estimates used in the model are summarised in Table 43.

Table 43: Resource use by health state

Resource	Resource use (mean number of units per year)				Cost per event	Source
	High response	Response	Partial response	Non-response		
Hospital admissions for HS surgery	0.13	0.22	0.54	0.80	£5,488.32	NHS Reference Costs 2013/14[38] – elective inpatient, JC41Z (major skin procedures)
Hospital admissions not for HS surgery	0.11	0.23	0.29	0.45	£2,202.14	NHS Reference Costs 2013/14[38] – elective inpatient, weighted mean of codes JD07D and JD07K
Outpatient visits associated with surgery	0.22	0.35	0.67	0.94	£97.63	NHS Reference Costs 2013/14[38] – outpatient service codes 330
Outpatient visits not associated with surgery	3.10	3.51	4.44	4.68	£97.63	
Outpatient wound care visits associated with surgery	0.12	0.17	0.40	0.85	£97.63	
Outpatient wound care visits not associated with surgery	0.67	0.47	0.64	0.45	£97.63	
A&E visits	0.12	0.20	0.47	0.57	£123.67	NHS Reference Costs 2013/14[38] – weighted mean of total HRG for all emergency medicine procedures

HS – hidradenitis suppurativa; A&E – accident and emergency; HRG – Healthcare Resource Group

Adverse event rates and costs

The company's model includes the costs associated with treatment-emergent adverse events (TEAEs) if they occurred in at least 5% of patients in the PIONEER I/II trials.[18, 19] Separate AE estimates were applied for the induction and maintenance phases of the model. During the maintenance phase, AE rates were estimated separately for patients receiving adalimumab, for those who have discontinued adalimumab and for patients receiving standard care. The source of the proportion of AEs which were classified as being severe is unclear from both the CS and the model. The costs associated with managing AEs were valued using 2013/14 NHS Reference Costs,[38] together with GP costs sourced from the PSSRU.[39] AE rates and costs used in the company's model are summarised in Tables 44 and 45, respectively.

Table 44: Annual AE rates assumed in the company's model based on PIONEER I/II[18, 19]

Adverse event	Induction period		Maintenance period			Percentage events severe
	Adalimumab	Standard care	Adalimumab	Standard care	Following discontinuation	
Headache	0.486	0.505	[redacted]	[redacted]	[redacted]	3%
Nasopharyngitis	0.250	0.365	[redacted]	[redacted]	[redacted]	1%
Influenza	0.069	0.084	[redacted]	[redacted]	[redacted]	5%
Gastroenteritis	0.069	0.056	[redacted]	[redacted]	[redacted]	6%
Viral gastroenteritis	0.000	0.028	[redacted]	[redacted]	[redacted]	20%
Diarrhoea	0.167	0.084	[redacted]	[redacted]	[redacted]	0%
Upper respiratory tract infection	0.180	0.182	[redacted]	[redacted]	[redacted]	0%
Bronchitis	0.028	0.084	[redacted]	[redacted]	[redacted]	0%
Toothache	0.028	0.028	[redacted]	[redacted]	[redacted]	0%
Hidradenitis*	0.291	0.575	[redacted]	[redacted]	[redacted]	11%

** As discussed in Section 4.2.4, exacerbations of HS were classified as an AE*

Table 45: Costs associated with managing AEs

Adverse event	Cost (severe)	Cost (mild/moderate)	Cost (per event)*	Source
Headache	£674.21	-	£20.03	NHS Reference Costs 2013/14,[38] - weighted mean of total HRGs for codes AA31C, AA31D and AA31E.
Nasopharingitis	£908.28	-	£12.62	NHS Reference Costs 2013/14,[38] - weighted mean of total HRGs for codes WA06A, WA06B and WA06C.
Influenza	£908.28	-	£43.25	
Gastroenteritis	£1,468.01	£46.00	£125.00	Severe - NHS Reference Costs 2013/14,[38] - weighted mean of total HRGs for codes FZ91A to FZ91M. Mild/moderate - PSSRU[39] - GP visit lasting less than 11.7 minutes.
Viral gastroenteritis	£1,345.99	£46.00	£306.00	Severe - NHS Reference Costs 2013/14[38] - weighted mean of total HRGs for codes FZ36G to FZ36Q. Mild/moderate - PSSRU[39] - GP visit lasting less than 11.7 minutes.
Diarrhoea	-	£46.00	£46.00	Mild/moderate - PSSRU[39] - GP visit lasting less than 11.7 minutes.
Upper respiratory tract infection	-	£147.22	£147.22	NHS Reference Costs 2013/14[38] - weighted mean of outpatient codes 340 and 341.
Bronchitis	-	£147.22	£147.22	
Toothache	-	-	-	Assumed to be zero
Hidrahentitis	-	-	-	Assumed to be zero

NHS – National Health Service; PSSRU – Personal Social Services Research Unit; GP – general practitioner; HRG – healthcare resource group
** Weighted by severity*

5.2.5 Methods for model evaluation

The CS[9] presents the results of the economic evaluation in terms of the incremental cost per QALY gained for adalimumab versus standard care. The base case results are presented deterministically, based on point estimates of parameters. The CS[9] also includes the results of probabilistic sensitivity analysis (PSA) and deterministic sensitivity analyses (DSA). The results of the PSA are presented in the form of cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs), based on 5,000 Monte Carlo simulations. The CS[9] presents the results of the DSA in the form of a tornado diagram. A number of alternative scenario analyses are presented to explore the impact of truncating the model time horizon, using different discount rates for costs and health gains, using alternative data sources for transition probabilities and discontinuation rates (including removing the assumption of 12-weeks continued use of adalimumab in non-responders), using alternative imputation rules for missing data and varying the adalimumab compliance rate.

5.2.6 Cost-effectiveness results presented within the CS

5.2.6.1 Base case cost-effectiveness results

Table 46 presents the company's base case results. Based on the probabilistic version of the company's base case model, adalimumab is expected to produce an additional 1.02 QALYs at an additional cost of £[redacted] compared with standard care; the ICER for adalimumab versus standard care is expected to be £[redacted] per QALY gained. The results of the deterministic model are similar, with adalimumab yielding an ICER of £[redacted] per QALY gained compared with standard care.

Table 46: Company's base case cost-effectiveness results

Probabilistic model*					
Option	**QALYs**	**Costs**	**Incremental QALYs**	**Incremental costs**	**Incremental cost per QALY gained**
Adalimumab	12.63	£[redacted]	1.02	£[redacted]	£[redacted]
Standard care	11.61	£128,674	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Incremental QALYs**	**Incremental costs**	**Incremental cost per QALY gained**
Adalimumab	12.61	[redacted]	1.00	£[redacted]	[redacted]
Standard care	11.61	£128,541	-	-	-

QALY – quality-adjusted life year
** derived from company's model*

5.2.6.2 Probabilistic sensitivity analysis results

Figures 4 and 5 present the cost-effectiveness plane and CEACs for adalimumab versus standard care, respectively. Assuming a WTP threshold of £20,000 per QALY gained, the company's base case model suggests that the probability that adalimumab produces more net benefit than standard care is approximately [redacted]. Assuming a WTP threshold of £30,000 per QALY gained, the probability that adalimumab produces more net benefit than standard care is approximately [redacted].

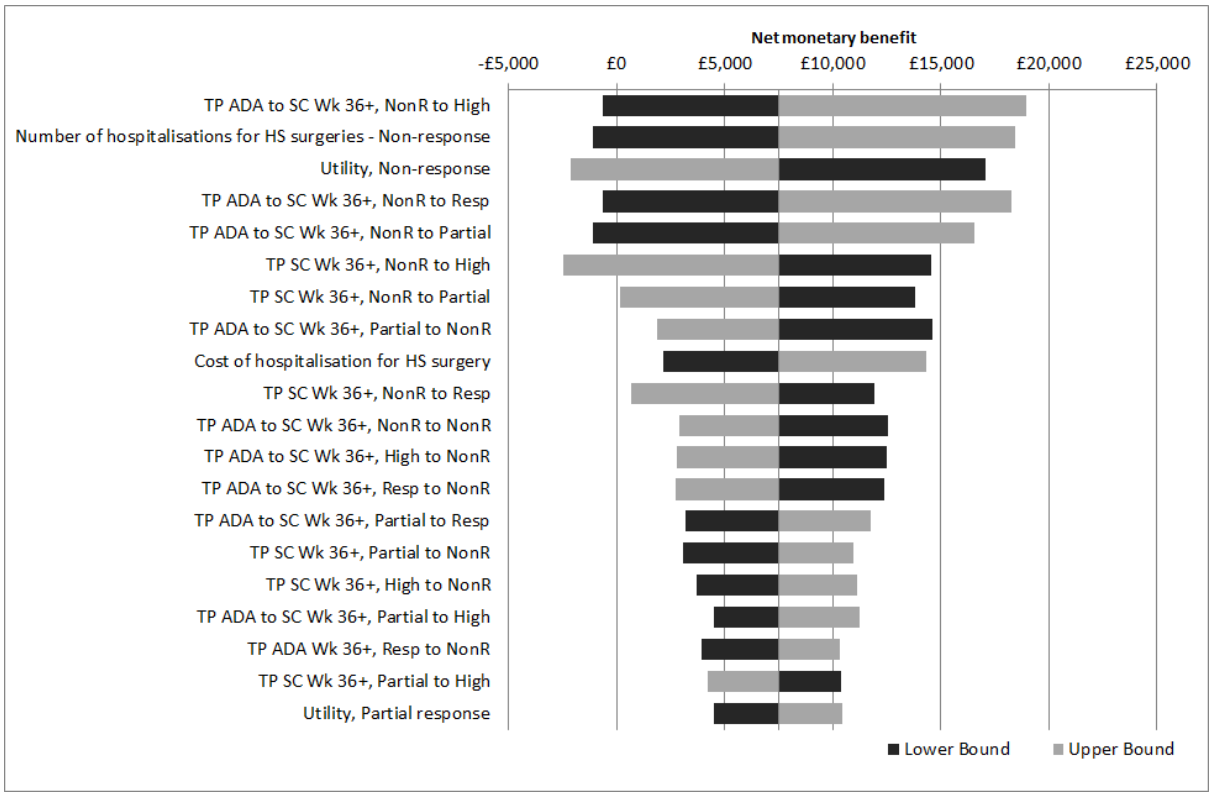
Figure 4 Cost-effectiveness plane (redrawn by the ERG)

Figure 5: Cost-effectiveness acceptability curves (redrawn by the ERG)

5.2.6.3 Deterministic sensitivity analysis results

Figure 6 presents the results of the company's DSA in the form of a tornado diagram using net monetary benefit as the economic outcome measure assuming a WTP threshold of £30,000 per QALY gained.

Figure 6: Deterministic sensitivity analysis - tornado diagram (redrawn by the ERG)



The company's DSA indicates that the key groups of uncertain parameters within the model relate to the long-term transition probabilities (from week 36 onwards, based on the GLMs), the number of HS surgeries assumed in the no response state, and the utility value applied to the no response state. When considered individually, the bounds of the 95% confidence intervals for these parameters produce a negative net benefit for adalimumab versus standard care at a WTP threshold of £30,000 per QALY gained.

5.2.6.4 Scenario analysis results

Table 47 presents the results of the company's scenario analyses.

Table 47: Company's scenario analysis results

Scenario	Incremental – adalimumab versus standard care		
	QALYs	Cost	ICER
Company's base case (deterministic)	1.00	[redacted]	[redacted]
Time horizon 20 years	0.74	[redacted]	[redacted]
Time horizon 30 years	0.86	[redacted]	[redacted]
Discount rate=0%	1.78	[redacted]	[redacted]
Discount rate=5%	0.84	[redacted]	[redacted]
Model based on PIONEER II[19] only for both adalimumab and standard care arms during induction, PIONEER II[19]only for both adalimumab and standard care arms during maintenance	0.90	[redacted]	[redacted]
LSCF extrapolation	0.90	[redacted]	[redacted]
Mean transition probability extrapolation	1.09	[redacted]	[redacted]
Transition probabilities for the adalimumab arm after week 36 estimated based on PIONEER I/II trial data[18, 19]	Model analysis unclear*		[redacted]
LOCF missing value imputation	1.32	[redacted]	[redacted]
Response specific discontinuation rates for adalimumab during weeks 12-36 from PIONEER I/II[18, 19]	0.99	[redacted]	[redacted]
Response specific discontinuation rates for adalimumab for week 36+ from PIONEER I/II[18, 19]	0.94	[redacted]	[redacted]
Discontinuation rate of adalimumab non-responders after week 36 based on OLE study[20]	1.34	[redacted]	[redacted]
Maintenance compliance rate of adalimumab (week 12+) equal to 100%	1.00	[redacted]	[redacted]

LOCF – last observation carried forward; LSCF – last state carried forward; OLE – open-label extension; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

Across most of the scenarios considered, the ICER for adalimumab versus standard care remains below £30,000 per QALY gained and in some instances the ICER is below £20,000 per QALY gained. The ICER for adalimumab is greater than £30,000 per QALY gained in the following scenarios: (i) when the time horizon is truncated to 20 years; (ii) when the model uses only data from PIONEER II;[19] (iii) when the LSCF imputation rule is used, and; (iv) when the discontinuation rate for adalimumab non-responders after week 36 is based on the OLE study.[20] The ERG notes that removing the company's approach to modelling 12 further weeks of adalimumab in non-responding patients, and instead basing this on the observed estimate in the OLE study, increases the ICER to £[redacted] per QALY gained.

5.3 Critical appraisal of the company's health economic analysis

5.3.1 Methods for reviewing the company's economic evaluation and health economic model

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic evaluation and the underlying health economic model upon which this was based. These approaches included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists[40, 41] to critically appraise the company's model and analysis.
- Scrutiny of the company's model by health economic modellers and discussion of issues identified amongst the members of the ERG.
- Double-programming of the deterministic version of the company's model to fully assess the logic of the company's model structure, to draw out any unwritten assumptions and to identify any apparent errors in the implementation of the model.
- Examination of correspondence between the description of the model reported within the CS[9] and the company's executable model.
- Replication of the base case results, PSA and scenario analysis presented within the CS.[9]
- Where possible, checking of parameter values used in the company's model against the original data sources.
- The use of expert clinical input to judge the clinical credibility of the company's economic evaluation and the assumptions underpinning the model.

5.3.2 Summary of main issues identified within the critical appraisal

Box 2 summarises the main issues identified within the ERG's critical appraisal of the company's economic analysis. These issues are discussed in further detail in the subsequent sections.

