CONFIDENTIAL UNTIL PUBLISHED
Evidence Review Group's Report

Guselkumab for treating active psoriatic arthritis after inadequate response to DMARDs

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD

Authors

Ana Duarte, Research Fellow, CHE, University of York

Mark Corbett, Research Fellow, CRD, University of York

Hollie Melton, Research Fellow, CRD, University of York

Kath Wright, Information specialist, CRD, University of York

Marta Soares, Senior Research Fellow, CHE, University of York

Claire Rothery, Senior Research Fellow, CHE, University of York

Mark Simmonds, Senior Research Fellow, CRD, University of York

Correspondence to
Mark Simmonds
Centre for Reviews and Dissemination
University of York
York, YO10 5DD

Date completed 15/10/2020

Source of funding

This report was commissioned by the NIHR Systematic Reviews Programme as project number ID1658.

Declared competing interests of the authors

None

Acknowledgements

We thank Dr Deepak Jadon, Consultant in Rheumatology, at the Cambridge University Hospitals NHS Foundation Trust, and Professor Laura Bojke, Centre for Health Economics at University of York for their expert advice during the project.

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:

Duarte A, Corbett M, Melton H, Wright K, Soares M, Rothery C, Simmonds M. Guselkumab for treating active psoriatic arthritis after inadequate response to DMARDs. A Single Technology Appraisal. CRD/CHE University of York ERG. 2020.

Contributions of authors

Ana Duarte performed the economic analyses and wrote Sections 4, 5 and 6 of the report.
Mark Corbett performed the clinical effectiveness review and wrote Sections 2 and 3 of the report.
Hollie Melton performed the clinical effectiveness review and wrote Sections 2 and 3 of the report.
Kath Wright performed database searches and provided information and library support.
Marta Soares provided expert advice on the economic analyses and the report as a whole.
Claire Rothery performed economic analyses, wrote Sections 4, 5 and 6 of the report, led the overall economic analyses and takes joint responsibility for the report as a whole.
Mark Simmonds reviews statistical components of the review (NMAs, Section 3), led the overall clinical effectiveness review and takes joint responsibility for the report as a whole.

Note on the text

All commercial-in-confidence (CIC) data have been [redacted], all academic-in-confidence (AIC) data are [redacted], all depersonalised data (DPD) are [redacted].

Copyright statement

Copyright belongs to the University of York. Copyright is retained by Janssen for Tables 3, 4 and 5 and figures 4 and 6Figure 6.

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

Abbreviation	Definition
ACR	American College of Rheumatology
ACR20/50/70 20%/50%/70%	Improvement in the ACR response criteria
ADA	Adalimumab
AE	Adverse event
APL	Apremilast
APR	Apremilast
bDMARD	Biological disease-modifying anti-rheumatic drug
BSC	Best supportive care
BUGS	Bayesian inference using Gibbs sampling
cDMARD	Conventional disease-modifying anti-rheumatic drug
CERT	Certolizumab pegol
CHE	Centre for Health Economics
CI	Confidence interval
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CONSORT	Consolidated Standards of Reporting Trials
CRP	C-reactive protein
CS	Company submission
CSR	Clinical study report
CZP	Certolizumab pegol
DAPSA	Disease activity in psoriatic arthritis
DAS28	Disease activity score - 28 joints
DIP	Distal interphalangeal
DMARD	Disease-modifying anti-rheumatic drug
EQ-5D	EuroQol-5 Dimensions
ERG	Evidence Review Group
ETN	Etanercept
GOL	Golimumab
GUS	Guselkumab
HAQ-DI	Health Assessment Questionnaire-Disability Index
IL	Interleukin
INF	Infliximab
IXE	Ixekizumab

LEI	Leeds enthesitis index
MDA	Minimal disease activity
MTX	Methotrexate
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
NSAIDs	Non-steroidal anti-inflammatory drugs
PASI	Psoriasis Area and Severity Index
PASI 50/75/90	% or greater improvement in PASI score
PBO	Placebo
PLA	Placebo
PRISMA	Preferred reporting index for systematic reviews and meta-analyses
PsA	Psoriatic Arthritis
PsARC	Psoriatic Arthritis Response Criteria
PsO	Psoriasis
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
SAE	Serious adverse event
SD	Standard deviation
SEC	secukinumab
SF36	Short-Form Health Survey 36-item
SJC	Swollen joint count
SmPC	Summary of Product characteristics
STA	Single Technology Appraisal
TNF	Tumour necrosis factor
TOF	Tofacitinib
UST	ustekinumab

1 EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report.

All issues identified represent the ERG's view, not the opinion of NICE.

ID1658	Summary of issue	Report sections
3.1	• Trial populations may not be representative of population seen in UK clinical practice • Guselkumab trials had more severe levels of skin symptoms (PASI) than usually seen in the UK, and lower prior use of cDMARDs.	Section 3.2.2.1
3.2	• "Early escape" may bias results in guselkumab trials • Treatment non-responders at 16 weeks were assumed to be non-responders in the 24 week analyses of guselkumab trials. This may bias results in favour of guselkumab	• Section 3.2.1 • Section 3.2.2.2 • Section 3.2.2.3
3.3	• DISCOVER trials had very low treatment discontinuation rates • Guselkumab trials had much lower discontinuation rates than most other trials/treatments.	• Section 3.3.1
3.4	• Assessment time point for guselkumab response to treatment at 24 weeks • The assessment time point for response to treatment for guselkumab is based on a stopping rule at 24 weeks, in anticipation of its expected marketing authorisation for PsA, but it is unclear why a 16-week assessment time point should not also be considered.	• Section 3.4.1 • Section 3.4.2.1
3.5	• Network meta-analyses adjust for placebo group response The company's preferred NMAs adjusted for varying placebo response rates in the included trials. • While the ERG agrees that this was a reasonable analysis, it affects treatment effects and rankings.	• Section 3.4.1 • Section 3.4.2.1
3.6	• Alternative response definition • Treatment continuation is defined based on achieving a PsARC response. However, consideration may be given to the possibility of continuation on treatment for patients whose PsARC response does not justify continuation but who demonstrate a PASI 75 response.	• Section 3.4.1 • Section 3.4.2.1
4.1	• Modelled treatment sequences • The modelled treatment sequences do not reflect the range of treatment sequences seen in UK clinical practice.	• Section 4.2.5

	Treatment with multiple lines of active therapy, instead of placing patients on best supportive care (BSC) at second and third line, is considered a valid treatment strategy.	
4.2	• Treatment-specific discontinuation rates • The use of treatment-specific discontinuation rates in the maintenance period of the model (after initial response to treatment) is informed by a limited and potentially biased evidence base towards guselkumab.	• Section 4.2.6 • Section 6.1.3
4.3	• Cost-effectiveness results for baseline unadjusted NMA • Results based on placebo-response unadjusted models for PsARC and PASI (with the exception of PsARC response in the biologic-experienced population) are not presented.	• Section 4.2.8 • Section 6.1.5
4.4	• Administration costs • Administration costs should be included for drugs administered subcutaneously.	• Section 4.2.12.2 • Section 6.1.6
4.5	• Monitoring costs • Monitoring resource use and costs should be the same across treatments to be consistent with previous TAs in PsA.	• Section 4.2.12.3 • Section 6.1.7
4.6	• Arthritis related costs • The Kobelt et al., 2002, algorithm should be used to estimate arthritis related costs in the base-case analysis to ensure consistency with previous TAs in PsA.	• Section 4.2.12.4
4.7	• Adverse events • The company's approach to include adverse events in the model is unlikely to reflect the safety profile of the different treatments and is not consistent with the assumptions of previous TAs in PsA.	• Section 4.2.9

1.1 Overview of the ERG's key issues

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are (i) the use of an equivalent annual treatment discontinuation rate across all therapies; (ii) the exclusion of adverse events; (iii) the use of an alternative source for arthritis related costs; (iv) inclusion of an administration cost for subcutaneous drugs; and (v) equivalent monitoring resource use across all treatments.