Box 2: Main issues identified within the critical appraisal of the company's model

1. Deviations from the NICE Reference Case
2. Disconnect between evidence for the efficacy and cost of the comparator
3. Modelling treatment effects according to depth of response
4. Modelling treatment continuation rules
5. Potential overestimation of number of surgical inpatient admissions
6. Uncertainty surrounding transition probabilities
7. Appropriateness of pooling data from PIONEER I and II trials
8. Conceptual inconsistency in handling time-variance in transition probabilities
9. Potential bias in the use of OLE study data for long-term adalimumab responders
10. Model errors and other issues surrounding model implementation

(1) Deviations from the NICE Reference Case

Table 48 summarises the extent to which the company's model adheres to the NICE Reference Case.[21]

Table 48: Adherence of the company's economic analysis to the NICE Reference Case

Element	Reference case	ERG comments
Defining the decision problem	The scope developed by NICE	The scope of the company's model is generally in line with the final NICE scope.[8] The population considered directly relates to the populations of the PIONEER I/II trials.[18,19] Clinical advice received by the ERG suggests that this is likely to be reflective of the UK HS population who may be eligible for treatment using adalimumab.
Comparator(s)	As listed in the scope developed by NICE	The final NICE scope[8] defines the comparator as *"established clinical management without adalimumab."* Whilst the placebo arms of the PIONEER I/II trials[18, 19] are used to model efficacy, costs are based on estimates of surgical and non-surgical resource use from an online survey of UK physicians undertaken by the company.[22] It is unclear whether the elicited survey estimates as applied in the model truly reflect standard care in the UK. The ERG notes that there is no established pathway of care for patients with active moderate to severe HS after the failure of systemic conventional therapy in the UK.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health gains for patients are modelled in terms of QALYs gained.
Perspective on costs	NHS and PSS	The CS[9] states that an NHS and PSS perspective was adopted, although no relevant PSS costs are included in the company's model. Excluding the management of certain AEs, all costs are assumed to be incurred in the secondary care setting.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The company's economic evaluation takes the form of a cost-utility analysis. The results of the analysis are presented in terms of the incremental cost per QALY gained for adalimumab versus standard care.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The model adopts a lifetime horizon. Scenario analyses are also presented for shorter time horizons (20 years and 30 years).
Synthesis of evidence on health effects	Based on systematic review	The model is largely based on data collected within the PIONEER I/II trials.[18, 19] Long-term extrapolation of transitions for adalimumab responders beyond 36 weeks are based on the M12-555 OLE study.[20] The company's use of arm based summaries to aggregate data from the PIONEER I/II trials breaks randomisation and may lead to bias.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Health utilities were based on EQ-5D estimates from the PIONEER II trial.[19]

Element	Reference case	ERG comments
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Health utilities were based on EQ-5D estimates from the PIONEER II trial.[19]
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to estimated QALY gains.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Resource use estimates according to depth of response were elicited via a survey of UK physicians. Cost estimates were based on the BNF,[37] NHS Reference Costs[38] and the PSSRU.[39]
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	All costs and QALYs are discounted at a rate of 3.5%

QALY – quality-adjusted life year; EQ-5D – Euroqol 5D; NHS – National Health Service; PSS – Personal Social Services; NICE – National Institute for Health and Care Excellence

The company's model is generally in line with the final NICE scope.[8] The company's model is principally based on data collected within the PIONEER I/II trials and therefore reflects the populations of patients recruited into these trials. Clinical advisors to the ERG considered that these patients generally reflect the population who may be considered eligible for treatment using biologic therapy in England. The ERG notes that there is a lack of clarity within the CS with respect to the comparator, particularly since standard care is assumed to relate only to secondary care resource use, and the costs of pharmacological therapies are not included in the model. The ERG also has some concerns regarding whether the elicited estimates of surgical resource use applied in the model truly reflect the typical experience of patients with HS in England. Clinical advisors to the ERG were satisfied that the types of resource use included in the model were generally relevant, but noted that some treatments (e.g. wound dressings, where needed) may be given in a primary care setting and that some patients will be prescribed antibiotics by their GPs for several years, yet these costs are not included in the model. The clinical advisors also noted that a comparison of adalimumab against infliximab may have been useful, but could not have been based on the HiSCR measure. The time horizon, perspective and discount rate used in the company's analysis are appropriate. No additional

equity weighting is applied to estimated QALY gains. Issues surrounding the appropriateness of the company's approach to modelling treatment benefits are detailed in the subsequent sections.

(2) Disconnect between evidence for the efficacy and cost of the comparator

The CS[9] highlights that until recently there were no published guidelines to help clinicians and patients determine potential treatment choices. The CS[9] also states that there are no licensed effective therapies for the treatment of HS in the UK and that various pharmacological therapies are commonly used off-label (including antiseptics, NSAIDs, immunosuppressants, corticosteroids, anti-androgens, retinoids and TNF-α inhibitors). The ERG considers that this in itself is an insufficient justification for excluding these options as potential comparators. The CS does however also note that there is limited robust evidence to demonstrate the efficacy of any of these therapies in the management of HS.

The CS argues that surgery does not represent a relevant comparator for adalimumab since adalimumab and surgery are not alternative or exclusive treatment choices and because within the PIONEER I/II trials,[18,19] patients were allowed to undergo surgery to control symptoms. Page 139 of the CS states that *"Patients receiving ADA in the clinical trials were allowed surgery for symptom control",* whilst the company's clarification response[17] suggests the opposite, stating that *"Surgery was not permitted in the PIONEER I and II studies per protocol. As such a change in the number of surgeries could not be observed"* and a second clarification response included incision and drainage in a list of permitted co-interventions (see clarification response,[17] question A12) and this was reported to have taken place in the M10-467 trial[25] (see Section 4.2.1). Consequently, the ERG remains unclear whether surgery was, or was not, allowed in the PIONEER trials. The CS further argues that antibiotics, dapsone, retinoids, immunomodulators and other biologics are not suitable comparators to adalimumab. As such, the CS argues that the main comparator for the analysis is standard care, as represented by the placebo arms in the PIONEER I/II trials.[18,19]

However, the ERG notes that there is a disconnect with respect to how the treatment benefits and costs of standard care are represented within the company's model. The progression of patients receiving standard care, which is characterised in terms of transitions between HiSCR-defined health states, are based directly on either cross-tabs of observed trial data or GLMs fitted to observed HiSCR outcomes for patients randomised to the placebo groups within the PIONEER I/II trials.[18, 19] In contrast, resource use estimates are instead based on the predicted use of surgery-related and non-surgery-related secondary care resources (inpatient admissions, outpatient appointments and A&E visits), estimated from a survey of UK physicians.[22] These estimates of resource use are assumed to differ according to depth of HiSCR response, which in turn, produces different health state costs for each HiSCR state. Higher resource use is assumed for patients achieving a weaker response or no response (see Table 43).

The ERG has several concerns regarding this approach:

(i) In general, the ERG considers it inconsistent to model the benefits of treatment using one source and the resources required to achieve these benefits using another source. From the evidence presented within the CS,[9] it is unclear whether the company's modelled predictions of overall resource use reflects the experience of patients enrolled into the PIONEER I/II trials.[18, 19] Whilst the CS[9] (page 118) makes the assertion that adalimumab may lead to the delay or reduction in the need for surgery, and this assertion flows through to the company's model to produce surgery-related cost savings for adalimumab, this potential treatment benefit is not substantiated by empirical evidence presented in the CS. In response to a request from the ERG for clarification on this matter (see clarification response,[17] question B5), the company undertook a *post hoc* analysis using combined data from the PIONEER I/II studies to assess the use of incision and drainage procedures and intralesional steroid injections as surrogate markers for surgical interventions. The results of the company's analysis showed that at week 12, a greater proportion of patients who received adalimumab, compared with placebo, experienced elimination of both draining fistulas (33% vs 19%; $p<0.001$) and non-draining fistulas (15% vs 9%; $p=0.017$). The ERG notes however that it is unclear whether this fully reflects an overall reduction in surgery, particularly inpatient surgical admissions, which are a key cost driver in the company's model (see critical appraisal point 5). Clinical advisors to the ERG noted that whilst the use of adalimumab could reduce the extent to which limited surgical procedures are required for patients with previously uncontrolled disease, in some instances adalimumab may be used as a preadjuvant "bridge" to more definitive surgery, thereby increasing the use of surgery. Consequently, there remains uncertainty regarding whether adalimumab will increase or decrease the lifetime costs of surgery for HS patients.

(ii) The company's approach ignores the costs of other concomitant pharmacological therapies. The CS[9] (page 173) claims that the costs of conventional therapies were likely to have been lower in the adalimumab groups relative to the placebo groups in PIONEER I/II,[18, 19] and that excluding these costs from the model is "conservative." This assertion is not however supported by any evidence within the CS.[9] In response to a request for clarification by the ERG (see clarification response,[17] question B9), the company provided estimates of the use of concomitant medications in >5% patients in Period A of the PIONEER I/II trials. On the basis of this additional information provided by the company, the ERG is satisfied that the proportions of patients receiving each therapy are broadly similar between the adalimumab and placebo groups. However, this information relates only to the first 12 weeks of treatment within the RCTs; it remains unclear whether the inclusion of the costs of concomitant medications would substantially impact upon the cost-effectiveness of adalimumab over a lifetime horizon.

(iii) Specifying resource use according to depth of response, whereby resource use is lower for better HiSCR outcomes, may induce a spurious (or at least an unproven) relationship within the model between the time spent receiving adalimumab, the time spent in a state of better response and resource use avoided. Within the PIONEER I/II trials,[18, 19] it is unclear whether the use of health care resources was lower in the adalimumab groups.

(iv) The ERG remains unclear whether estimates of surgery-related and non-surgery-related resource use could or could not have been drawn directly from the PIONEER I/II trials.[18, 19] If certain types of surgery were indeed allowed in the PIONEER I/II trials, using estimates from this source may have allowed for a greater degree of consistency between the modelled estimates of QALYs gained and the resources required to generate such health gains. The ERG notes however that if this information was not adequately collected in the trials, the company would have had no alternative but to use an alternative evidence source to inform resource use estimates within the model.

(3) Modelling treatment benefits according to depth of response

As detailed in Section 5.2, the company's model structure is based on four main health states, defined according to the depth of HiSCR response: (i) high response; (ii) response; (iii) partial response, and; (iv) no response. With respect to the company's decision to adopt this depth-based structure, the CS[9] (page 135) states: *"Preliminary analyses of the EuroQol (EQ-5D) data collected in the PIONEER II trial[30] indicated that there was a statistically significant difference in the utility values between the high response and response state, and between the values of the partial response and non-response health states. Therefore, to better evaluate the impact of treatment on HRQOL, the analysis considered four separate response health states."*

With the exception of data from the poster presented at WCD 2015 reported in Section 4.2.7.4 of the CS,[9] the company's systematic review of clinical evidence reports only on the full HiSCR measure, as pre-specified in the final statistical analysis plans of the PIONEER I/II trials. This leads to a degree of discordance between the evidence presented in the company's clinical efficacy review and the way in which evidence from the included studies is used within the company's model. Consequently, the ERG requested further clarification regarding the justification of the company's model structure (see clarification response,[17] question B2). In response, the company stated: *"...the selection of four response health states was due to the following considerations: 1) there were statistically significant differences in the response rates of adalimumab and placebo in "high response", "response" and "non-response", and 2) the utility and resource use differed across the four response health states 3) a post-hoc analysis of the PIONEER I and II studies identified a population where continued treatment with ADA could be beneficial. Therefore, to evaluate the cost-effectiveness of adalimumab, it was reasonable to segregate the model into four response health states."*

The ERG considers that disaggregating the full HiSCR measure (which is a dichotomous outcome) according to depth of response (which is an ordered categorical outcome), represents a *post hoc* analysis of a pre-planned endpoint. The ERG also notes that the Kimball *et al*[29] validation study of the HiSCR measure relates only to the full HiSCR threshold (≥50% reduction in ANs, with no increase in abscesses or draining fistulas from baseline). Kimball *et al*[29] report that patients with worsening disease or minimal improvement in ANs (<30% reduction) did not have a meaningful improvement on the DLQI and reported some worsening in pain despite demonstrating some improvement in total work impairment and total activity impairment (see Figures 7 and 8). Kimball *et al* also report that no substantial incremental benefits were observed on patient reported outcomes beyond the ≥50% AN reduction threshold for HiSCR.

Figure 7: Change in pain VAS and DLQI (reproduced from Kimball *et al*[29])

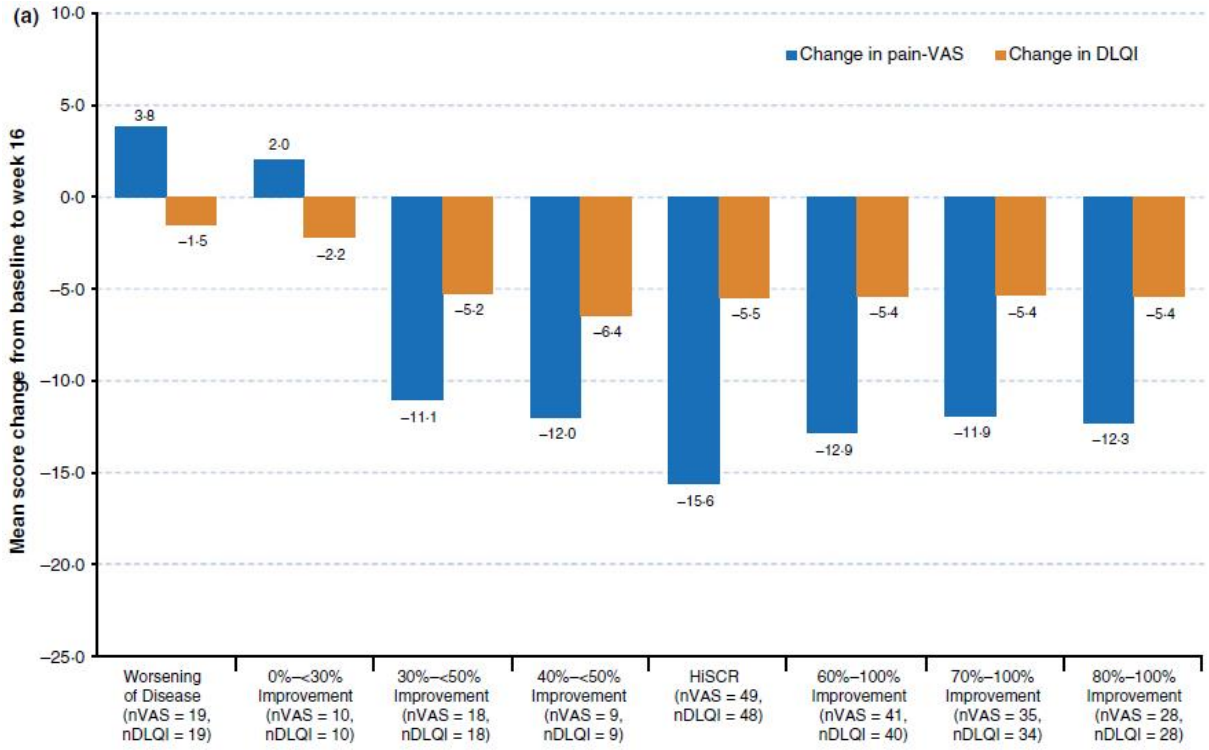
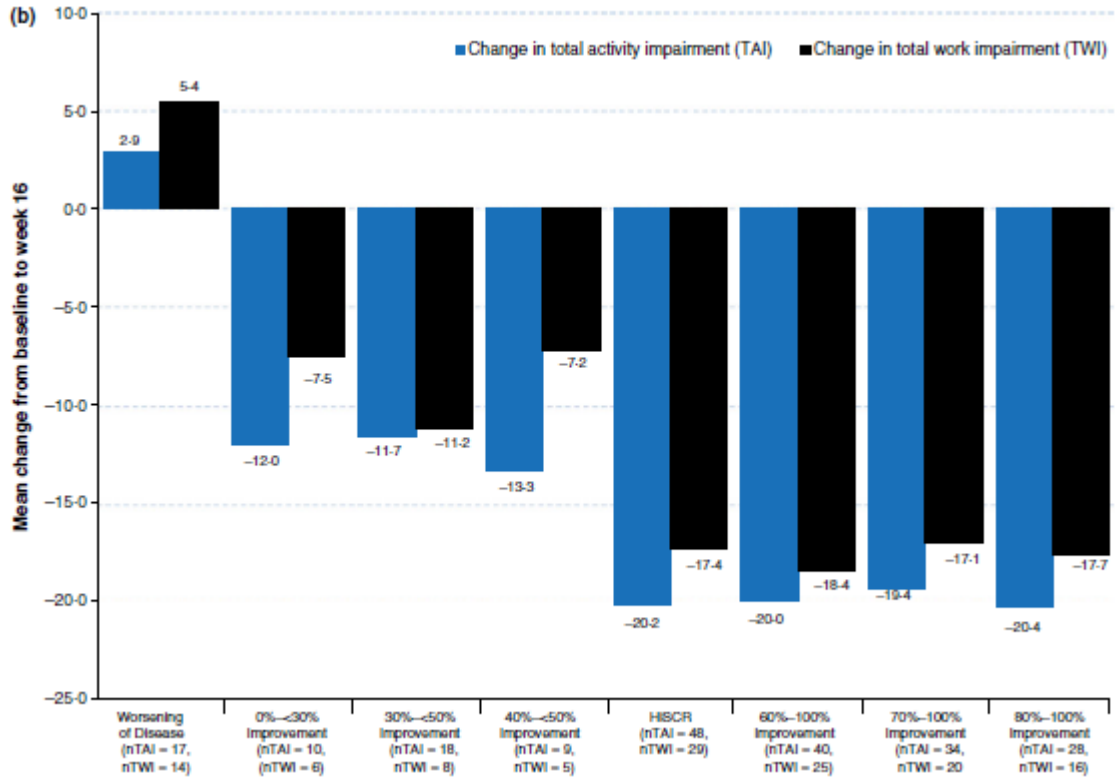


Figure 8 Change in Work Productivity and Activity Impairment Questionnaire – Total Activity Impairment and Total Work Impairment (reproduced from Kimball *et al*[29])



With respect to the PIONEER II EQ-5D valuations, the CS states that the differences in HRQoL between no response and partial response, and high response and response, were statistically significant, but does not provide evidence to support this statement. In response to a request for clarification, the company provided *p*-values for these comparisons which confirm the company's original claim ($p<0.05$ for both comparisons). Whilst the instruments used in these two sets of analyses are not the same, the apparent distinction between health states evident in the *post hoc* analysis of the PIONEER II EQ-5D data does not appear to be entirely consistent with the analyses reported by Kimball *et al.*[29]

From the perspective of model structuring, splitting the HiSCR outcome data according to depth of response within the model would allow for a more granular representation of EQ-5D benefits over time, and in principle, the consideration of alternative discontinuation rules for patients achieving different levels of treatment benefit (although this has not been done). There are however also some negative consequences associated with this approach: (i) the available efficacy data are "stretched" across four rather than two states, hence several cells in the transition matrices are populated with small numbers of patients (see Appendix 1); (ii) patients who would be classed as partial responders in the model would have been considered to be non-responders in the clinical analysis based on the

pre-specified HiSCR threshold, thereby producing some inconsistency in the interpretation of the data from the PIONEER I/II trials,[18, 19] and; (c) the definition of health states in the model is not consistent with the aims and findings of the Kimball *et al* validation study.

The ERG notes also that within the company's model, the criterion for continuing treatment with adalimumab at 12-weeks and during subsequent maintenance therapy requires patients to achieve only a partial response, rather than a full HiSCR response. Had the company's model been structured according to the full HiSCR ≥50% AN reduction threshold, this would have necessarily led to the use of different treatment continuation rules during induction and thereafter, as only patients achieving and maintaining this level of response would continue adalimumab therapy. Clinical advisors to the ERG noted that it was unclear whether patients achieving a partial HiSCR response would obtain a clinically meaningful benefit sufficient to warrant continuing adalimumab treatment. The advisors also noted that some patients may achieve a level of benefit which is only slightly below the threshold for response, whilst at the other end of the spectrum, some patients may accrue little benefit from continued adalimumab treatment.

Based on the definition of health states and the treatment continuation rules assumed in the company's model, it could be argued that the model implicitly suggests that the 50% AN reduction threshold determined in the Kimball validation study, and later pre-specified as the primary endpoint in the PIONEER I/II trials, has been set at the wrong level for clinical practice.

Given the above discussion, there are arguments both for and against structuring the model according to depth of HiSCR response. The ERG therefore considers this to be an area of structural uncertainty. In light of this, the ERG requested that the company undertake a separate analysis in which the modelled costs and health outcomes for adalimumab and standard care were based only on HiSCR responders and non-responders at the ≥50% AN reduction threshold (see clarification response,[17] question B2). In response, the company stated the following:

"Unfortunately due to time constraints AbbVie was not able to make structural changes to the cost effectiveness model (ie. change the structure from a 4 model response state to a 2 model response state), however AbbVie was able to provide a health economic analysis which would use only the outcomes of response or no response as per the PIONEER trials by implementing the following changes to the existing model structure:

1. *Assign the same utility value to the High response and Response (HiSCR responders as per the PIONEER trials) health states based on a re-analysis of the EQ-5D data at week 12 and 36 from the PIONEER II trial*

2. *Assign the same utility value to the Partial response and non-Response (HiSCR non-responders as per the PIONEER trials) health states based on a re-analysis of the EQ-5D data at week 12 and 36 from the PIONEER II trial*
3. *Assign the same resource use cost to the High response and Response (HiSCR responders as per the PIONEER trials) health states (average the cost across the two health states)*
4. *Assign the same resource use cost to the Partial response and non-Response (HiSCR non-responders as per the PIONEER trials) health states (average the cost across the two health states)*
5. *Assign same week 36+ discontinuation rate for partial responders as per non responders based on discontinuation rate using OLE"* (Clarification response,[17]question B2).

The results of the company's re-analysis of the model based on the ≥50% AN reduction threshold are presented in Table 49.

Table 49: Results of company's analysis based on HiSCR response/no response (deterministic model, taken from company's clarification response[17])

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	13.12	£[redacted]	0.69	£[redacted]	£[redacted]
Standard care	12.43	£113,068	-	-	-

QALY – quality-adjusted life year; HSiCR - Hidradenitis Suppurativa Clinical Response

Within this re-analysis, the incremental QALY gain for adalimumab versus standard care is reduced considerably (from 1.00 QALYs in the company's base case to 0.69 QALYs in the HiSCR-based analysis) whilst the incremental cost is increased (from £[redacted] in the company's base case to £[redacted] in the HiSCR-based analysis). Consequently, the ICER is increased to £[redacted] per QALY gained. The ERG notes that ideally the analysis should also have included the re-estimation of all transition matrices to reflect the HiSCR ≥50% AN reduction threshold. More importantly, the ERG notes that whilst the company's re-analysis assumes that there is no difference in utility or resource use between partial responders and non-responders, patients who achieve only a partial response at the end of induction or during maintenance are assumed to continue adalimumab as a maintenance therapy. This is somewhat inconsistent given that within this analysis, these patients are assumed to gain the same health utility as non-responders. Consequently, the value of the company's re-analysis is limited. Had the company's re-analysis extended the continuation rules at induction and maintenance to apply only to those patients achieving a full HiSCR response, this would have likely improved the ICER for adalimumab. This cannot however be confirmed given the company's model structure.

(4) Modelling treatment continuation rules

The company's model assumes that at the end of the induction phase (week 12), patients receiving adalimumab who achieve high response, response or partial response will go on to receive adalimumab as a maintenance therapy. In addition, according to the CS,[9] beyond week 36, patients who are non-responsive to adalimumab are assumed to receive an additional 12 weeks adalimumab therapy prior to discontinuation.

The SmPC for adalimumab states: *"Continued therapy beyond 12 weeks should be carefully reconsidered in a patient with no improvement within this time period. Should treatment be interrupted, Humira 40 mg every week may be re-introduced (see section 5.1). The benefit and risk of continued long-term treatment should be periodically evaluated (see section 5.1)."*[12]

Given that the PIONEER I/II trials[18, 19] used the full HiSCR ≥50% AN reduction threshold as their primary endpoints, this might be reasonably inferred to reflect the definition of improvement within the SmPC. However, the company's model differs in that patients with only a partial response are also assumed to continue to receive adalimumab treatment into the maintenance phase. It is unclear whether the achievement of a partial response would lead to a health gain which is sufficient to warrant the continuation of adalimumab treatment and indeed whether the modelled continuation rule reflects what would typically occur in usual practice.

The ERG also notes that the SmPC does not stipulate how the balance of benefits and risks of continued long-term treatment with adalimumab should be assessed; the wording of the marketing authorisation is not in disagreement with the company's assumption of a further 12 weeks therapy for non-responders, yet it does not specifically recommend such a treatment approach. Whilst the clinical advisors to the ERG were satisfied with the company's assumption that, in practice, clinicians may continue to use adalimumab for a further time period (up to 3 months) if patients have lost a prior response to treatment, they did have concerns that using the HiSCR alone (particularly the achievement of only a partial response) may not represent a sufficiently broad assessment of whether the treatment is working. As discussed in Section 4.2.1, commentators on the validity of the HiSCR measure have highlighted the need to capture other aspects of treatment benefit such as pain and improvements on the DLQI.[33]

(5) Potential overestimation of number of surgical inpatient admissions

Within the company's model, the incidence and costs of surgical inpatient admissions are key drivers of the total costs in both the adalimumab and standard care groups. As detailed in Section 5.2, annual surgical inpatient admissions according to HiSCR response state were based on the company's survey of UK physicians,[22] whilst the costs were based on NHS Reference Costs 2013/14[38] (elective inpatient

code JC41Z - major skin procedures). The ERG has some concerns regarding the estimated lifetime costs associated with inpatient admissions predicted by the company's model. Within the standard care group, the model predicts that the average patient will require approximately 33.87 inpatient surgical admissions over their remaining lifetime. The equivalent number in the adalimumab group is approximately 29.78 admissions. The tariff cost applied to each inpatient admission is £5,488.32 and is assumed to be associated with a length of stay of 5.1 days; this might be considered to be broadly reflective of a wide excision procedure. The costs of these inpatient surgical admissions account for 69.47% of the total discounted lifetime costs in the standard care group and 50.86% of the total discounted lifetime costs in the adalimumab group. As discussed earlier, the CS[9] does not report any evidence to demonstrate that adalimumab reduces the requirement for overall surgical admissions relative to standard care. The ERG notes also that the questionnaire elicited information from respondents on their patients' average use of surgery over the past 12 months and did not consider an upper limit on the number of inpatient surgical admissions per patient.

Clinical advisors to the ERG suggested that excluding the management of surgical complications, the maximum number of sites which may require wide excision for a patient with very extensive disease would be 6-10 (including breasts, groin, the perineum, armpits and buttocks). Patients with less extensive disease would require fewer wide excisions than this maximum number and in some cases more than one region can be treated in the same surgical episode. The ERG's clinical advisors also suggested that patients may undergo a comparatively higher number of smaller procedures such as incision and drainage and narrow margin excision. Incision and drainage may not require inpatient admission and narrow margin excisions are likely to require a shorter length of stay thereby resulting in a comparatively lower cost than that assumed within the company's model. Lowering the cost of surgical inpatient admissions reduces the total costs for both the standard care and adalimumab groups, although given that the company's model suggests that adalimumab will reduce the number of inpatient admissions relative to standard care, this would ultimately lead to a less favourable ICER for adalimumab.

During the clarification process, the ERG queried the number of inpatient surgical admissions predicted using the model (see clarification response,[17] question B7). Within their response, the company stated:

"Considering that a typical HS patient is diagnosed in its early 20s it is not unreasonable to assume that over a lifetime patients who receive no active treatment could undergo approximately 34 inpatient admissions for surgery. Furthermore evidence from the literature suggests that patients with moderate to severe HS undergo surgical procedures quite frequently. Menderes et al 2010[42] reported 54 operative procedures among 27 HS moderate to severe patients from 2004 to 2009. In an

observational cross-sectional study conducted by AbbVie out of 41 patients with surgeries there were 86 surgeries over a 5-year period.[22]*"*

The ERG notes that the starting age assumed in the model is 35 years of age (not early 20s). In addition, in both of the sources cited by the company,[22, 42] the crude surgery rate is around 2 procedures over approximately 5 years (~0.4 procedures per year). This is lower than the estimates predicted by the company's model. Furthermore, the ERG notes that several alternative surgical procedures may be used in the treatment of HS (for example, local destruction, incision and drainage, and narrow margin excision) which require fewer health resources than an inpatient length of stay of 5.1 days and a cost of £5,488 per procedure. Clinical advisors to the ERG noted also that wide excision surgery has a low recurrence rate and does not usually need to be repeated. Overall, the ERG accepts that the true lifetime cost of HS surgery for the population under consideration is highly uncertain, but considers that the assumed cost of each procedure is likely to have been overestimated within the company's model.

(6) Uncertainty surrounding transition probabilities

There is considerable uncertainty surrounding the long-term transition probabilities for adalimumab responders, for patients discontinuing adalimumab and for patients receiving standard care. The company's DSA (see Figure 6) indicates that altering some of these probabilities individually has the potential to considerably worsen the cost-effectiveness of adalimumab. This issue is recognised in the CS[9] (page 122): *"… the main limitation is the paucity of data for the licensed dose beyond 12 or 16 weeks due to re-randomisation at 12 or 16 weeks and protocol-driven discontinuation during period B for patients with LOR or WOAI in the PIONEER studies."* The number of patients with available data for each period are summarised in Table 50.