The ERG's preferred assumptions are aimed at ensuring consistency with previous Technology Appraisals (TAs) in psoriatic arthritis (PsA). Where the company has not presented compelling evidence to support a change from recent previous TAs (namely, TA445, TA537 and TA543) the ERG's preferred base case is in line with the assumptions used in these TAs.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Treatment response based on PsARC, HAQ-DI and PASI scores
- All-cause treatment discontinuation rates
- HAQ-DI associated with progression of uncontrolled psoriatic arthritis.

Overall, the technology is modelled to affect costs by:

- Having a longer assessment time point of 24 weeks for response to treatment compared to most of the comparator therapies (12-16 weeks)
- Having a lower treatment discontinuation rate compared to the comparator therapies
- [redacted]
- Disease-related costs based on changes in HAQ-DI scores, which are conditional on PsARC response.

The modelling assumptions that have the greatest effect on the ICER are:

- Treatment-specific annual discontinuation rates, with a much lower discontinuation rate for guselkumab compared to the other biologically therapies
- A maximum of two active therapies modelled before final line best supportive care.

1.3 The decision problem: summary of the ERG's key issues

The ERG notes that patients receiving guselkumab after one previous cDMARD (Population 1) were excluded from the submission due to lack of evidence. Clinical advice suggests that guselkumab is unlikely to be used in such patients, so the ERG did not consider this population further.

The ERG also notes that the patient population considered in the submission may not be representative of the UK. This is covered in Section 1.4.

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

The ERG's primary concerns with the clinical effectiveness evidence relate to whether the trial populations are representative of UK populations, and the preference for assessing outcomes at 24 weeks for guselkumab.

Issue 3.1 Trial populations may not be representative of population seen in UK clinical practice

Report section	Section 3.2.2.1
Description of issue and why the ERG has identified it as important	Trials of guselkumab, and comparator treatments, were generally conducted in Eastern Europe. Clinical advice to the ERG suggested that the populations had more severe disease (particularly skin symptoms as measured by PASI) than would usually be seen in the UK. Patients in the trials were also less likely to have received prior cDMARD therapies than is typical for the UK. As guselkumab appears more effective at improving PASI, this may bias overall conclusions in favour of guselkumab.
What alternative approach has the ERG suggested?	For clinical evaluation no alternatives are feasible, given the absence of UK trial data. The ERG conducted an exploratory scenario for the level of psoriasis severity and baseline PASI and HAQ-DI score in the population based on clinical opinion relevant to UK practice.
What is the expected effect on the cost-effectiveness estimates?	Total costs decrease and total QALYs increase for all treatments across all populations because of the alternative assumption of a higher proportion of patients with less severe psoriasis in the population. The ICER for guselkumab vs. the next best option increased [redacted]
What additional evidence or analyses might help to resolve this key issue?	Further clarification would require trial or observational evidence from the UK (or similar population), which is unlikely to exist at present.

Issue 3.2 "Early escape" may bias results in guselkumab trials

Report section	Section 3.2.1 Section 3.2.2.2 Section 3.2.2.3
Description of issue and why the ERG has identified it as important	The DISCOVER trials of guselkumab permitted an "early escape" at 16 weeks for patients not responding to treatment (mostly in the placebo arm). These patients were assumed to be non-responders in the main analyses at 24 weeks. The 24 week analysis may therefore have overestimated the effectiveness of guselkumab.
What alternative approach has the ERG suggested?	The ERG analysed the DISCOVER trials at 16 weeks only, and including outcome data on early escape patients at 24 weeks. In both cases the effectiveness of guselkumab relative to placebo was reduced.
What is the expected effect on the cost-effectiveness estimates?	Refer to modifications to model using 16 week data (Issue 3.4)
What additional evidence or analyses might help to resolve this key issue?	The ERG considers that existing data presented for DISCOVER trials are sufficient for this issue.

Issue 3.3 DISCOVER trials had very low treatment discontinuation rates

Report section	Section 3.3.1
Description of issue and why the ERG has identified it as important	The DISCOVER trials had very low rates of discontinuation of treatment than for most other trials and treatments, including in the placebo arms. If this is not representative of discontinuation in actual practice it may bias the economic analyses.
What alternative approach has the ERG suggested?	See Issue 4.2 in Section 1.5
What is the expected effect on the cost-effectiveness estimates?	See Issue 4.2 in Section 1.5
What additional evidence or analyses might help to resolve this key issue?	Data from the UK (or a similar country) on discontinuation rates with guselkumab would be required.

Issue 3.4 Assessment time point for guselkumab response to treatment

Report section	Section 3.4.1 Section 3.4.2.1 Section 4.2.5 Section 4.2.8 Section 6.1.4 Section 6.2.3
Description of issue and why the ERG has identified it as important	The assessment time point for response to treatment for guselkumab is based on a stopping rule at 24 weeks, in anticipation of its expected marketing authorisation for PsA, but it is unclear why a 16-week assessment time point should not also be considered in decision making. Furthermore, the NMAs combined trials assessing outcomes at different time points: guselkumab at 24 weeks, while other treatments were assessed at 16 weeks or earlier (usually according to licenced indication). This could potentially bias results of NMAs in favour of guselkumab given the longer treatment exposure. The NMA results suggest that while response for guselkumab in terms of arthritis symptoms is mostly achieved by 16 weeks, psoriasis symptoms continue to improve at least until 24 weeks. However, there remains uncertainty as to whether this improvement in PASI response from 16 to 24 weeks for patients treated with guselkumab is not confounded (and to what extent) by the bias potentially introduced by allowing "early escape" in the DISCOVER trials (see issue 3.2).
What alternative approach has the ERG suggested?	The ERG requested re-analyses of the NMAs using guselkumab outcomes assessed at 16 weeks. These were provided by the company at points for clarification. However, the structure of the electronic model submitted by the company does not allow capturing the potential continued improvement in PASI response between 16 and 24 weeks for patients treated with guselkumab. Therefore, the model is not suitable to explore the full impact on outcomes of using a 16-week stopping rule for guselkumab, as it could misrepresent the QALY gains associated with an improvement in PASI response from 16 to 24 weeks. The ERG undertook an exploratory analysis to assess the impact on cost of treatment at first line of therapy for a 16-week stopping rule.
What is the expected effect on the cost-effectiveness estimates?	The difference in guselkumab acquisition costs for a 16-week vs. 24 week stopping rule ranges from [redacted]
What additional evidence or analyses might help to resolve this key issue?	The structure of the electronic model would need to be revised to allow flexibility to consider alternative response criterion and assessment time point of 16 weeks for guselkumab.

Issue 3.5 Network meta-analyses adjust for placebo group response

Report section	Section 3.4.1 Section 3.4.2.1
Description of issue and why the ERG has identified it as important	The company's preferred NMAs adjusted for varying placebo response rates in the included trials. While the ERG agrees that this was a reasonable analysis, it did lead to larger treatment effects and higher treatment rankings for guselkumab compared to other treatments than when using unadjusted analyses.
What alternative approach has the ERG suggested?	The ERG requested re-analyses of the NMAs without placebo adjustment where these had not been reported. These were provided by the company.
What is the expected effect on the cost-effectiveness estimates?	See Issue 4.3 in Section 1.5.
What additional evidence or analyses might help to resolve this key issue?	The ERG considers that additional analyses provided by the company are sufficient for this issue.