Table 50: Number of patients included in transition matrices

Time period	Adalimumab responders	Adalimumab discontinuers	Standard care
Weeks 0-12	316	N/a	317
Weeks 12-36	68	100	151
Weeks 36+	Unclear	100 (6 observations per patient)	151 (6 observations per patient)

As shown in Table 50, whilst the number of patients with available HiSCR data during induction is fairly large (adalimumab n=316; standard care n=317), the available dataset during the maintenance phase is notably smaller. In particular, only 68 patients were used to model the time-dependent transition matrices for adalimumab responders during weeks 12-36. Whilst this is not a criticism of the model itself, it does suggest considerable uncertainty in the cost-effectiveness estimates produced

from it. The ERG notes also that there are no data on long-term outcomes for patients who have discontinued adalimumab or for patients receiving standard care alone.

(7) Appropriateness of pooling data from PIONEER I and II trials

Where relevant data are available (see Table 40), the company model uses arm-based aggregate data from the PIONEER I/II trials[18, 19] to inform the transition matrices. Within the CS, the company argues against conducting a conventional NMA of trials of all treatments because of differences between trials in baseline characteristics that are potential treatment effect modifiers. The ERG notes that there are methods available which may enable such comparisons to be made, for example, matching-adjusted treatment comparisons or simulated treatment comparisons. However given that only the adalimumab trials assessed response according to the HiSCR measure, the value of using such comparisons to inform the company's model would be limited (or an entirely different model would be required).

The CS also argues against conducting a pairwise meta-analysis of the placebo-controlled adalimumab trials because, in addition to differences in baseline characteristics between PIONEER I and PIONEER II that are potential treatment effect modifiers, *"There are also differences in study design between PIONEER I and PIONEER II … which means that the results of PIONEER I and PIONEER II are not directly comparable."* (CS,[9] page 122). The differences in study design that the company alludes to are the use of concomitant oral antibiotics at study entry and subsequently, and the inclusion of concomitant oral antibiotics as a stratification factor within PIONEER II[19] but not PIONEER I.[18] Nevertheless, the company did combine the data from PIONEER I and PIONEER II and the ERG requested clarification regarding this approach (see clarification response,[17] questions B14 and B15). In their response[17] (question B14), the company focusses largely on the similarities in study design, the limited sample size within the individual studies and the similarities in baseline characteristics between the two trials, stating that: *"From a clinical perspective, both studies are of very similar study design which allows many direct comparisons as well as pooling of data."* This inconsistency in perspective ignores the fact that we expect heterogeneity in treatment effect between trials because the two trials included patients with different baseline characteristics that are potential or known treatment effect modifiers.

Whether it is appropriate to combine the evidence from PIONEER I and PIONEER II given the issue of antibiotic use raises some important issues. Firstly, it is never appropriate to perform an arm-based synthesis of data from different trials because this breaks the randomisation within trials; an appropriate analysis involves combining trial-specific treatment effects. Secondly, if variation in treatment effect is expected between trials, then this should be acknowledged in the analysis, ideally by conducting a random effects meta-analysis. Thirdly, consideration should be given to the

appropriate estimate of the treatment effect in the target patient population. If PIONEER I and PIONEER II estimate different treatment effects, then neither trial provides an estimate of the treatment effect in the target patient population. In addition, the mean of a random effects distribution does not relate to any specific patient population and the predictive distribution of a new trial is generally recommended as an estimate of treatment effect. Interestingly, the company performed a sensitivity analysis using only the data from PIONEER II and reported that the ICER for adalimumab versus standard care was £[redacted] per QALY gained (see Table 47). The ERG considers that the implications of these issues are that the estimate of treatment effect provided by the company is likely to be biased, understates uncertainty and lacks clarity regarding the population for whom the decision is being made.

The ERG notes that these same issues also apply to the company's estimates of the AE incidence rates during the induction and maintenance phases of the model (see Table 44).

(8) Conceptual inconsistency in handling time-variance in transition probabilities

The ERG considers that the company's approach to handling time dependency in the health state transition probabilities is conceptually inconsistent. Within the company's model, up to week 36, 2- or 4-week time-dependent transition matrices are used to characterise the trajectories of patients receiving adalimumab and standard care. Thereafter, the company's model uses a single time-independent transition matrix for: (a) adalimumab responders; (b) adalimumab discontinuers, and; (c) patients receiving standard care, based on separate GLMs for each of these three groups. For the adalimumab discontinuers and standard care group, the logit models are each based on all transitions previously observed during the maintenance phase (weeks 12-36), which were treated as being time-variant during the earlier cycles in the model. The ERG considers this approach to be somewhat inconsistent as time-dependency is assumed for one portion of the model, but is then ignored for the remainder of the time horizon, even though the time-variant and time-invariant matrices are based on the same data. The company's decision to adopt this approach was likely driven by the limitations of the available evidence. The ERG notes that whilst the company's scenario analyses consider the use of alternative methods for projecting long-term HiSCR outcomes, the impact of incorporating time-variance in the post-36 week transition matrices is unclear.

(9) Potential biases in the use of OLE study data for long-term adalimumab responders

Within the company's model, long-term HiSCR outcomes for adalimumab responders beyond week 36 are modelled using a GLM based on the M12-555 OLE study.[20] The populations recruited into the OLE study included people not achieving a response by or after week 16 in PIONEER I/II[18, 19] and those who achieved response and completed the PIONEER trials.

The ERG has some concerns regarding the use of these data in the model.

(i) The population recruited into the M12-555 OLE study includes a mix of patients who achieved and maintained a response to adalimumab within the PIONEER trials, as well as non-responders. This is not directly in line with the experience of the patient group for whom the matrix is applied in the model as these patients are specifically those who have achieved at least a partial response to adalimumab up to week 36. The impact of including patients with a history of response or non-response to adalimumab, rather than only long-term adalimumab responders, is unclear. The ERG notes that including only the specific group of patients with a prior response to adalimumab would reduce the available sample size for the GLM further, thereby increasing uncertainty.

(ii) Whilst patients in the M12-555 OLE study were previously enrolled within the PIONEER I/II trials,[18, 19] the OLE study adopted an unblinded observational design. Since the use of the OLE data in the model is not based on relative treatment effects drawn from a randomised blinded study design, there is a risk of bias and confounding.

(iii) The data from the OLE study used in the model have been derived from an interim analysis. Given the immaturity of these data, particularly in terms of length of follow-up for the overall OLE cohort, these transition probabilities are subject to further uncertainty.

(iv) The company used an LOCF imputation rule, whereby patients' final observations are carried forward to the final timepoint, to account for missing data in the OLE cohort, noting specifically that *"less than half of the patients had follow-up up to 24 weeks at the time of the interim data cut"* (see CS,[9] page 142). The ERG notes that the single imputation LOCF approach only produces unbiased estimates of treatment effect in certain situations. In particular, the approach may produce optimistic estimates of treatment effect if the patient's condition is expected to worsen following withdrawal from treatment. The ERG has further concerns about the use of LOCF without adequate justification or an assessment of the robustness of results based on sensitivity analyses using alternative approaches.[12] At the request of the ERG, the company provided the results of an alternative GLM which did not include imputation (see Table 51). As shown in the table, the GLM-derived transition probabilities are affected by the LOCF imputation, although this impact does not appear to be substantial. Based on these additional data, the company's deterministic ICER for adalimumab versus standard care was decreased slightly to £[redacted] per QALY gained. The company's clarification response also notes that when the PIONEER I/II trial data are used instead of the OLE study, the ICER for adalimumab versus standard care was reduced to £[redacted] per QALY gained.[17]

(v) The ERG considers the use of the OLE study to model the trajectory of long-term adalimumab responders beyond 36 weeks to be somewhat inconsistent with the approach

used for all other clinical parameters within the model (which are all based on PIONEER I/II[18,19]), but accepts the company's reasons for using these data.

Table 51: Alternative GLM model excluding imputation

OLE study GLM-derived transition matrix including LOCF imputation (base case)				
From/to state	High response	Response	Partial response	Non-response
High response	0.91	0.04	0.04	0.00
Response	0.23	0.70	0.02	0.05
Partial response	0.08	0.09	0.81	0.02
Non-response	0.03	0.03	0.12	0.82
OLE study GLM-derived transition matrix excluding LOCF imputation				
From/to state	High response	Response	Partial response	Non-response
High response	0.88	0.04	0.06	0.01
Response	0.28	0.63	0.02	0.07
Partial response	0.11	0.10	0.75	0.04
Non-response	0.04	0.05	0.14	0.77

OLE – open-label extension; LOCF – last observation carried forward; GLM – generalised logit model

(10) Model errors and other issues surrounding model implementation

The ERG rebuilt the deterministic version of the company's model in order to assess the logic of the company's model structure, to draw out any unwritten assumptions and to identify any errors in the implementation of the model. Table 52 presents a comparison of total QALYs and costs for adalimumab and standard care, as estimated using the company's model and the ERG's rebuilt model. As shown in Table 52, the ERG was able to produce very similar estimates of costs, health gains and cost-effectiveness to those estimated using the company's model.

Table 52: Comparison of company's base case model and ERG's rebuilt model

Option	**Company's model**			**ERG's rebuilt model**		
	QALYs	**Costs**	**ICER**	**QALYs**	**Costs**	**ICER**
Adalimumab	12.61	[redacted]	[redacted]	12.61	[redacted]	[redacted]
Standard care	11.61	£128,541	-	11.61	£128,541	-

QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

During the model double-programming exercise, the ERG identified four errors in the implementation of the company's model, as detailed below.

(i) Inconsistent handling of time

The company's model is inconsistent with respect to the number of days per year. For example, the QALY calculations correctly divide the cycle duration by 365.25 days; however, the weekly discount rate, the per-cycle mortality calculations, the age tracker and the cost calculations all assume that there

are exactly 52 weeks per year (364 days). These minor errors produce a small bias for both the adalimumab and standard care groups.

(ii) The cost of adalimumab is implemented incorrectly

Within the company's model, the health state costs and AE costs are applied from the first cycle (weeks 0-2); however the model only includes the costs of adalimumab from the beginning of the second cycle (during weeks 2-4). This is incorrect, as the initial dose of 160mg should have been included during the first cycle (for the period week 0-2) and an additional cost of maintenance therapy should have been applied to all patients except discontinuing non-responders for the cycle beginning at week 12.

(iii) Incorrect implementation of half-cycle correction

The implementation of the half-cycle correction within the company's model is incorrect. Whilst the company correctly subtract half of the QALY gain and cost for the final cycle from the unadjusted totals, the model includes the full QALY gains and cost for the first model cycle (at cycle 0). Only half of this QALY gain and cost should have been included in the cycle-corrected totals.

(iv) Incorrect implementation of the adalimumab non-responder continuation rule during the maintenance phase

According to the CS,[9] the company's model includes an assumption whereby patients receiving adalimumab who continue to achieve no response from treatment receive an additional 12 weeks of adalimumab treatment prior to discontinuation. The CS[9] (page 138) states that this assumption was based on input from clinical experts who suggested that patients who do not respond to adalimumab treatment will be discontinued in clinical practice after a re-assessment period and 12 additional weeks of treatment. Clinical advisors to the ERG were satisfied that the principle of continuing adalimumab treatment in these patients for a short period is reasonable.

However, the ERG notes that the implementation of this assumption within the company's model is incorrect. In the model, the probability of transiting from the adalimumab no response state to the adalimumab no response state drawn from the OLE GLM (probability = 0.82) is raised to the power of 3 (leading to a probability of 0.56) and is assumed to reflect the probability of discontinuing adalimumab (i.e. transiting to the standard care no response state). All other transitions in the row of the matrix are then adjusted accordingly. The model then applies the first unadjusted matrix to the cycles beginning at weeks 40 and 44, followed by the adjusted matrix from week 48 onwards. The impact of the company's assumption is that the use of this higher discontinuation rate leads to patients discontinuing adalimumab more quickly, thereby substantially reducing the total adalimumab treatment costs.

The ERG sought clarification regarding the mathematical logic underpinning the company's approach (see clarification response,[17] questions B3 and B17). In their response, the company stated: *"…the assumption is made that when patients are in the non-response health state for 12 weeks, they discontinue treatment. 12 weeks equals three model cycles of four weeks. The probability of a patient staying in the non-response health state for three consecutive cycles is the probability of a patient remaining in the non-response health state for 1 cycle cubed. Therefore the transition probability in "cell N130" is the probability of a patient remaining in the non-response health state for 4 weeks cubed."*

The ERG does not consider the company's approach of cubing the discontinuation probability to be mathematically correct. The cubed probability reflects the 12-week probability of consistently remaining in the no response state for three 4-week cycles. As shown in Table 53, the ICER for adalimumab versus standard care is increased substantially as the discontinuation rate is reduced.

Table 53: Modelled time on adalimumab treatment based on OLE discontinuation rate and applying company's 12-week continuation approach

Model scenario	Mean time spent receiving adalimumab (years)	ICER (adalimumab versus standard care)
Model including company's 12-week continuation approach	2.47	£[redacted]
Model based on observed OLE discontinuation rate	5.51	£[redacted]

OLE – open-label extension; ICER – incremental cost-effectiveness ratio

The mathematically correct approach to modelling the company's intended adalimumab continuation rule for non-responding patients would be to include tunnel states to reflect the number of prior cycles in which adalimumab non-responders have remained non-responsive before discontinuing treatment, whilst also accounting for the probability that a patient regains a response within the 12-week period.

Following receipt of the company's clarification response,[17] the ERG asked NICE to request further clarification from the company regarding their implementation of this continuation rule for non-responding patients. In response to this further request, the company sent an additional brief explanation together with a mock-up Excel file[43] which compares their implemented approach against an alternative approach in which 3 consecutive non-response tunnel states are modelled using aggregate HiSCR response/non-response data from the OLE GLM. The company's documented response is reproduced in full below:

"In the base case analysis the same discontinuation rate is assumed during weeks 12-36 for all ADA patients, regardless of health states, since all patients remaining on ADA during this period were

week 12 responders and if a loss of response might occur, an attempt would most likely be made to regain response instead of aggressive discontinuation as suggested by the experts consulted during this submission. However, after week 36 the discontinuation rate is based on the response-specific discontinuation rates since the discontinuation rate of ADA would most likely be driven by loss of response to treatment in the long term, given that patients who remained on ADA treatment for 36 weeks were likely to be those who tolerated the biologic treatment well.

Clinical experts consulted during this submission suggested that patients who do not respond to ADA treatment will most likely be discontinued in clinical practice after a re-assessment period and 12 additional weeks of treatment. Furthermore the ADA drug label also indicates that "the benefit and risk of continued treatment should be periodically evaluated after week 12". As such in the model base case all patients who are in the non-response health state at week 36 discontinue ADA treatment at week 48. In order to implement this assumption into the model patients who were in a non-response health state at week 36 were assigned the non-response discontinuation rate as per the OLE trial in weeks 36-40, 40-44 and 44-48 (first 12 weeks) and then at week 48 were discontinued using the cubing approach.

Beyond week 48 all patients who move to the non-response health state also discontinue treatment at a rate of 0.56 per cycle, taking in the assumption that patients who have been unresponsive for 12 consecutive weeks should discontinue treatment (the probability of adalimumab discontinuation for non-responders is 0.56 at week 48+). The assumption around the use of a higher discontinuation rate beyond week 48 was necessary in order to stop treatment in all patients who would gain no further benefit with ADA treatment, as was suggested by the clinical experts consulted. Using the discontinuation rates as observed in the OLE trial (annual rate of 44.99%) beyond week 48 would result in some patients not responding at week 36 continuing treatment with ADA for far more than 12 weeks.

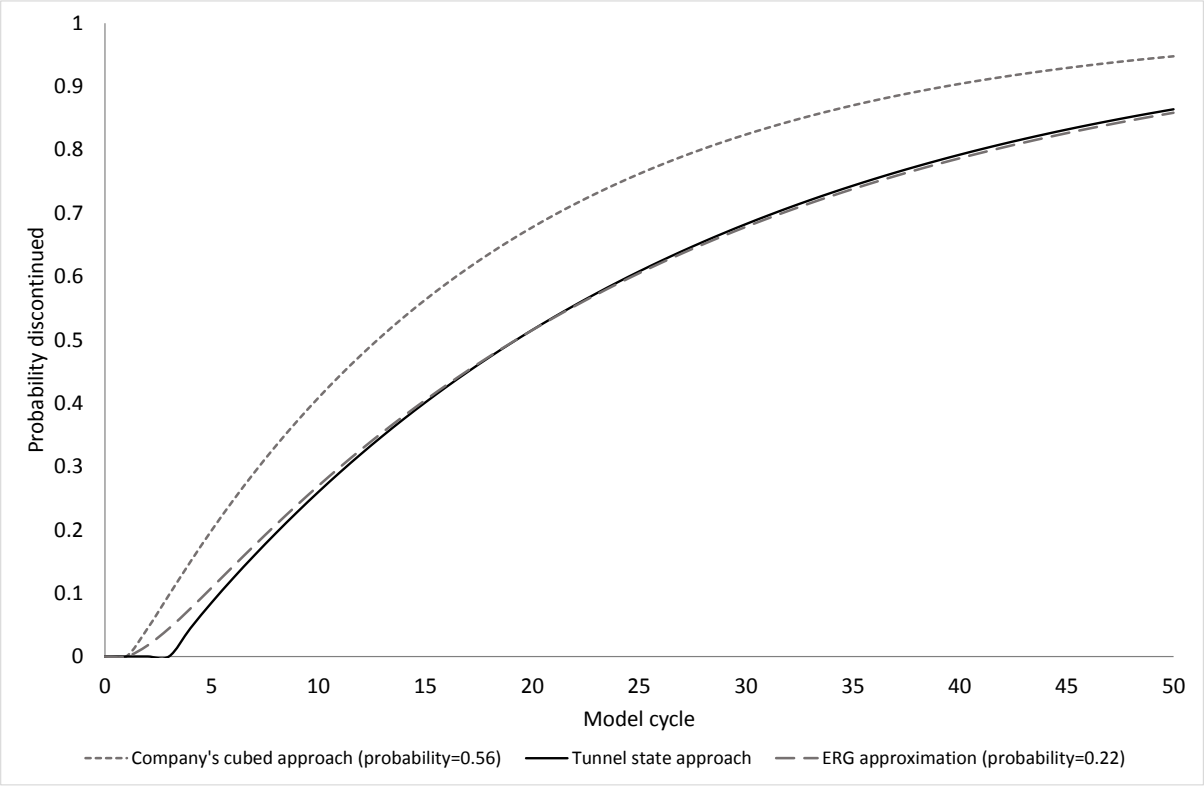
The cubed transition probability is used to reflect the assumption that patients that have been unresponsive for 12 consecutive weeks discontinue treatment. This approach was used in order to avoid introducing multiple tunnel states into the model. The ERG seem to suggest that cubing the probability of remaining unresponsive would overestimate the proportion of patients discontinuing, however the proportion discontinuing will equal out in the long term. AbbVie has provided an example with and without tunnel states to demonstrate the impact of using a model with and without tunnel states. From the calculations provided we can notice that there is initially a difference between the proportion of patients that have discontinued with and without using tunnel states, however this difference becomes smaller in the long term."[43]

With respect to the company's additional response, the ERG notes the following:

- The justification for attempting to incorporate a continuation rule for adalimumab non-responders is reasonable. Clinical advisors to the ERG were satisfied that this is likely to reflect how adalimumab may be used in clinical practice. The ERG's concerns relate to the mathematical implementation of this continuation rule.
- The ERG also agrees that using the observed discontinuation probability for non-responders beyond week 36 may result in an unrealistic proportion of patients remaining on adalimumab yet deriving no further benefit from it.
- The company's response appears to accept that using tunnel states is appropriate, but attempts to justify not using this approach due to the increased complexity associated with its implementation. This is not a satisfactory justification. The use of such an approach may lead erroneous model results.
- The company's approach of cubing the 4-week probability of transiting from non-response to non-response produces a 12-week probability of remaining non-responsive; this cannot be directly used in a model which uses a 4-weekly cycle length. Whilst the ERG understands how the probability of 0.56 has been derived, its use in the model reflects an error of logic.
- The company's Excel mock-up differs slightly from the company's model with respect to how the other transition probabilities from the no response state are normalised. In the Excel mock-up model, the probability of remaining on treatment is calculated as the probability of being non-responsive minus the probability of discontinuing adalimumab. In the company's submitted economic model, all transitions from the non-response state to the high response, response and partial response states are normalised by multiplying the transition probability by one minus the probability of discontinuation.
- The company's Excel mock-up demonstrates that using the cubed probability of 0.56, the probability of adalimumab discontinuation is consistently and substantially overestimated relative to the tunnel state approach. Whilst the company's response indicates that this difference diminishes over time, this is only because there are few patients left on treatment by that point. The ERG notes that within the company's Excel mock-up, manually reducing the 4-week probability of adalimumab discontinuation to a value of 0.22 (estimated by trial and error) produces a much closer approximation of the correct tunnel state approach (see Figure 9). It is also noteworthy that converting the company's cubed (12-week) probability of 0.56 to a rate and then converting this back to a 4-week probability gives an estimated discontinuation probability of 0.24, which is similar to the ERG's manually derived estimate (the slight difference is likely to be due to the small probability of non-responders regaining response during each cycle). The ERG considers the value of 0.22 to be a more reasonable,

but not ideal, approximation of the company's assumed 12-week adalimumab non-responder continuation rule during the maintenance phase of the model.

Figure 9 Time to treatment discontinuation using the company's Excel mock-up



Overall, the ERG considers that the company should have adopted a model structure which includes tunnel states to account for the assumed maintenance phase continuation rule. Based on the evidence presented within the CS and subsequent clarification responses, the impact of the company's approach on the expected ICER for adalimumab versus standard care is unclear.

5.4 Additional exploratory analyses undertaken by the ERG

This section presents additional exploratory analyses using the company's model undertaken by the ERG.

5.4.1 Methods for exploratory analyses

Based on the issues discussed in the ERG's critical appraisal of the company's model (see Section 5.3), eight sets of additional analyses were undertaken. The first three sets of analyses reflect the ERG's base case, whilst the subsequent five sets of analyses were undertaken to examine remaining uncertainties using the ERG's base case. Specific amendments made to the company's model within these analyses are detailed in Appendix 2.

ERG Exploratory Analysis 1 - Correction of model errors

As detailed in Section 5.3, the ERG identified several minor errors in the implementation of the company's model. Within this exploratory analysis, the ERG corrected the inconsistencies in the number of days in a year, resolved the issues surrounding the implementation of the half-cycle correction and altered the timing of the adalimumab acquisition costs to reflect the licensed dosing schedule.[12] The issues surrounding the 12-week adalimumab non-responder continuation rule during the maintenance phase of the model are not addressed within this analysis.

ERG Exploratory Analysis 2 – Incorporation of tunnel states to reflect the maintenance phase adalimumab non-responder continuation rule

The company's base case model attempts to apply a continuation rule during the maintenance phase whereby patients who are non-responsive to adalimumab continue to receive an additional 12-weeks of adalimumab therapy prior to withdrawing from treatment. The mathematical implementation of this assumption within the company's model is flawed and leads to a rapid withdrawal rate for patients receiving adalimumab (see Figure 9). In this analysis, major structural changes were made to the company's model to implement the company's adalimumab non-responder continuation rule from week 40 onwards, as described in the CS and subsequent clarification responses. This involved the following steps:

1. Restructuring the 36+ week adalimumab transition matrix to include three tunnel states for adalimumab non-responders.
2. Re-calculating the transition probabilities based the original 36+ week matrix whereby the probability of transiting from each tunnel state to the next tunnel state (or eventually discontinuing) is defined as the complement of each row of probabilities. The original and amended week 36+ adalimumab transition matrices are shown in Tables 54 and 55, respectively.

3. Re-generating the Markov trace for the high response, response and partial response maintenance phase submodels using a looping approach to account for state transitions followed by adjustments to account for other-cause mortality.
4. Condensing the new Markov trace for each submodel back to the original states defined in the company's model, whereby health state occupancy in the no response state is calculated as the sum of the health state occupancy in all three no response tunnel states in each cycle.
5. Replacing the entire Markov trace for each submodel with the new trace generated by the ERG from week 40 onwards. The ERG notes that whilst this is in line with the assumptions of the company's model, in practice, this continuation rule may apply immediately following the start of adalimumab maintenance therapy (from week 16 onward).

Table 54: Company's original week 36+ matrix for patients receiving adalimumab

			To state							
			ADA				SC			
			High response	Response	Partial response	No response	High response	Response	Partial response	No response
From state	ADA	High response	[redacted]							
		Response								
		Partial response								
		No response								
	SC	High response								
		Response								
		Partial response								
		No response								

Table 55: Week 36+ matrix for patients receiving adalimumab including tunnel states

			To state									
			On adalimumab						On standard care			
			High response	**Response**	**Partial response**	**No response1**	**No response 2**	**No response3**	**High response**	**Response**	**Partial response**	**No response**
From state	ADA	High response										
		Response										
		Partial response										
		No response 1										
		No response2										
		No response3										
	SC	High response										
		Response										
		Partial response										
		No response										

** Calculated as complement of all other transitions in the row*
† In line with the transitions from the high response, response and partial response states, spontaneous discontinuation from each tunnel state is also assumed

This analysis also includes the minor model amendments detailed in ERG Exploratory Analysis 1.

ERG Exploratory Analysis 3 - Use of alternative assumptions regarding the costs of HS surgery inpatient admissions (ERG-preferred base case)

As discussed in Section 5.3, the ERG has concerns that the costs of HS surgical inpatient admissions are likely to be considerably overestimated. An exploratory analysis was therefore undertaken based on revised HS surgery costings developed by the ERG in conjunction with the clinical advisors involved in the assessment. Within this analysis, the following assumptions were made:

(i) The company's modelled estimate of total lifetime HS surgeries (33.87 procedures) for patients receiving standard care, based on the company's resource use survey, is reasonable;

(ii) Based on the company's retrospective cohort study using HES data described in the CS[9] (page 30), [redacted] of all HS surgeries are intermediate procedures which are undertaken in a day case setting;

(iii) Of the remaining [redacted] of HS surgeries, patients on average have 2 wide excisions over their lifetime;

(iv) All other remaining surgeries are comprised of an equal mix of elective and non-elective intermediate skin procedures with an average length of stay (LOS) of 2 days.

Table 56 presents revised estimates of the average cost of HS surgery, valued using 2013/14 NHS Reference Costs.[38] These alternative assumptions result in an estimated cost of £1,525.74 per surgical procedure. Within the economic analysis, this cost is applied as the unit cost for all HS surgical admissions.

Table 56: Revised HS surgery costing assumptions

Parameter	Value	Source
Lifetime number of surgeries for patients receiving standard care	33.87	Company's model prediction
Average number of wide excisions over patient's lifetime	2	Expert opinion (JRI)
Proportion of all surgeries which are undertaken in day case setting	[redacted]	Company's survey (page 30)
Proportion of all surgeries which are wide excisions	[redacted]	Assumption
Proportion of all surgeries which are intermediate procedures requiring inpatient admission (including procedure plus 24 hours i.v. antibiotics)	[redacted]	Assumption
Cost day case intermediate procedure	£943.17	NHS Reference Costs 2013/14 - JC42A (day case)
Cost wide excision	£5,488.32	NHS Reference Costs 2013/14 - JC41Z (inpatient elective)
Cost intermediate skin procedure requiring admission	£2,102.73	NHS Reference Costs 2013/14 - (average of JC42A elective and JC42A non-elective assuming length of stay=2 days)
Mean HS surgery cost	**£1,525.74**	-

HS – hidradenitis suppurativa

This analysis also includes the model corrections and incorporation of tunnel states for non-responders detailed in ERG Analyses 1 and 2.

ERG Additional Exploratory Analysis 4 – Use of PIONEER II data only

As discussed in Section 5.3, the CS presents contradictory arguments regarding whether PIONEER I and II should be pooled. Within this analysis, the ERG presents a scenario which includes data only from the PIONEER II trial.

This additional exploratory analysis also includes the amendments detailed in ERG Exploratory Analyses 1-3.

ERG Additional Exploratory Analysis 5 – Alternative assumptions regarding transition probabilities beyond week 36

Within this analysis, two alternative scenarios were considered to explore the uncertainty surrounding the long-term extrapolation of health state transitions within the company's model. The first analysis uses the GLM transition matrix derived from the M12-555 OLE study but excludes the use of LOCF

imputation (see Table 51). The second scenario uses the alternative transition matrices for adalimumab discontinuers and patients receiving standard care based on the mean transition probabilities for weeks 12-36 derived from the PIONEER I/II trials.

This additional exploratory analysis also includes the amendments detailed in ERG Exploratory Analyses 1-3.

ERG Additional Exploratory Analysis 6 – Discontinuation of partial responders and non-responders at 12-weeks

Within this analysis, patients who achieve a response or a high response on adalimumab at 12 weeks are assumed to continue adalimumab treatment, whilst those achieving only a partial response or no response are assumed to discontinue at this timepoint. It should be noted that due to the structural limitations of the model, it was not possible to apply the company's intended maintenance continuation rule to both partial responders and non-responders as this would require a further set of tunnel states for partial responders.

This additional exploratory analysis also includes the amendments detailed in ERG Exploratory Analyses 1-3.