Issue 3.6 Alternative response definition

Report section	Section 3.2.1
Description of issue and why the ERG has identified it as important	Treatment continuation is defined based on achieving a PsARC response. However, consideration may be given to the possibility of continuation on treatment for patients whose PsARC response does not justify continuation but who demonstrate a PASI 75 response. This is particularly relevant for guselkumab where the probability of PsARC response is comparable to other bDMARDs, but it is demonstrated to have the highest PASI 75 response at week 24.
What alternative approach has the ERG suggested?	Modelling the continuation of treatment based on achieving PsARC or PASI 75 response. This may be an important consideration for guselkumab in relation to the 16 week versus 24 week stopping rule, where the strength of guselkumab relative to the other DMARDsbDMARDs appears to be in achieving a higher PASI 75 response, but this is achieved at week 24 rather than week 16.
What is the expected effect on the cost-effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	The trial data would need to be reanalysed to provide an alternative response definition of PsARC or PASI 75 and the structure of the electronic model revised to allow additional flexibility to consider this response criterion.

1.5 The cost-effectiveness evidence: summary of the ERG's key issues

Issue 4.1 Modelled treatment sequences

Report section	Section 4.2.5
Description of issue and why the ERG has identified it as important	The modelled treatment sequences do not reflect the range of treatment sequences seen in UK clinical practice. Clinical opinion suggests that switching among different TNFi therapies represents a valid treatment strategy as well as switching to IL modulators (ustekinumab, secukinumab and ixekizumab) or tofacitinib. Importantly, treatment with multiple lines of active therapy, instead of placing patients on best supportive care (BSC) at second and third line, is considered a valid treatment strategy.
What alternative approach has the ERG suggested?	The ERG suggests modelling of multiple lines of therapy instead of placing patients on BSC at second and third line; patients are unlikely to receive only two active therapies in the biologic-naïve or TNFi contraindicated populations, or only one active therapy after ≥1 TNFi's in the biologic-experienced population. However, the ERG recognises that there is unlikely to be a standardised approach to treatment sequencing in UK clinical practice and, therefore, identifying the choice of treatment sequences to model is challenging to address. The approach used by the company aligns with previous TAs in PsA.
What is the expected effect on the cost-effectiveness estimates?	The ERG has not considered alternative treatment sequences due to the absence of a standardised approach and lack of evidence to inform the effectiveness of switching between active therapies. However, the ERG highlights that the treatment discontinuation rates used in the model have a material impact on the ICERs because the modelled treatment duration may not be sufficiently long enough with multiple lines of active therapy before reaching final line BSC.
What additional evidence or analyses might help to resolve this key issue?	In the absence of registry data on switching to IL modulators (ustekinumab, secukinumab and ixekizumab) or tofacitinib, the frequency of therapy switching and outcomes among PsA patients is unknown.

Issue 4.2 Treatment-specific discontinuation rates

Report section	Section 4.2.6 Section 6.1.3
Description of issue and why the ERG has identified it as important	The use of treatment-specific discontinuation rates in the maintenance period of the model (after initial response to treatment) is informed by a limited and potentially biased evidence base towards guselkumab. In addition, treatment-specific discontinuation rates should only be modelled when the appropriate range of treatment sequences are considered that reflect the full duration of disease.
What alternative approach has the ERG suggested?	In the absence of strong evidence to support treatment-specific discontinuation rates, and the fact that more than two lines of active therapy are not modelled, the ERG considers an equivalent annual treatment discontinuation rate for all active therapies to be appropriate. This approach would align with the discontinuation rates used in previous TAs in PsA.
What is the expected effect on the cost-effectiveness estimates?	The discontinuation rate is delaying time to BSC, which has the lowest QALY benefits and highest disease-related costs. Therefore, treatments with a lower discontinuation rate have higher QALY benefits and lower disease-related costs. When the discontinuation rates are equivalent across therapies, the difference in time on treatment between the alternative active therapies is reduced making the total QALY benefits and non-drug related costs more similar across all treatments. If the modelled treatment duration was sufficiently long enough with multiple lines of active therapy before reaching final line BSC, the impact of the discontinuation rate would be less important.
What additional evidence or analyses might help to resolve this key issue?	In the absence of long-term registry or trial data on discontinuation rates after initial response, the treatment-specific discontinuation rates are uncertain.

Issue 4.3 Cost-effectiveness results for baseline unadjusted NMA

Report section	Section 4.2.8 Section 6.1.5
Description of issue and why the ERG has identified it as important	Cost-effectiveness results based on placebo-response unadjusted models for PsARC and PASI (with the exception of PsARC response in the biologic-experienced population) are not presented. In baseline risk-adjusted models, the relative effectiveness of treatments trialled under a high placebo response (such as guselkumab) are adjusted upwards, while the relative effectiveness of treatments trialled under a low placebo response are adjusted downwards. Without a clear rationale for the placebo effect, the results must be interpreted with caution, especially in relation to trying to distinguish (and rank) treatments that achieve fairly similar response rates.
What alternative approach has the ERG suggested?	The ERG considers that the cost-effectiveness results for both placebo-response adjusted and unadjusted models should be presented, in line with the approach used in previous TAs for PsA.
What is the expected effect on the cost-effectiveness estimates?	Lower total costs and QALYs for guselkumab compared to the company's base-case results for the biologic-naïve subpopulation. Etanercept is no longer [redacted] and is the cost-effective treatment, while the ICER of guselkumab compared to etanercept [redacted]. For the TNFi-contraindicated subpopulation, guselkumab remains the cost-effective treatment, albeit with a slightly higher ICER compared to BSC.
What additional evidence or analyses might help to resolve this key issue?	None required.

1.6 Other key issues: summary of the ERG's view

Issue 4.4 Administration costs

Report section	Section 4.2.12.2 Section 6.1.6
Description of issue and why the ERG has identified it as important	Administration costs should be included for drugs administered subcutaneously.
What alternative approach has the ERG suggested?	Inclusion of administration costs for drugs administered subcutaneously (i.e., all treatments except infliximab, tofacitinib, apremilast and BSC).
What is the expected effect on the cost-effectiveness estimates?	Small increase in the costs of drugs administered subcutaneously without a sizeable impact on the ICERs.
What additional evidence or analyses might help to resolve this key issue?	None required

Issue 4.5 Monitoring costs

Report section	Section 4.2.12.3 Section 6.1.7
Description of issue and why the ERG has identified it as important	Monitoring resource use and costs should be the same across treatments to be consistent with previous TAs in PsA.
What alternative approach has the ERG suggested?	Inclusion of the same monitoring costs across all treatments
What is the expected effect on the cost-effectiveness estimates?	Reduction of costs across all treatments with the magnitude of this reduction depending on the treatment route of administration, the length of the trial period, and the number of treatments in the sequence. Overall, this has a marginal impact on the estimates of cost-effectiveness.
What additional evidence or analyses might help to resolve this key issue?	None required.