ERG Additional Exploratory Analysis 7 – Assumption of no difference in utility, resource use or discontinuation rates for non-responders and partial responders, and for high responders and responders

Within this analysis, the utility values, resource use estimates and discontinuation rates for the high response and response states, and for the partial response and no response states, are assumed to be the same based on the alternative model submitted by the company during the clarification stage (see Table 49). The ERG notes however that within this analysis, partial responders are assumed to continue adalimumab treatment, yet they derive no more benefit than non-responders.

This additional exploratory analysis also includes the amendments detailed in ERG Exploratory Analyses 1-3.

ERG Additional Exploratory Analysis 8 – Assumption of no difference in utility, resource use or discontinuation rates for non-responders and partial responders, and for high responders and responders, including the discontinuation of partial responders and non-responders at 12-weeks

This analysis is the same as the previous analysis, except that patients who achieve a partial response at 12 weeks are assumed to discontinue adalimumab induction therapy. This provides some indication of the impact of discontinuing adalimumab in both partial responders and non-responders, but only at

the end of induction. It would have been preferable to apply a consistent continuation rule to partial responders during the maintenance phase, however, this was not possible within the company's model structure.

This additional exploratory analysis also includes the amendments detailed in ERG Exploratory Analyses 1-3.

5.4.2 Results of the ERG's additional exploratory analyses

ERG Exploratory Analysis 1: Correction of model errors

Table 57 presents the results of ERG Exploratory Analysis 1 which includes only the correction of model errors discussed in Section 5.3 (see critical appraisal point 10 and Appendix 2).

Table 57: ERG Exploratory Analysis 1 – correction of model errors

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	12.64	£[redacted]	1.00	£[redacted]	£[redacted]
Standard care	11.64	£128,430	-	-	-

HiSCR – Hidradenitis Suppurativa Clinical Response; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

Based on the corrected version of the company's model, the deterministic ICER for adalimumab is estimated to be £[redacted] per QALY gained; this is marginally higher than the company's base case estimate presented within the CS[9] (original ICER=£[redacted] per QALY gained).

ERG Exploratory Analysis 2: Incorporation of tunnel states to reflect the maintenance phase adalimumab non-responder continuation rule (including ERG Exploratory Analysis 1)

Table 58 presents the results of the company's model which includes the addition of tunnel states to better reflect the proposed adalimumab non-responder continuation rule during the maintenance phase. The analysis also includes the model corrections presented in ERG Exploratory Analysis 1.

Table 58: ERG Exploratory Analysis 2 – incorporation of tunnel states to reflect the maintenance phase adalimumab non-responder continuation rule

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	12.72	£[redacted]	1.07	£[redacted]	£[redacted]
Standard care	11.64	£128,430	-	-	-

HiSCR – Hidradenitis Suppurativa Clinical Response; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

The results presented in Table 58 demonstrate that the incorporation of tunnel states within the company's model increases both the incremental QALY gains and the incremental costs of

adalimumab relative to the company's base case estimates. The incorporation of tunnel states for adalimumab non-responders in the corrected version of the model increases the ICER for adalimumab versus standard care to £[redacted] per QALY gained.

The ERG notes that using the corrected version of the company's submitted model together with an adalimumab non-responder 4-week discontinuation probability of 0.22 (see Figure 9) produces a similar ICER to the results presented in Table 58 (ICER=£[redacted] per QALY gained).

ERG Exploratory Analysis 3: Revised assumptions regarding costs of HS surgery (including ERG Exploratory Analyses 1 and 2)

Table 59 presents an exploratory analysis in which the cost of surgical inpatient admissions is assumed to be £1,525.74 per procedure (see Table 56). This analysis also incorporates the model corrections applied in ERG Exploratory Analysis 1 and the tunnel states applied in ERG Exploratory Analysis 2. This analysis represents the ERG's preferred base case (given the constraints of the company's adopted model structure).

Table 59: ERG Exploratory Analysis 3 – revised assumptions regarding costs of HS surgery (ERG base case)

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Probabilistic model					
Adalimumab	12.75	£[redacted]	1.09	£[redacted]	**£**[redacted]
Standard care	11.66	£63,909	-	-	-
Deterministic model					
Adalimumab	12.72	£[redacted]	1.07	£[redacted]	**£**[redacted]
Standard care	11.64	£64,018	-	-	-

HiSCR – Hidradenitis Suppurativa Clinical Response; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

As shown in Table 59, the estimated QALY gains for adalimumab and standard care are the same as those estimated within ERG Analysis 2. However, the total discounted lifetime costs in both treatment groups are reduced considerably. Since the ERG's preferred estimate of the costs of HS surgery are lower than those used in the company's model, and because the company's base case analysis suggests that adalimumab produces cost savings by avoiding HS surgery due to patients spending more time in the better response states, this analysis produces a higher incremental cost for adalimumab versus standard care. Within this analysis, the deterministic ICER for adalimumab versus standard care is estimated to be £[redacted] per QALY gained. Based on the probabilistic version of the model, the ICER for adalimumab versus standard care is expected to be £[redacted] per QALY gained.

ERG Additional Exploratory Analysis 4: Use of PIONEER II data only (using the ERG-preferred base case)

Table 60 presents an exploratory analysis using only the PIONEER II data. This analysis uses the ERG's base case version of the model (ERG Exploratory Analysis 3).

Table 60: ERG Additional Exploratory Analysis 4 – use of PIONEER II data only

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	12.63	£[redacted]	0.99	£[redacted]	**£[redacted]**
Standard care	11.64	£64,007	-	-	-

HiSCR – Hidradenitis Suppurativa Clinical Response; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

The results presented in Table 60 suggest that deriving the transition matrices and AE probabilities only from the PIONEER II trial increases the ICER for adalimumab versus standard care to £[redacted] per QALY gained. This is partly a consequence of patients remaining on adalimumab for a longer period of time compared with the ERG's base case analysis.

ERG Additional Exploratory Analysis 5 – Alternative assumptions regarding transition probabilities beyond week 36

Table 61 presents the results of two exploratory analyses using alternative long-term transition probabilities.

Table 61: ERG Additional Exploratory Analysis 5 – alternative assumptions regarding transition probabilities beyond week 36

OLE GLM for adalimumab responders (excluding imputation), PIONEER I/II GLMs for adalimumab discontinuers and patients receiving standard care					
Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	12.68	£[redacted]	1.04	£[redacted]	**£[redacted]**
Standard care	11.64	£64,018	-	-	-
OLE GLM for adalimumab responders (including LOCF), mean of week 12-36 data from PIONEER I/II for adalimumab discontinuers and patients receiving standard care					
Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	12.58	£[redacted]	1.17	£[redacted]	**£[redacted]**
Standard care	11.41	£65,650	-	-	-

As shown in Table 61, the results of these analyses suggest that the cost-effectiveness of adalimumab versus standard care is slightly reduced when alternative long-term transition matrices are used to project HiSCR outcomes. When LOCF imputation is removed from the GLM for patients receiving adalimumab beyond week 36, the ICER for adalimumab versus standard care is estimated to be

£[redacted] per QALY gained. When the transition matrices for patients who have discontinued adalimumab and for patients receiving standard care are based on the mean of week 12-36 data from the PIONEER I/II trials, the ICER is reduced to £[redacted] per QALY gained.

ERG Additional Exploratory Analysis 6: Discontinuation of partial responders and non-responders at 12-weeks (using the ERG-preferred base case)

Table 62 presents the results of an analysis in which only patients achieving response or high response are assumed to continue adalimumab treatment beyond 12 weeks.

Table 62: ERG Additional Exploratory Analysis 6 – discontinuation of partial responders and non-responders at 12 weeks

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	12.62	£[redacted]	0.98	£[redacted]	**£[redacted]**
Standard care	11.64	£64,018	-	-	-

HiSCR – Hidradenitis Suppurativa Clinical Response; QALY – quality-adjusted life year; ICER – incremental cost-effectiveness ratio

Relative to the ERG's preferred base case, the discontinuation of patients who have achieved only a partial response at 12-weeks results in an estimated ICER for adalimumab versus standard care of £[redacted] per QALY gained. This is more favourable than the ERG's base case analysis. The ERG notes however that the impact of discontinuing treatment for partial responders during the maintenance phase is unclear.

ERG Additional Exploratory Analysis 7: Assumption of no difference in utility, resource use and discontinuation rates for non-responders and partial responders, and for high responders and responders (using the ERG-preferred base case)

Table 63 presents the results of an analysis in which the model corrections, non-responder tunnel states and lower surgery cost (ERG Exploratory Analyses 1, 2 and 3) are applied to a version of the model in which health utilities, resource use and discontinuation rates are assumed to be the same for partial responders and non-responders, and high responders and responders.

Table 63: ERG Additional Exploratory Analysis 7 – assumption of no difference in utility, resource use and discontinuation rates for non-responders and partial responders, and for high responders and responders

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	13.25	£[redacted]	0.79	£[redacted]	**£[redacted]**
Standard care	12.46	£57,065	-	-	-

The results of this analysis suggest a considerably higher ICER than both the ERG's base case and the company's base case. However, it is important to note that whilst partial responders are assumed to continue adalimumab as maintenance therapy, their health utility is assumed to be the same as that for non-responders, hence this analysis assumes that these patients remain on treatment without obtaining further benefit from it. The ERG would have preferred that the company had incorporated adalimumab continuation rules based on the 50% HiSCR AN reduction threshold.

ERG Additional Exploratory Analysis 8: Assumption of no difference in utility, resource use and discontinuation rates for non-responders and partial responders, and for high responders and responders with discontinuation of patients achieving only partial response or no response at 12-weeks (using the ERG-preferred base case)

Table 61 presents the results of the scenario described in ERG Additional Exploratory Analysis 7, combined with an additional assumption that both non-responders and partial responders discontinue adalimumab at 12 weeks.

Table 64: ERG Additional Exploratory Analysis 8 – assumption of no difference in utility, resource use and discontinuation rates for non-responders and partial responders, and for high responders and responders with discontinuation of patients achieving only partial response or no response at 12 weeks

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Adalimumab	13.17	£[redacted]	0.71	£[redacted]	£[redacted]
Standard care	12.46	£57,065	-	-	-

The results presented in Table 63 indicate that assuming no difference in utility, resource use and discontinuation rates for no response and partial response, and for high response and response, together with the discontinuation of partial responders and non-responders at 12-weeks, the ICER for adalimumab versus standard care is estimated to be £[redacted] per QALY gained. This is lower than the previous scenario in which only non-responders discontinue at 12-weeks (ERG Additional Exploratory Analysis 7, Table 63). As noted above, due to its structure, it was not possible to apply the company's assumed discontinuation rule to partial responders within the maintenance phase of the model. The ERG does however note that increasing the discontinuation rate for partial responders lowers the ICER for adalimumab. However, the true impact of applying the discontinuation rules to both adalimumab non-responders and adalimumab partial responders in both the induction and maintenance phases of the model is unclear. This represents an important uncertainty which cannot be fully addressed given the evidence provided within the CS.

5.5 Discussion

The CS includes a systematic review of economic evaluations of treatments for HS together with a *de novo* model-based economic evaluation of adalimumab versus standard care in adult patients with an inadequate response to conventional systemic HS therapy.

The company's systematic review of existing economic evaluations did not identify any relevant studies for inclusion.

The company's *de novo* economic model adopts a Markov approach to estimate costs and health outcomes for adalimumab and standard care from the perspective of the NHS and PSS over a lifetime horizon. All analyses presented in the CS relate to the full population specified in the marketing authorisation for adalimumab; no subgroup analyses are presented within the CS. The company's model includes five mutually exclusive health states, based on depth of HiSCR response: (i) high response; (ii) response; (iii) partial response; (iv) no response, and; (v) dead. The model uses a 2-week cycle length for the first 2 cycles, and a 4-week cycle length thereafter. Health state transitions are modelled up to week 36 using data from PIONEER I/II, including a discontinuation rule for patients who do not achieve at least a partial response by week 12. The long-term HiSCR trajectory of adalimumab responders (including partial responders) beyond 36 weeks is subsequently modelled using a time-invariant GLM fitted to LOCF-imputed data from the M12-555 OLE study. The long-term HiSCR trajectories for patients receiving standard care and for those who have previously discontinued adalimumab beyond 36 weeks are modelled using separate time-invariant GLMs fitted to data from weeks 12-36 from the PIONEER I/II trials. Health utilities are modelled according to depth of HiSCR response using a *post hoc* analysis of EQ-5D data collected within PIONEER II. Resource use estimates, which are also differentiated by depth of HiSCR response, were based on a survey of UK physicians and were assumed to include inpatient visits due to HS surgery, outpatient visits due to HS surgery, visits to wound care due to HS surgery, non-surgical inpatient visits, non-surgical outpatient visits, visits to wound-care not due to HS surgery, A&E visits and costs associated with AEs. Unit costs were taken from the BNF, the PSSRU and NHS Reference Costs. AEs are not assumed to have an additional impact on HRQoL.

Based on the probabilistic version of the company's base case model, adalimumab is expected to produce an additional 1.02 QALYs at an additional cost of £[redacted] as compared with standard care; the probabilistic [redacted] for adalimumab versus standard care is expected to be £[redacted] per QALY gained. The results of the deterministic model are similar, with adalimumab yielding an ICER of £[redacted] per QALY gained compared with standard care. The company's PSA suggests that assuming a WTP threshold of £20,000 per QALY gained, the probability that adalimumab produces

more net benefit than standard care is approximately [redacted]. Assuming a WTP threshold of £30,000 per QALY gained, the probability that adalimumab produces more net benefit than standard care is approximately [redacted]. Within the company's deterministic scenario analysis, the ICER for adalimumab was greater than £30,000 per QALY gained in four scenarios: (i) when the time horizon was truncated to 20 years; (ii) when the model was based only on data from PIONEER II; (iii) when the LSCF imputation rule was used, and; (iv) when the discontinuation rate for adalimumab non-responders after week 36 was based on the OLE study.

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's model. The ERG's critical appraisal identified a number of issues relating to the company's model and analysis. The most pertinent of these relate to: (i) the use of a model structure in which health gains and treatment continuation rules are defined according to depth of response, which does not reflect the pre-planned and validated HiSCR endpoint used in the PIONEER trials; (ii) the likely overestimation of the lifetime costs of HS surgery predicted by the company's model; (iii) the incorrect implementation of a continuation rule for adalimumab non-responders which does not mathematically reflect the actual assumptions stated in the CS; (iv) the use of arm-based aggregate data from the PIONEER I/II trials rather than a formal meta-analysis, and; (v) uncertainty surrounding the long-term transition probabilities derived from the PIONEER I/II trials and the M12-555 OLE study.

The ERG undertook eight sets of exploratory analyses based on the company's submitted model. The first three of these analyses relate to: (i) correction of technical programming errors in the company's model; (ii) applying structural amendments to the model to correctly reflect the company's intended adalimumab non-responder continuation rule during the maintenance phase, and; (iii) re-estimation of the costs of HS surgery. The combination of these three exploratory analyses represent the ERG's preferred base case. Five additional sets of analyses were undertaken using this base case to explore uncertainty surrounding the transition probabilities employed in the model, the likely impact of discontinuing non-responders and partial responders to adalimumab (during the induction phase only) and the potential structural uncertainty around the company's adopted modelling approach. The latter two analyses could not however be fully implemented due to the limitations of the company's model structure.

The ERG's exploratory analyses indicate that the technical programming errors have only a minor impact on the model results and lead to a small increase in the ICER for adalimumab versus standard care. The incorporation of tunnel states for adalimumab non-responders within the maintenance phase of the corrected model increases the ICER for adalimumab versus standard care more substantially (ICER=£[redacted] per QALY gained). The ERG's base case, which comprises a scenario whereby

these two sets of corrections are combined with a lower cost of HS surgery, results in an estimated deterministic ICER for adalimumab versus standard care of £[redacted] per QALY gained. The probabilistic ICER for this analysis is slightly higher (£[redacted] per QALY gained). The ERG's base case ICER for adalimumab versus standard care is markedly less favourable than that presented within the CS.

The additional exploratory analyses undertaken using the ERG's base case model suggest the following:

- Using only PIONEER II to inform the model increases the ICER for adalimumab to £[redacted] per QALY gained. The ERG notes however that this analysis excludes the PIONEER I data; this is not ideal. The ERG would have preferred an analysis whereby treatment effects are based on a formal meta-analysis which maintains the randomised design of the PIONEER trials.
- The exclusion of LOCF imputation within the M12-555 OLE GLM for patients receiving adalimumab and using the mean transition data from the maintenance phase of PIONEER I/II for adalimumab discontinuers and patients receiving standard care, may reduce the ICER for adalimumab versus standard care.
- The discontinuation of partial responders as well as non-responders at the end of induction improves the ICER for adalimumab versus standard care. This can be partly explained in that high responders and responders are assumed to accrue greater benefits than partial responders, yet all three groups incur the same cost of adalimumab whilst receiving treatment. Importantly, owing to the structure of the company's model, this analysis does not apply the company's maintenance phase discontinuation rule to partial responders. Increasing the rate of discontinuation for this group may improve the ICER for adalimumab, however the ERG is unable to fully demonstrate this due to the limitations of the company's model structure.
- Based on the approach used in the company's clarification response, assuming that health utility, resource use and discontinuation rates are the same for partial responders and non-responders, and for high responders and responders, increases the ICER for adalimumab versus standard care to £[redacted] per QALY gained. It is important to note however that this analysis only applies the discontinuation rule to non-responders; whilst partial responders are assumed to continue on adalimumab beyond induction and thereafter, these patients are assumed to accrue the same utility as non-responders. Withdrawing partial responders and non-responders at the end of induction improves the ICER for adalimumab, however, the ERG was unable to apply a consistent continuation rule during the maintenance phase of the company's model. Consequently, it is not possible to fully assess the impact of this uncertainty within the company's model.

There remain several potentially important areas of uncertainty:

1. The impact of using relative treatment effects for adalimumab versus placebo based on a formal meta-analysis of data from PIONEER I and II within the model is unclear. Further, there is no comparative evidence regarding the long-term benefits of adalimumab relative to any other therapy.
2. The company's implemented model is subject to structural uncertainty, in particular around the definition of health states and the use of evidence to populate these. An alternative simpler model would have involved defining health utility, resource use, discontinuation rates, baseline transitions, relative treatment effects and adalimumab continuation rules according to the HiSCR ≥50% AN reduction threshold validated by Kimball *et al* and used as the primary endpoint in the PIONEER I/II trials.
3. The impact of adalimumab on the subsequent requirement or opportunity for surgical intervention is unclear. There is uncertainty around whether reductions in the overall costs of surgery predicted by the company's model will manifest in clinical practice. The impact of taking into account the use of other pharmacological therapies on the cost-effectiveness of adalimumab is also unknown.

6. END OF LIFE

End of life criteria are not relevant to this submission.

7. OVERALL CONCLUSIONS

The CS consisted of three separate reviews: (1) a review of the clinical efficacy evidence from RCTs of treatments for HS, specifically RCTs comparing adalimumab with placebo; (2) a review of the evidence from a non-controlled, OLE study, and; (3) a review of safety evidence from the RCTs of adalimumab versus placebo and the non-controlled, open-label extension study.

The principal clinical efficacy review is a poorly-reported systematic review of three relevant RCTs comparing adalimumab with placebo in adults with moderate to severe HS: these were comprised of a "dosing" Phase II trial, M10-467, and two Phase III trials, PIONEER I and II. The three trials all have two periods: an initial period (Weeks 0-12 in the PIONEER I/II trials and Weeks 0-16 in the M10-467 trial) comparing adalimumab 40mg EW with placebo, and a second period (Weeks 12-36 in the PIONEER trials and Weeks 16-52 in the M10-467 trial), initiated by re-randomisation of patients to arms of adalimumab 40mg EW, placebo or adalimumab 40mg EOW (PIONEER trials only). The included trials are generally consistent with the final NICE scope. The primary efficacy outcome was clinical response to treatment, principally using the company's own HiSCR measure. Clinical advice received by the ERG confirms that the HiSCR measure has been validated but, in terms of clinical decision-making, its findings must be viewed alongside the results of patient-reported outcome measures, in particular quality of life assessed by the DLQI and a pain measure. The trials' secondary outcomes included assessments of disease severity and symptoms, using the MSS score and AN counts, pain and quality of life (various measures).

The ERG considers the M10-467 trial to be at low risk of bias across all domains for the relevant Period 1 (up to week 16). The ERG also considers the results from Period A (i.e. up to week 12) in PIONEER I and II to be generally at low risk of bias. However, the ERG considers there to be a moderate or unclear risk of selection and attrition bias affecting the results of Period B in the PIONEER trials. There is also a low-to-moderate risk of reporting bias in Period B in the two trials. It should also be noted that whilst M10-467 has been published, the PIONEER trials have not.

In PIONEER I and II, significantly more patients in the adalimumab 40mg EW group achieved a clinical response (defined as achieving HiSCR [at least a 50% reduction in the total AN count with no increase in abscess count and no increase in draining fistula count relative to baseline] at week 12) than patients receiving placebo: 41.8% for adalimumab vs 26.0% for placebo, p=0.003 in PIONEER I, and 58.9% for adalimumab vs 27.6% for placebo, p<0.001 in PIONEER II. Subgroup analyses indicated that patients achieved benefit with adalimumab 40mg EW regardless of their baseline characteristics, although for some subgroups had small patient numbers. Significant or clinically relevant differences in favour of adalimumab 40mg EW that were reported for secondary outcomes in

PIONEER II were not always found in PIONEER I, especially for AN count, MSS score, pain and some components of quality of life measured by the SF-36. The treatment effect varied between the trials. This might be explained in part by different patient demographics across trials. The company is conducting ongoing analyses of the data from the PIONEER trials and the OLE study to understand these differences. An NMA was not considered feasible.

An arm-based integrated summary, which breaks randomisation, was conducted for the two PIONEER trials to tabulate Period B response (12-36 weeks) for all patients and for a group of HiSCR "responders" and "partial responders." According to this analysis, improvements in response were maintained or reduced in this second period. However, the "partial responder" group (defined as HiSCR responders with ≥25% reduction rather than ≥50% reduction) are a *post hoc* analysis group. This group was not defined in protocols or published descriptions of study design or pre-specified analysis methods for the PIONEER trials. It was also not considered in the published validation study for the HiSCR measure, nor was it justified or explained in the company's clinical review. A small number of secondary outcomes were reported for PIONEER I and II only for patients who had had clinical response at week 12, but the results were based on analyses with small sample sizes (range of 15 to 22 patients across all outcomes for both PIONEER trials).

These trials were supplemented by a single, unpublished, non-randomised, non-controlled, un-blinded cohort study, which was an OLE study of the PIONEER trials (M12-555 OLE). In terms of efficacy, the results suggested that [redacted]

[redacted] Details of the results for secondary outcomes such as MSS and NRS30 were not reported. The ERG considers these efficacy results to be uncertain because they are the result of interim analyses of unpublished study data with a sizeable amount of missing data. The study also only potentially offers efficacy data for up to 72 weeks for a drug that might be taken for many years by patients with moderate or severe HS.

The submission of safety evidence was a review of the three generally good quality RCTs, supplemented by the single arm cohort study. There were no obvious safety concerns, with most AEs being balanced across adalimumab 40mg EW and placebo trial arms, and small numbers of SAEs. Longer-term data are required to determine whether reported AE rates are maintained for patients on long-term maintenance doses of adalimumab 40mg EW, whether or not certain subgroups of patients are at higher risk of certain events, and to confirm whether or not there are any differences between the interrupted and uninterrupted regimens. The submission notes the M12-555 OLE is the only

ongoing study of adalimumab in this indication. Final data from this study are expected to be available in 2016.

Based on the probabilistic version of the company's base case model, adalimumab is expected to produce an additional 1.02 QALYs at an additional cost of £[redacted] as compared with standard care; the probabilistic ICER for adalimumab versus standard care is expected to be £[redacted] per QALY gained. The results of the deterministic model are similar, with adalimumab yielding an ICER of £[redacted] per QALY gained compared with standard care. The company's PSA suggests that assuming a WTP threshold of £20,000 per QALY gained, the probability that adalimumab produces more net benefit than standard care is approximately [redacted]. Assuming a WTP threshold of £30,000 per QALY gained, the probability that adalimumab produces more net benefit than standard care is approximately [redacted]. Within the company's DSA, the ICER for adalimumab was greater than £30,000 per QALY gained in four scenarios: (i) when the time horizon was truncated to 20 years; (ii) when the model was based only on data from PIONEER II; (iii) when the LSCF imputation rule was used, and; (iv) when the discontinuation rate for adalimumab non-responders after week 36 was based on the OLE study.

The ERG's critical appraisal identified a number of issues relating to the company's model and analysis. The most pertinent of these relate to: (i) the use of a model structure in which health gains and treatment continuation rules are defined according to depth of response, which does not reflect the pre-planned and validated HiSCR endpoint used in the PIONEER trials; (ii) the likely overestimation of the lifetime costs of HS surgery predicted by the company's model; (iii) the incorrect implementation of a continuation rule for adalimumab non-responders which does not mathematically reflect the actual assumptions stated in the CS; (iv) the use of arm-based aggregate data from the PIONEER I/II trials rather than a formal meta-analysis of relative treatment effects, and; (v) uncertainty surrounding the long-term transition probabilities derived from the PIONEER I/II trials and the M12-555 OLE study.

The ERG undertook eight sets of exploratory analyses based on the company's submitted model. The first three of these analyses represent the ERG's base case analysis. These include: (i) correction of technical programming errors in the company's model; (ii) applying structural amendments to the model to correctly reflect the company's intended adalimumab non-responder continuation rule during the maintenance phase; (iii) re-estimation of the costs of HS surgery. Further analyses were also undertaken to explore uncertainty surrounding the transition probabilities employed in the model, the likely impact of discontinuing non-responders and partial responders to adalimumab (during the induction phase only) and the potential structural uncertainty around the company's adopted modelling approach. The latter two analyses could not however be fully implemented due to the limitations of the company's model structure.

The ERG's base case analysis suggests that the probabilistic ICER for adalimumab versus standard care is expected to be £[redacted] per QALY gained. This is less favourable than the company's base case ICER. Additional analyses undertaken using this revised base case model indicate that: (i) using only PIONEER II to inform the model increases the ICER for adalimumab to £[redacted] per QALY gained; (ii) the exclusion of LOCF imputation using the M12-555 OLE GLM for patients receiving adalimumab and using the mean transition data from the maintenance phase of PIONEER I/II for adalimumab discontinuers and patients receiving standard care may reduce the ICER for adalimumab versus standard care, and; (iii) the discontinuation of partial responders at induction improves the ICER for adalimumab versus standard care. Owing to limitations in the structure of the company's model, the ERG was not fully able to assess the impact of modelling health gains, costs and adalimumab continuation rules according to the HiSCR ≥50% AN reduction threshold.

With respect to the company's economic analysis and the ERG's additional exploratory analyses, there remain several potentially important areas of uncertainty:

1. The impact of using relative treatment effects for adalimumab versus placebo based on a formal meta-analysis of data from PIONEER I and II within the model is unclear. Further, there is no comparative evidence regarding the long-term benefits of adalimumab relative to any other therapy.
2. The company's implemented model is subject to structural uncertainty, in particular around the definition of health states and the use of evidence to populate these. An alternative simpler model would have involved defining health utility, resource use, discontinuation rates, baseline transitions, relative treatment effects and adalimumab continuation rules according to the HiSCR ≥50% AN reduction threshold validated by Kimball *et al* and used as the primary endpoint in the PIONEER I/II trials.
3. The impact of adalimumab on the subsequent requirement or opportunity for surgical intervention is unclear. There is uncertainty around whether reductions in the overall costs of surgery predicted by the company's model will manifest in clinical practice. The impact of taking into account the use of other pharmacological therapies on the cost-effectiveness of adalimumab is also unknown.

8. REFERENCES

1. Jemec G. Clinical practice. Hidradenitis suppurativa. *New England Journal of Medicine* 2012;366:158-164.

2. Revuz J. Hidradenitis suppurativa. *Journal of the European Academy of Dermatology and Venereology* 2009;23:985-998.

3. Dufour D, Emtestam L, Jemec G. Hidradenitis suppurativa: a common and burdensome, yet under-recognised, inflammatory skin disease. *Postgraduate Medical Journal* 2014;90:216-21.

4. Von Der Werth J, Jemec G. Morbidity in patients with hidradenitis suppurativa. *British Journal of Dermatology* 2001;144:809-813.

5. Von der Werth J, Williams H. The natural history of hidradenitis suppurativa. *Journal of the European Academy of Dermatology and Venereology* 2000;14:389-392.

6. Kromann C, Deckers I, Esmann S. Risk factors, clinical course and long-term prognosis in hidradenitis suppurativa: A cross-sectional study. *British Journal of Dermatology* 2014;171:819-824.

7. Schrader A, Deckers I, van der Zee H. Hidradenitis suppurativa: a retrospective study of 846 Dutch patients to identify factors associated with disease severity. *Journal of the American Academy of Dermatology* 2014;71:460-467.

8. National Institute for Health and Care Excellence. *Adalimumab for treating moderate to severe hidradenitis suppurativa – final scope*. NICE: London; 2015.

9. AbbVie Ltd. *Adalimumab for treating moderate to severe hidradenitis suppurativa. ID812. Company evidence submission to the National Institute for Health and Care Excellence*. AbbVie: Berkshire, UK; 2015.