Issue 4.6 Arthritis related costs

Report section	Section 4.2.12.4
Description of issue and why the ERG has identified it as important	The Kobelt et al., 2002, algorithm should be used to estimate arthritis related costs in the base-case analysis to ensure consistency with previous TAs in PsA.
What alternative approach has the ERG suggested?	In the absence of strong evidence to support the use of McHugh et al., 2019, the Kobelt et al., 2002, algorithm is used to estimate the arthritis related cost
What is the expected effect on the cost-effectiveness estimates?	Overall, the use of an alternative algorithm reduces total costs with the magnitude of the reduction being conditional on the treatment specific effect on the HAQ-DI outcomes. This alternative assumption does not, however, have a significant impact on ICERs.
What additional evidence or analyses might help to resolve this key issue?	A larger study with a design similar to McHugh et al., 2019, including sufficient patients across the full range of arthritis severity, as measure by HAQ-DI scores.

Issue 4.7 Adverse events

Report section	Section 4.2.9
Description of issue and why the ERG has identified it as important	The company's approach to include adverse events in the model is unlikely to reflect the safety profile of the different treatments and is not consistent with the assumptions of previous TAs in PsA.
What alternative approach has the ERG suggested?	Exclusion of SAEs.
What is the expected effect on the cost-effectiveness estimates?	Mean total costs decrease for all treatments, and mean total QALYs increase.
What additional evidence or analyses might help to resolve this key issue?	An analysis where individual treatment effects are modelled, as opposed to one assumed homogeneous type of adverse event.

1.7 Summary of ERG's preferred assumptions and resulting ICER

Scenario	Incremental cost	Incremental QALYs	ICER (change from company base case)
Company's base case – all biologic-naïve and TNFi contraindicated	[redacted]	3.497	[redacted]
Company's base case – all biologic-experienced	[redacted]	2.546	[redacted]
ERG's preferred base case – with placebo response adjusted effectiveness estimates • All biologic-naive	-	-	[redacted]
ERG's preferred base case – with placebo response adjusted effectiveness estimates • All biologic-experienced	[redacted]	1.259	[redacted]
ERG's preferred base case – with placebo response adjusted effectiveness estimates • All TNFi contra-indicated	[redacted]	2.370	[redacted]
ERG's preferred base case – with effectiveness estimates unadjusted for placebo response • All biologic-naive			[redacted]
ERG's preferred base case – with effectiveness estimates unadjusted for placebo response • All biologic-experienced	[redacted]	1.261	[redacted]
ERG's preferred base case – with effectiveness estimates unadjusted for placebo response • All TNFi contra-indicated	[redacted]	2.260	[redacted]

Modelling errors identified by the ERG and corrected by the company are described in Section 5.3.

For further details of the exploratory and sensitivity analyses done by the ERG, see Section 6.1.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

In this report the ERG has reviewed the clinical and cost-effectiveness evidence submitted by Janssen in support of guselkumab (Tremfya) for adults with active psoriatic arthritis following inadequate response to previous disease-modifying anti-rheumatic drugs (DMARDs). The ERG notes that at the time of writing the report, the marketing authorisation for guselkumab for the treatment of psoriatic arthritis had not been granted. In this section the ERG critiques the company's proposed positioning of guselkumab in the treatment pathway and its definition of the decision problem when compared with the NICE scope.

2.2 Background

Figure 7 in the CS reported the proposed positions of guselkumab in the treatment pathway, which were based on the subpopulations detailed in the NICE scope. The ERG agrees with the company's approach of aligning the proposed use of guselkumab with NICE guidance on biological DMARDs which states that they should be given after failure of two or more conventional non-biological DMARDs, i.e. analyses were not presented in the CS for subpopulation 1 (after one conventional DMARD). Guselkumab was therefore proposed as being used in biologic-naïve patients who have not responded (or who have been intolerant) to two or more prior cDMARDs (subpopulation 2), biologic-experienced patients (subpopulation 3) and patients for whom TNFα inhibitors were contraindicated (subpopulation 4).

The CS stated that guselkumab is an IL-23 inhibitor which has a new mechanism of action which works 'upstream' of existing treatments, potentially leading to important treatment improvements from a patient perspective. The company added that they expected that guselkumab will be regarded as especially appropriate for patients who have a strong desire for 'holistic control' of their symptoms, especially skin symptoms, or who have concerns about adverse events and contraindications related to IL-17 modulators. Guselkumab was also reported to address unmet needs by having several features such as low discontinuation rates and high rates of resolution of skin psoriasis. The CS also stated that guselkumab's convenient dosing schedule (most patients will self-administer 8-weekly doses subcutaneously) is also valued by patients who do not want to be 'defined' by their disease.

The ERG acknowledges that guselkumab offers a new mechanism of action and therefore provides a different treatment option which will be welcomed by both patients and clinicians. The submission from the Psoriasis Association stated that there was an unmet need in patients with psoriatic arthritis since "*the heterogeneity of the disease means that there is not a "one size fits all" in terms of which treatment will work, for how long and with manageable side effects*." The ERG's clinical adviser also

considered that having another treatment option, such as guselkumab, would be helpful. It is also possible that the use of 8-weekly dosing may improve treatment compliance when compared to other biologics which are given weekly or fortnightly.

The CS stated that it is highly likely that guselkumab is better than all TNFα inhibitors and many IL therapies for skin responses assessed using PASI scores. However, the ERG's clinical adviser considers that, in the NHS, the decision to initiate a biologic for psoriatic arthritis is not driven by the severity of skin symptoms but by the severity of joint symptoms and that in his large clinic he very rarely sees patients with PASI scores greater than 5. The ERG's clinical adviser also stated that, for many NHS patients, adalimumab or etanercept would be considered as a first line biologic due to their well-known efficacy and safety profiles and the availability of biosimilars. The ERG therefore thinks that guselkumab would very rarely be considered as a first-line biologic in the NHS. The ERG also has concerns about the company's data to support the statement about the low discontinuation rates ascribed to guselkumab (see Section 3.3.1).

2.3 Critique of company's definition of decision problem

The ERG's comments on the company's definition of the decision problem are summarised in Table 1.

Table 1 Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Adults with active psoriatic arthritis whose disease has not responded adequately or who have been intolerant to a previous conventional DMARD therapy or biologic DMARD therapy or for whom DMARD therapy is contraindicated.	Adults with active psoriatic arthritis: • whose disease has not responded adequately or who have been intolerant to two previous cDMARD therapies but have not received bDMARD therapy (**biologic naïve; subpopulation 2**). • whose disease has not responded adequately or who have been intolerant to two previous cDMARD therapies and at least one bDMARD therapy (**biologic experienced; subpopulation 3**). • for whom TNFi therapy is contraindicated (**TNFi contraindicated; subpopulation 4**).	Patients whose disease has not responded adequately to only one previous cDMARD therapy (subpopulation 1) have not been included in this submission due to limitations in data availability.	The applicability of the guselkumab trials to the NHS setting was poor. See Section 3.2.2.1. Of particular note: • Most patients had not received two or more prior cDMARDs • Baseline PASI scores were much higher than would be seen in NHS clinics • Many patients were ineligible for the trials for having C-reactive protein levels of <3mg/l or <6mg/l at screening visits • Only a small subgroup of biologic-experienced patients was recruited • No UK trial sites
Intervention	Guselkumab alone or in combination with conventional disease modifying anti-rheumatic drugs (DMARDs)	Guselkumab alone or in combination with cDMARDs.	Same as final scope issued by NICE.	An appropriate stopping rule will be needed for guselkumab. The company anticipates this will be recommended as being 24 weeks in the SmPC though the ERG considers that the evidence to support this is limited – see Section 3.2.2.4