10. Jemec G, Revuz J. Treatment. In: Jemec G, Revuz J, Leyden J, editors. *Hidradenitis Suppurativa*. Heidelberg, Germany: Springer-Verlag; 2006.

11. Finlay A, Khan G. Dermatology Life Quality Index (DLQI) – a simple practical measure for routine clinical use. *Clinical and Experimental Dermatology* 1994;19:210-216.

12. European Medicines Agency. *Summary of product characteristics – adalimumab (Humira)*. EMA: London; 2015.

13. Zouboulis C, Desai N, Emtestam L. European S1 guideline for the treatment of hidradenitis suppurativa/acne inversa. *Journal of the European Academy of Dermatology and Venereology* 2015;29:619-644.

14. Blok J, van Hattem S, Jonkman M, Horváth B. Systemic therapy with immunosuppressive agents and retinoids in hidradenitis suppurativa: a systematic review. *British Journal of Dermatology* 2013;168:243-252.

15. Ingram J, McPhee M. Management of hidradenitis suppurativa: a UK survey of current practice. *British Journal of Dermatology* 2015;173:1070-1072.

16. Ingram JR, Woo PN, Chuna SL, Ormerod AD, Desai N, Kai AC. Interventions for hidradenitis suppurativa. *Cochrane Database of Systematic Reviews* 2015; October:10.

17. AbbVie Ltd. *Adalimumab for treating moderate to severe hidradenitis suppurativa – response to clarification questions*. AbbVie: Berkshire, UK; 2015.

18. AbbVie Ltd. *A Phase 3 Multicenter Study of the Safety and Efficacy of Adalimumab in Subjects with Moderate to Severe Hidradenitis Suppurativa – PIONEER I. Clinical Study Report (Interim), Adalimumab/Protocol*; 2015.

19. AbbVie Ltd. *A Phase 3 Multicenter Study of the Safety and Efficacy of Adalimumab in Subjects with Moderate to Severe Hidradenitis Suppurativa – PIONEER II (Clinical Study Report (Interim), Adalimumab/Protocol M11-810)*; 2015.

20. AbbVie Ltd. Phase 3 Open-Label Study of the Safety and Efficacy of Adalimumab in Subjects with Moderate to Severe Hidradenitis Suppurativa – PIONEER (Open-Label Extension) Clinical Study Report (Interim), Adalimumab/Protocol M12-555. 2015.

21. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013*; NICE: London; 2013.

22. Abbvie Ltd. *An observational cross-sectional study to describe the patient profile and secondary care management pathways for patients with hidradenitis suppurativa in the UK, investigator meeting* 2015.

23. Moher D, Tetzlaff J, Altman D, Group. TP. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 2009;6:e1000097. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2707599/pdf/pmed.1000097.pdf

24. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]*. Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org

25. Kimball A, Kerdel F, Adams D, Mrowietz U, Gefland J, Gniadecki R, *et al*. Adalimumab for the treatment of moderate to severe hidradenitis suppurativa: a parallel randomized trial. *Annals of Internal Medicine* 2012;157:846-855.

26. Miller I, Lynggaard C, Lophaven S, Zachariae C, Dufour DN, Jemec GB. A double-blind placebo-controlled randomized trial of adalimumab in the treatment of hidradenitis suppurativa. *British Journal of Dermatology* 2011;165:391-398.

27. European Medicines Agency. *Guideline on Missing Data in Confirmatory Clinical Trials*. London; 2011. Available from: http://www.ema.europa.eu/docs

28. Critical Appraisal Skills Programme (CASP). *Cohort Study Checklist*; 2013. Available from: http://www.casp-uk.net/#!casp-tools-checklists/c18f8

29. Kimball A, Jemec GB, Yang M, Kageleiry A, Signorovitch JE, Okun MM, *et al*. Assessing the validity, responsiveness and meaningfulness of the Hidradenitis Suppurativa Clinical Response (HiSCR) as the clinical endpoint for hidradenitis suppurativa treatment. *British Journal of Dermatology* 2014;171:1434-1442.

30. AbbVie Ltd. Safety and efficacy of adalimumab in patients with moderate to severe hidradenitis suppurativa: Results from first 12 weeks of PIONEER I, a phase 3, randomized, placebo-controlled trial (poster). 44th Annual European Society for Dermatological Research meeting; 2014; Copenhagen, Denmark. 2014.

31. AbbVie Ltd. Evaluating optimal medium-term dosing strategy for adalimumab in patients with moderate-to-severe Hidradenitis Suppurativa based on analysis of integrated results from the PIONEER I and II phase 3, randomized, placebo-controlled trials (poster). 23rd World Congress of Dermatology 2015; Vancouver, Canada. 2015.

32. AbbVie Ltd. Safety and efficacy of adalimumab in patients with moderate to severe hidradenitis suppurativa: Results from PIONEER II, a phase 3, randomized, placebo-controlled trial (poster). Advances in Cosmetic and Medical Dermatology (ACMD) Conference; 2015; Maui Hawaii. 2015.

33. Ingram JR. Hidradenitis suppurativa outcome measures and treatment goals. *British Journal of Dermatology* 2014:1293-1294.

34. Basra M, Salek M, Camilleri L, Sturkey R, Finlay A. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology* 2015;230:27-33.

35. Ingram J, Woo P, Chua S. *Interventions for hidradenitis suppurativa*; 2015. (unpublished)

36. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. *Value in Health* 2013;16:686-695.

37. BMJ Group, RCPCH Publications Ltd and The Royal Pharmaceutical Society of Great Britain. British National Formulary. *BNF June 2015*. Available from: www.bnf.org/.

38. Department of Health. *NHS Reference Costs 2013/14*. Available from: https://www.gov.uk/government/collections/nhs-reference-costs

39. Curtis L. *Costs of Health and Social Care 2014*. PSSRU: University of Kent, UK; 2014.

40. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. *Methods for the economic evaluation of health care programmes*. 3rd edition. Oxford University Press: Oxford; 2005.

41. Weinstein M, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C. Principles of good practice for decision analytic modeling in health care evaluation: Report of the ISPOR Task Force on good research practices--modeling studies. *Value in Health* 2003;6:9-17.

42. Menderes A, Sunay O, Vayvade H, Yilmaz M. Surgical management of hidradenitis suppurativa. *International Journal of Medical Sciences* 2010;7:240.

43. AbbVie Ltd. *Adalimumab for treating moderate to severe hidradenitis suppurativa – additional response to clarification questions*. Berkshire: AbbVie; 2015.

9. APPENDICES

Appendix 1 – Transition probabilities used in the company's model

(1) Adalimumab transition matrices

Table A1: Transition probabilities, weeks 0-2, adalimumab induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A2: Transition probabilities, weeks 2-4, adalimumab induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Partial-response	[redacted]		[redacted]		[redacted]		[redacted]	
Non-response	[redacted]		[redacted]		[redacted]		[redacted]	

Table A3: Transition probabilities, weeks 4-8, adalimumab induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]		[redacted]		[redacted]		[redacted]	
Response	[redacted]		[redacted]		[redacted]		[redacted]	
Partial-response	[redacted]		[redacted]		[redacted]		[redacted]	
Non-response	[redacted]		[redacted]		[redacted]		[redacted]	

Table A4: Transition probabilities, weeks 8-12, adalimumab induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]		[redacted]		[redacted]		[redacted]	
Response	[redacted]		[redacted]		[redacted]		[redacted]	
Partial-response	[redacted]		[redacted]		[redacted]		[redacted]	
Non-response	[redacted]		[redacted]		[redacted]		[redacted]	

Table A5: Transition probabilities, weeks 12-16, adalimumab maintenance (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response	[redacted]				
Partial-response	[redacted]				
Non-response	[redacted]				

Table A6: Transition probabilities, weeks 16-20, adalimumab maintenance (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response	[redacted]				
Partial-response	[redacted]				
Non-response	[redacted]				

Table A7: Transition probabilities, weeks 20-24, adalimumab maintenance (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response	[redacted]				
Partial-response	[redacted]				
Non-response	[redacted]				

Table A8: Transition probabilities, weeks 24-28, adalimumab maintenance (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response	[redacted]				
Partial-response	[redacted]				
Non-response	[redacted]				

Table A9: Transition probabilities, weeks 28-32, adalimumab maintenance (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response	[redacted]				
Partial-response	[redacted]				
Non-response	[redacted]				

Table A10: Transition probabilities, weeks 32-36, adalimumab maintenance (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A11: Transition probabilities, weeks 36+, adalimumab maintenance (GLM. PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	NR
Response	[redacted]	[redacted]	[redacted]	[redacted]	NR
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	NR
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	NR

** not reported*

(2) Adalimumab discontinuation transition matrices

Table A12: Transition probabilities, weeks 12-16, post-discontinuation (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A13: Transition probabilities, weeks 16-20, post-discontinuation (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A14: Transition probabilities, weeks 20-24, post-discontinuation (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A15: Transition probabilities, weeks 24-28, post-discontinuation (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response					
Partial-response					
Non-response					

Table A16: Transition probabilities, weeks 28-32, post-discontinuation (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response					
Partial-response					
Non-response					

Table A17: Transition probabilities, weeks 32-36, post-discontinuation (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response					
Partial-response					
Non-response					

Table A18: Transition probabilities, weeks 36+, post-discontinuation (GLM, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response					
Partial-response					
Non-response					

(3) Standard care transition matrices

Table A19: Transition probabilities, weeks 0-2, standard care induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]				
Response					
Partial-response					
Non-response					

Table A20: Transition probabilities, weeks 2-4, standard care induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A21: Transition probabilities, weeks 4-8, standard care induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A22: Transition probabilities, weeks 8-12, standard care induction (cross-tab, PIONEER I and II)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A23: Transition probabilities, weeks 12-16, standard care maintenance (cross-tab, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A24: Transition probabilities, weeks 16-20, standard care maintenance (cross-tab, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A25: Transition probabilities, weeks 20-24, standard care maintenance (cross-tab, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A26: Transition probabilities, weeks 24-28, standard care maintenance (cross-tab, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A27: Transition probabilities, weeks 28-32, standard care maintenance (cross-tab, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A28: Transition probabilities, weeks 32-36, standard care maintenance (cross-tab, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A29: Transition probabilities, weeks 36+, standard care maintenance (GLM, PIONEER II only)

From	To				N observations
	High response	Response	Partial response	Non-response	
High response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Appendix 2: Technical details of amendments to company's model within the ERG's exploratory analyses

ERG Exploratory Analysis 1: Correction of model errors

Item no.	Worksheet reference	Cell reference	Description of amendment	Rationale for amendment
1	Life Table	N10:N80 and O10:O80	Amended to reflect number of days in year	Model previously assumed 364 days per year
2	Base Case Results	O11	Weekly discount rate amended to reflect number of days in year	Model previously assumed 364 days per year
3	Markov Trace – ADA	DR10:DR13, DR15:DR869	Amended to reflect number of days in year	Model previously assumed 364 days per year
4	Markov Trace – SC	BQ10:BQ13, BQ15:BQ869	Amended to reflect number of days in year	Model previously assumed 364 days per year
5	Markov Trace – ADA	BK9:BK13, BL9:BL13, BM9:BM13, BN9:BN13, BP9:BP13, BQ9:BQ13, BR9:BR13, BS9:BS13, BK15:BK869, BL15:BL869, BM15:BM869, BN15:BN869, BP15:BP869, BQ15:BQ869, BR15:BR869, BS15:BS869, DU9:DU13, DU15:DU869	Amended to reflect number of days in year	Model previously assumed 364 days per year
6	Markov Trace – SC	S9:S13, T9:T13, U9:U13, V9:V13, X9:X13, Y9:Y13, Z9:Z13, AA9:AA13, S15:S869, T15:T869, U15:U869, V15:V869, X15:X869, Y15:Y869, Z15:Z869, AA15:AA869, BT9:BT13, BT15:BT869	Amended to reflect number of days in year	Model previously assumed 364 days per year
7	Markov Trace – ADA	BK876:BY876, CG876:CU876	Half of cycle 0 costs subtracted from total costs	Half-cycle correction applied incorrectly
8	Markov Trace – ADA	CX876:DG876	Half of cycle 0 QALYs subtracted from total QALYs	Half-cycle correction applied incorrectly
9	Markov Trace – SC	S876:AG876, AK876:AY876	Half of cycle 0 costs subtracted from total costs	Half-cycle correction applied incorrectly

Item no.	Worksheet reference	Cell reference	Description of amendment	Rationale for amendment
10	Markov Trace – SC	BB876:BK876	Half of cycle 0 QALYs subtracted from total QALYs	Half-cycle correction applied incorrectly
11	Markov Trace - ADA	BE9:BH13	Additional 4-weeks of adalimumab included for patients in high response, response and partial response states	Adalimumab costs applied in wrong cycle
12	Markov Trace - ADA	BE876:BI876, CA876:CE876	Lifetime costs of treatment to include treatment received in cycle 0	Lifetime costs of treatment only included costs beginning in cycle 1

ERG Analysis 2: Incorporation of tunnel states to reflect the maintenance phase adalimumab non-responder continuation rule (including ERG Exploratory Analysis 1)

Structural amendments not shown in appendix; see "ERG_tunnels" worksheet in ERG base case model.

ERG Exploratory Analysis 3: Revised assumptions regarding costs of HS surgery (including ERG Exploratory Analyses 1 and 2)

Including the amendments detailed above, apply a cost of £1525.74 to worksheet "Costs & Resource Use" cell J53.

ERG Additional Exploratory Analysis 4: Use of PIONEER II data only

Using ERG base case model, select "M11-810 only" option in worksheet "Base Case Results" cells J17 and J18.

ERG Additional Exploratory Analysis 5: Alternative assumptions regarding transition probabilities beyond week 36

(i) Using ERG base case model, apply lower transition matrix shown in Table 51 to worksheet "ADA – TP" cells H104:K107.

(ii) Using ERG base case model, select "Mean TP of Weeks 12-36 applied forward" option in worksheet "Base Case Results" cell J19.

ERG Additional Exploratory Analysis 6: Discontinuation of partial responders and non-responders at 12-weeks

Using ERG base case model, set worksheet "Markov Trace – ADA" cell AR14=0 and BA14=I14.

ERG Additional Exploratory Analysis 7: Assumption of no difference in utility, resource use and discontinuation rates for non-responders and partial responders, and for high responders and responders

Using the ERG base case model, apply amendments to utilities, resource use and discontinuation rates as per the company's re-analysis provided in response to clarification question B2.

ERG Additional Exploratory Analysis 8: Assumption of no difference in utility, resource use and discontinuation rates for non-responders and partial responders, and for high responders and responders with discontinuation of patients achieving only partial response or no response at 12-weeks

Using the ERG base case model, apply amendments to utilities, resource use and discontinuation rates as per the company's re-analysis provided in response to clarification question B2. Set worksheet "Markov Trace – ADA" cell AR14=0 and BA14=I14.