				[redacted]
Comparators	For people who have only received 1 previous conventional disease modifying anti-rheumatic drug (DMARD) • Conventional DMARDs For people whose disease has not responded adequately to at least 2 conventional DMARDs: • Biological DMARDs (with or without methotrexate • including etanercept, adalimumab, infliximab, • golimumab, certolizumab pegol, ixekizumab and secukinumab) • Apremilast • Tofacitinib For people whose disease has not responded adequately to conventional DMARDs and 1 or more TNF-alpha inhibitors: • Ustekinumab • Secukinumab • Certolizumab pegol • Tofacitinib • Ixekizumab	For people whose disease has not responded adequately to at least 2 cDMARDs (biologic naïve; subpopulation 2): • bDMARDs (with or without methotrexate including etanercept, adalimumab, infliximab, golimumab, certolizumab pegol, ixekizumab and secukinumab) • Apremilast • Tofacitinib For people whose disease has not responded adequately to cDMARDs and 1 or more TNFi (biologic experienced; subpopulation 3): • Ustekinumab • Secukinumab • Certolizumab pegol • Tofacitinib • Ixekizumab • Best supportive care	Same as the final scope issued by NICE for the populations included in the submission. cDMARDs were considered the relevant comparator for subpopulation 1 only and as such are not included in the decision problem.	All relevant comparators were studied, though the ERG notes the lack of direct evidence comparing guselkumab with any of the comparators. The evidence used to compare treatments in the CS is derived from indirect comparisons via network meta-analyses. These estimates of efficacy are less reliable than those derived from direct evidence.

	• Best supportive care For people in whom TNF-alpha inhibitors are contraindicated or not tolerated: • Ustekinumab • Secukinumab • Ixekizumab • Tofacitinib • Best supportive care	For people in whom TNFi are contraindicated (TNFi contraindicated; subpopulation 4): • Ustekinumab • Secukinumab • Ixekizumab • Tofacitinib • Best supportive care		
Outcomes	The outcome measures to be considered include: • disease activity functional capacity • disease progression • periarticular disease (for example enthesitis, tendonitis, dactylitis) • axial outcomes • mortality • adverse effects of treatment • health-related quality of life.	The following outcomes are included: • disease activity • functional capacity • disease progression • periarticular disease (enthesitis and dactylitis) • mortality • adverse effects of treatment • health-related quality of life	Axial outcomes are not included in the submission as they were not recorded in the pivotal trial. This outcome has not been requested in any previous NICE appraisals for PsA ((3-5).	The ERG considers that all important outcomes have been assessed. The exclusion of axial outcomes is reasonable.
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. If the technology is likely to provide similar or greater health benefits at similar or lower cost than technologies recommended in published NICE technology appraisal guidance for the same indication, a cost-comparison may be carried out. 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 perspective.	The cost effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year (QALY). The time horizon for estimating clinical and cost effectiveness was set to 40 years to be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs are considered form an NHS and Personal Social Services perspective. The availability of commercial arrangements for the intervention or comparator technologies and	Same as the final scope issued by NICE.	

	The availability of any commercial arrangements for the intervention or comparator technologies and subsequent treatments will be taken into account. For the comparators the availability and cost of biosimilars should be taken into consideration.	subsequent treatments has been considered where details are publicly available and list prices used where confidential discounts are not available. The availability and cost of biosimilars has been taken into consideration for comparators.		
Subgroups	If evidence allows the following subgroups will be considered: • the reason for previous treatment failure (for example due to lack of efficacy, intolerance or adverse events). • presence or severity of concomitant psoriasis (no psoriasis, mild, moderate or severe psoriasis) • presence or severity of axial involvement Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	The following subgroups are included: • patients who have discontinued previous treatments (not separated by reason for discontinuation or type of previous treatment). • presence or severity of concomitant psoriasis (minimal psoriasis, mild to moderate, or moderate to severe, see Section B.3.2.1 in MS).	In line with the final scope issued by NICE. Reasons for discontinuation and presence of severity of axial involvement have not been considered as there was either no or insufficient data to analyse these subgroups. This is consistent with previous NICE PsA appraisals.	The psoriasis severity categories used to report subgroup results in the clinical effectiveness section of the CS did not match those used in the economic modelling. The ERG therefore requested subgroup results based on the categories used in the model. See Section 3.2.2.5.
Special considerations including issues related to equity or equality				The ERG notes no specific issues in this area.

The ERG sought more information from the company with respect to the recommended timepoint for discontinuing guselkumab due to lack of efficacy, given that the SmPC not yet available. The CS stated a timepoint for discontinuing of 16 weeks in Table 2, but also stated that it was anticipated to be 24 weeks on page 130. In response a clarification question, the company stated that the 16 weeks reference was misleading as it refers to the SmPC for the plaque psoriasis indication and that the SmPC for the psoriatic arthritis indication is anticipated to be 24 weeks. This seems to imply that guselkumab improves psoriasis symptoms more quickly than it improves joint symptoms. The company added that the 24-week timepoint was necessary since it may take guselkumab slightly longer to reach a full level of response compared to other therapies, because the impact on IL-23 needs to filter through to downstream inflammation modulators. The ERG considers this mechanistic explanation to be somewhat at odds with the aforementioned difference in stopping rules (16 weeks for plaque psoriasis versus 24 weeks for psoriatic arthritis). No explanation was provided as to why improvement in skin symptoms would occur before improvement in joint symptoms. The ERG therefore evaluated the evidence for differences in changes in outcome results data over time, which is presented in Section 3.2.2.4.

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

An assessment of the company's searches to identify studies on clinical effectiveness is presented in Table 2. Overall the literature searching was of a high standard and the details were well-reported.

Table 2 ERG assessment of the company's literature searches

Topic	ERG response	Note
Is the report of the search clear and comprehensive?	Partly	The search description lists the databases used and provides the full search strategies with search dates and numbers of records identified
Were appropriate sources searched?	Yes	The search used: bibliographic databases (MEDLINE, Embase, Cochrane CENTRAL) Trial Registers (e.g. ClinicalTrials.gov) Conference Proceedings identified from the sources above Reference lists from systematic reviews were scanned for relevant studies
Was the timespan of the searches appropriate?	Yes	The original search was conducted in October 2018 and the update search in January 2020 with no date restriction apart from a two year restriction for conference papers
Were appropriate parts of the PICOS included in the search strategies?	Yes	The search strategies appropriately combine terms for psoriatic arthritis; guselkumab (the technology in question) and comparators with study type terms for RCTs
Were appropriate search terms used?	Yes	In line with best practice, the search strategies combined thesaurus terms and free text terms
Were any search restrictions applied appropriate?	NA	
Were any search filters used validated and referenced?	Yes	The MEDLINE search strategy used the Cochrane Highly Sensitive Search Strategy to restrict the results to RCTs The Embase search strategy used the CADTH RCT search filter to restrict to records of RCTs Both search filters were fully referenced in the search description

3.1.2 Inclusion criteria

Full inclusion and exclusion criteria are presented in Table 1 of Appendix D of the CS. Many interventions were included which were not listed in NICE's final scope document. However, this was not an issue as these were subsequently excluded from the quantitative analyses (see Table 3 of Appendix D of the CS).

3.1.3 Data extraction

Data extraction methods were presented in Appendix D of the CS and are appropriate.

3.1.4 Quality assessment

Studies identified for inclusion in the systematic review of effectiveness were quality assessed with the results reported in Appendix D (Table 10) of the CS. Although appropriate aspects of internal validity were assessed, no further details were provided to justify the individual "Yes", "No" or "Unclear" decisions. No external validity assessments were made to evaluate trial applicability to the NHS setting.

3.1.5 Evidence synthesis

Results from the guselkumab studies were not pooled in pairwise meta-analyses. The synthesis presented in the CS was a network meta-analysis (see Section 3.4).

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation

The efficacy data in the CS were based on results from three double-blind, 24-week, placebo-controlled RCTs of guselkumab: one was a phase II trial (n=149)[1] and two were phase III trials called DISCOVER-1 (n=382) and DISCOVER-2 (n=741).[2]

3.2.1 Design and methods of the guselkumab trials

The DISCOVER trials had both 4-weekly and 8-weekly dosing arms for guselkumab, and the Deodhar trial used 8-weekly dosing. Given that the recommended dose is currently 8-weekly the ERG focussed on results for the 8-weekly dose arms. [redacted] Results were presented in the CS for the 24-week and 52-week timepoints. Another trial, named COSMOS, is ongoing and has not yet reported week 24 results. COSMOS (NCT03796858) is evaluating the efficacy and safety of guselkumab 100mg q8w in patients who have had an inadequate response to a TNFα-inhibitor.

Eligibility criteria were reported in Table 11 of the CS. These were broadly similar to the criteria used in other trials of systemic treatments for psoriatic arthritis. The trial appeared to have robust internal validity (see Table 14 of the CS). However, the ERG sought more detail from the company about the 'early escape' process at week 16. The company stated that at week 16, patients with less than 5%

improvement in both swollen and tender joint counts were eligible for early escape, in which the investigator could initiate or increase the dose of NSAIDs or other analgesics, oral corticosteroids, or certain non-biologic DMARDs. The decision to increase or initiate concomitant medicine was at the discretion of the investigator, and there was no protocol requirement that investigators discuss or refrain from discussing early escape status with participants. Furthermore, the ERG was concerned about how early-escape patient data were analysed at 24 weeks in both the DISCOVER trial analyses as this may have biased some effect estimates. These issues are discussed further in Section 3.2.2.3.

Modified rather than full intention-to-treat datasets were used for the analyses though the differences between the denominators for the two datasets were very small. Non-responder imputations were used for missing binary data and multiple imputation methods were used for continuous outcomes.

3.2.2 Results of the guselkumab trials

Figures 11 (DISCOVER 1) and 12 (DISCOVER 2) in the CS are CONSORT diagrams of patient flow and indicated that many patients were 'screen failures' i.e. screened but not randomised. The ERG requested more data from the company about why patients were excluded at the screening phase. These revealed that the most common reason for ineligibility was having a CRP level of <3mg/l at the screening visit.

3.2.2.1 Baseline characteristics

Baseline characteristics of patients in the two DISCOVER trials were presented in Table 12 of the CS, adapted in this report and presented in Table 3. In terms of the applicability of the trial populations to NHS settings, the ERG's clinical adviser thought it was unlikely that guselkumab would be used as a first line biologic therapy for most NHS patients, given the availability of established biologics with well-known safety profiles. Therefore, in clinical practice a high proportion of patients will have received another biologic before commencing guselkumab. The ERG therefore considers it likely that the proportion of patients in the pivotal guselkumab trials who have previously been treated with a biologic therapy (31% in DISCOVER 1, 0% in DISCOVER 2) is not reflective of clinical practice in the NHS.

A similar issue is seen for prior cDMARDs. NICE recommends that bDMARDs are given after two cDMARDs have been tried. Limited data on prior cDMARDs were presented in Table 12 of the CS so the ERG requested further data (included in Table 3). This showed that in the guselkumab trials less than a third of patients had received two or more prior cDMARDs. Moreover, just under 10% of patients had received no prior cDMARD. It is unclear how many of these patients had contraindications to cDMARDs, though the ERG's clinical adviser estimated that in the NHS contraindication to cDMARDs would be seen in <0.5% of patients. This further limits the applicability of the guselkumab trials to the NHS setting.

Another applicability concern was baseline PASI scores. The ERG's clinical advisor believes that the average PASI score in the DISCOVER trials were high compared to patients seen in NHS settings. No UK sites were used in any of the guselkumab trials. The ERG's clinical adviser stated that in his large clinic he rarely sees psoriatic arthritis patients with PASI scores above 5. Although both the DISCOVER trials were multi-site international trials, the ERG noted that a very large proportion of patients were recruited from Poland, Russia and Ukraine. The proportion of patients recruited from any Eastern European or Asian country was 82% in DISCOVER 1, 97% in DISCOVER 2, and 80% in the phase II trial. The ERG's clinical adviser said that Eastern European patients do not have the same primary care interface as NHS patients and were therefore less likely to have used two synthetic DMARDs before entering a trial. The more limited access to primary care in Eastern European countries may result in fewer alternative treatment options for patients who do not respond well to guselkumab. This could be an alternative explanation for low discontinuation rates associated with guselkumab (see Section 3.3.1); on p175 of the CS the company stated that "*the low rate of discontinuations for guselkumab observed in the DISCOVER trials are based on low levels of discontinuations of the drug for efficacy and tolerability reasons, with few patients developing antibodies to guselkumab as well*". Finally, many patients were ineligible for the trials for having C-reactive protein levels below a certain cut-off (<3mg/l or <6mg/l) at screening visits, a criterion which would not be used in the NHS. Taken together, these factors suggest that the guselkumab trial results have poor applicability to the NHS setting.

Table 3 Baseline Characteristics of participants across treatment arms in DISCOVER-1 and DISCOVER-2 (adapted from Table 12 of the CS and response to ERG clarification question A14)

	DISCOVER-1			DISCOVER-2		
Characteristics	**PBO (N=126)**	**GUS 100mg q8w (N=127)**	**GUS 100mg q4w (N=128)**	**PBO (N=246)**	**GUS 100mg q8w (N=248)**	**GUS 100mg q4w (N=245)**
Age (years), mean (SD)						
Male, n (%)						
Race (white), %						
Weight (kg), mean (SD)						
BMI (kg/m^2), mean (SD)						
Duration of PsA (years), mean (SD)						
PsA duration ≥3 years, n (%)						
Age of PsA diagnosis (years), mean (SD)						
Tender/ painful joints, mean (SD)						
Swollen joints, mean (SD)						
Participants with PsA subtypes, n (%)						

	DISCOVER-1			DISCOVER-2		
Characteristics	**PBO (N=126)**	**GUS 100mg q8w (N=127)**	**GUS 100mg q4w (N=128)**	**PBO (N=246)**	**GUS 100mg q8w (N=248)**	**GUS 100mg q4w (N=245)**
Polyarticular arthritis with absence of rheumatoid nodules						
Spondylitis with peripheral arthritis						
Asymmetric peripheral arthritis						
DIP joint involvement						
Arthritis mutilans						
HAQ-DI, mean (SD)						
CRP (mg/dL), median (IQR)						
DAS28 (CRP) score, mean (SD)						
DAPSA score, mean (SD)						
Participants with enthesis using LEI score, n (%)						
Participants with dactylitis, n (%)						
PASI score, mean (SD)						
Prior TNFi agent, n (%)						
1 therapy						
2 therapies						
Prior DMARDs, n (%)						
No prior cDMARD, n (%)						
1 prior cDMARD, n (%)						
2 prior cDMARDs, n (%)						
≥3 prior cDMARDs, n (%)						
Participants receiving cDMARDs, n (%)						
Participants receiving oral corticosteroids, n (%)						
Participants receiving NSAID, n (%)						

3.2.2.2 Main efficacy results of the guselkumab trials

Clinical efficacy results data for guselkumab trials were presented in section B.2.6 of the CS which reported results for DISCOVER 1 and DISCOVER 2, but not for the phase II trial. Some of the phase II trial results were reported in Appendix P of the CS; the ERG also requested the phase II clinical study report, which was accidentally omitted from the original submission. The phase II trial results appear similar to those reported for the DISCOVER trials.

The DISCOVER trial results showed that treatment with guselkumab resulted in statistically significant and clinically relevant improvements at 24 weeks when compared with placebo across a range of outcomes including ACR20, ACR 50, HAQ-DI, PASI 75, PASI 90, PASI 100, PsARC, and MDA. There was some variation across the DISCOVER trials in results for ACR 70, radiographic progression of disease, enthesitis and dactylitis. For ACR 70 statistically significant results were seen at 24 weeks in DISCOVER-2, but not in DISCOVER-1. Radiographic progression of disease was assessed only in DISCOVER-2, using the modified van der Heijde/Sharp score. At 24 weeks there was a statistically significant difference between the guselkumab q4w group and placebo (p=0.006) but not between the guselkumab q8w group and placebo (p=0.07, see Table 23 of the CS). The guselkumab q8w versus placebo results for resolution of enthesitis and resolution of dactylitis were not statistically significant in DISCOVER 1, but were in DISCOVER-2 ($p<0.05$, Table 29 of the CS).

Patients entering early-escape at 16 weeks

Patients entered early-escape based on lack of response at week 16, although investigators were not obliged to tell patients that they had qualified for early escape (see Section 3.2.1). In DISCOVER 1, 30 patients qualified for early escape but only 14 actually entered early escape; the corresponding figures for DISCOVER 2 were 63 and 27. As would be expected, most patients who qualified for early escape were taking placebo. The ERG had concerns about whether there may be systematic differences between patients who initiated early escape (i.e. received a new co-intervention, or change in dose of an existing co-intervention) and patients who qualified for early escape but were not told. For example, patients who had been experiencing a flare at week 16 may have been more likely to have been informed about their early escape status (and so initiate or change a co-intervention). This issue is important because the main strategy for analysing results data in the CS, including the guselkumab data used in the NMAs, involved assuming that patients who initiated early escape at week 16 were non-responders at week 24, even if they did actually respond at week 24. Observed data were used for patients who qualified for early escape, but who were not informed of this (and so could not initiate early escape). Week 24 results using observed data for early escapers were not presented in the CS but were available in the clinical study reports. In light of these issues the ERG sought to compare results using these different approaches.

3.2.2.3 Comparison of key results from the DISCOVER trials using different analysis strategies

Appendix M of the CS describes two different strategies for analysing the DISCOVER trial results with respect to the handling of data from patients classed as 'treatment failures'. The composite strategy - termed the 'composite estimand' - was the main strategy used for analysing all efficacy endpoints up to week 24. Using this approach, if a patient met any of the treatment failure criteria, they were considered a non-responder for response variables and had a score of no improvement (i.e.

no change from baseline) for continuous variables from the time the patient met any treatment failure criteria. The treatment failure criteria were:

- Discontinued study agent injections for any reason
- Terminated study participation for any reason
- Initiated or increased the dose of non-biologic DMARDs or oral corticosteroids over baseline for PsA (includes 'early escapers')
- Initiated protocol-prohibited medications/therapies for PsA

The 'treatment policy estimand' was a supplementary estimand which used all the observed data collected for an outcome, regardless of whether the patient had met any treatment failure criterion. Given that, in DISCOVER 1 in particular, there was a notable group imbalance in treatment failures due to early escape (10.3% for placebo versus 0.8% guselkumab q8w) the use of the composite estimand (when compared to using the treatment policy estimand) may result in biased effect estimates at 24 weeks, as the placebo group results may be diminished disproportionately more than the guselkumab results. There was only a small imbalance in the proportion of early escapers in DISCOVER 2 (5.7% placebo versus 2.4% guselkumab q8w).

In order to assess whether the use of the composite estimand had affected the internal validity of the DISCOVER trials the ERG extracted results data from the two DISCOVER CSRs to compare results based on the composite estimand with those based on the treatment policy estimand for both the 16- and 24-week timepoints (Table 4). Any bias in results was anticipated to be most likely seen in outcomes where achievement of a placebo response was not uncommon. The grey cells in Table 4 highlight the different placebo responses between the two analysis strategies at 24 weeks. The corresponding differences for the guselkumab q8w group (below the grey cells) are smaller, or do not exist (in the case of PASI 75). The resulting changes in the relative risks (or difference in means) calculated by the ERG could not be considered as being negligible in DISCOVER 1. The ERG considers that the use of analyses which impute data when the data are not missing is inappropriate as a main strategy and therefore has concerns about the validity of the 24-week results for DISCOVER 1 using the composite estimand.

Table 4 Comparison of guselkumab trial results using composite and treatment policy estimands (guselkumab q8w vs placebo) in outcomes where a placebo response is not uncommon

Outcome	Trial arm	Analysis strategy and results							
		DISCOVER 1				DISCOVER 2			
		16 weeks: composite estimand	16 weeks: treatment policy estimand	24 weeks: composite estimand	24 weeks: treatment policy estimand	16 weeks: composite estimand	16 weeks: treatment policy estimand	24 weeks: composite estimand	24 weeks: treatment policy estimand
PsARC	PLA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Q8w	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Relative risk (95% CI)		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI 75	PLA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Q8w	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Relative risk (95% CI)		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ACR 20	PLA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Q8w	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Relative risk (95% CI)		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ-DI mean change from baseline	PLA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Q8w	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in mean change in HAQ, Q8w v PLA		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

3.2.2.4 Analysis of 16, 20 and 24-week results

Given that the recommended stopping rule when guselkumab is used to treat plaque psoriasis is 16 weeks, and that a 24-week stopping rule timepoint was deemed necessary for psoriatic arthritis (see Section 2.3) the ERG critiqued the submission for evidence to support the difference in stopping timepoints. The ERG noted several references in the CS to guselkumab providing "rapid" control of symptoms, often quite some time before week 16. This was reported for joint symptoms (p64-65, p117), skin symptoms (p62, p118) and physical functioning (p66), with the latter being "*consistent with guselkumab being able to address the unmet need of patients for rapid improvement in symptoms*". The ERG finds this assessment by the company to be inconsistent with using a 24-week stopping rule because it "*may take guselkumab slightly longer to reach a full level of response compared to other therapies, because the impact on IL-23 needs to filter through to downstream inflammation modulators*."

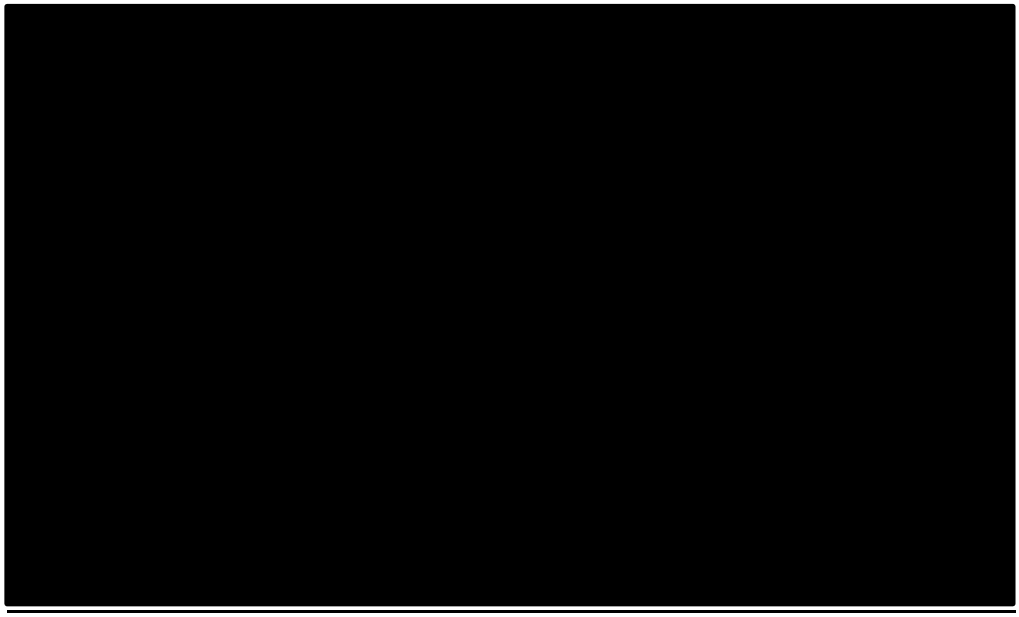
The ERG also examined whether trial data supported the company's assertion that guselkumab may take slightly longer to reach a full level of response, by looking at the more reliable treatment policy estimand results. Figure 1and Figure 2 show response rates over time for PsARC and ACR 50 in the two DISCOVER trials. The continued improvement in placebo responses up to week 20 imply that the improvement in guselkumab responses from week 16 to week 20 is not specific to guselkumab. Moreover, the lack of improvement in response rates between weeks 20 and 24 for both guselkumab q8w and placebo do not suggest that a 24-week stopping rule is required.

Figure 1 Proportion of patients achieving a PsARC response up to week 24 in the DISCOVER 1 and DISCOVER 2 trials using treatment policy estimand data

Figure 2 Proportion of patients achieving an ACR 50 response up to week 24 in the DISCOVER 1 and DISCOVER 2 trials using treatment policy estimand data

For PASI 75, only 16 and 24-week results data were available in both the DISCOVER CSRs (presented here in Table 4). With respect to changes in response rates over time, the results from these two trials differed: in DISCOVER 2 there was an increase in response rates between 16 and 24 weeks for both guselkumab and placebo, whereas in DISCOVER 1 the placebo response rate decreased by 3% and the guselkumab response rate increased by 13%. Given both this conflicting evidence about how PASI 75 rates may change over time, and the company's 'mechanism of action' explanation for longer times needed for guselkumab response (compared to other therapies), consideration was made of the evidence from the guselkumab trials in patients with plaque psoriasis, where PASI 75 was also an outcome. Guselkumab was compared to adalimumab in the VOYAGE 1 and VOYAGE 2 trials in patients with moderate-to-severe psoriasis. Figure 3c of the Blauvelt et al 2016 paper on VOYAGE 1 (included in the company's reference pack) suggests that peak PASI 75 response rates are reached at week 16, with no tangible improvements in rates thereafter.[3] Figure 3 of the Reich et al 2017 paper on VOYAGE 2 (also included in the company's reference pack) shows a slight rise in response rates between weeks 16 and 20 for both guselkumab and adalimumab and a flattening (guselkumab) or slight decrease (adalimumab) at week 24.[4]

Figure 3 HAQ-DI change from baseline up to week 24 in the DISCOVER 1 and DISCOVER 2 trials using treatment policy estimand data



Looking at HAQ-DI changes between weeks 16 and 24, for placebo, HAQ-DI change reaches its peak at week 20. For guselkumab (q8w) the peak change in HAQ-DI is also seen at week 20 in DISCOVER 1, although HAQ-DI continues to fall at week 24 in DISCOVER 2.

In summary, the trial data analysed using the treatment policy estimand suggest that most of the patients who are going to respond to guselkumab, respond by week 16; responses after week 16 may be seen, but they do not appear to be specific to guselkumab. Therefore, there is little evidence of a specific time-to-response effect to suggest that guselkumab warrants a 24-week stopping rule.

3.2.2.5 Subgroup analyses

Results for subgroups were reported in section B.2.7 of the CS (beginning on p72) and in Appendix E. The CS stated that data could not support a subgroup analysis of the reason for prior treatment failure or the presence or severity of axial involvement, which were listed as subgroups of interest in the NICE scope.

Psoriasis severity

The results presented in the CS for the psoriasis severity subgroups related to ACR 20, using severity thresholds which were different to those used in the economic model. In a clarification question the ERG requested subgroup results using the psoriasis severity subgroup categories used in the cost-effectiveness model for the following outcomes: ACR 20, ACR 50, ACR 70, PASI 75, PASI 90, PsARC, HAQ-DI, MDA, enthesitis, and dactylitis.

The results provided by the company, presented in Table 3, are based on the following subgroup definitions:

- Minimal psoriasis, BSA <3% (no PASI cut-off)
- Mild-to-moderate psoriasis, BSA ≥3% and PASI ≤10
- Moderate-to-severe psoriasis, BSA ≥3% and PASI >10

Interpretation of these results is somewhat hindered by the lack of denominators in the Excel data file provided by the company. It was evident though that in DISCOVER 2 around 20% of patients had minimal psoriasis, around 50% had mild-moderate psoriasis and around 30% had moderate-severe psoriasis. Although there is a suggestion in these results that efficacy might increase as psoriasis severity increases, the number of events is sometimes small in the minimal and moderate-to-severe subgroups and, as the company noted, this trend could be because these patients have greater room for experiencing an improvement (at least for some outcomes).

Prior TNFα inhibitor exposure

The CS also reported subgroup results based on prior TNFα inhibitor exposure (Table 31, p74, CS). The ERG requested that further outcomes were added, which are presented here in Table 6. Comparison of results between these subgroups is difficult given that only 80 patients were included in the prior TNFα inhibitor group. Effect estimates from this subgroup therefore tend to have very wide confidence intervals.

Table 5 Trial results by psoriasis severity subgroup (adapted from Table 3 of the company's response to clarification questions)

Outcome	Subgroup	Relative risks (95% CI) for DISCOVER 1		Relative risks (95% CI) for DISCOVER 2	
		GUS q8w vs PBO Week 16	GUS q8w vs PBO Week 24	GUS q8w vs PBO Week 16	GUS q8w vs PBO Week 24
ACR50	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]
PASI75	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]
PSARC	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]
MDA	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]
Resolution of enthesitis	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]

Outcome	Subgroup	Relative risks (95% CI) for DISCOVER 1		Relative risks (95% CI) for DISCOVER 2	
		GUS q8w vs PBO Week 16	GUS q8w vs PBO Week 24	GUS q8w vs PBO Week 16	GUS q8w vs PBO Week 24
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]
Resolution of dactylitis	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]
HAQ-DI (LS Mean difference)	All patients	[redacted]	[redacted]	[redacted]	[redacted]
	Minimal PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mild-Mod PsO	[redacted]	[redacted]	[redacted]	[redacted]
	Mod-Sev PsO	[redacted]	[redacted]	[redacted]	[redacted]

LS least squares, Mod moderate, PsO psoriasis, Sev severe

Table 6 Subgroup results by prior biologic use at Week 24 in DISCOVER-1 (adapted from Table 4 of the company's clarification response)

Outcome	Treatment arm	Results for prior TNFi subgroup (PBO=39, q8w=41)	Results for no Prior TNFi subgroup (PBO=87, q8w=86)
ACR 20	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
ACR 50	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
ACR 70	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
PASI 75	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
PASI 90	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
PASI 100	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
PASI 50	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
PsARC	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
HAQ	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
MDA	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]
	Placebo	[redacted]	[redacted]

Outcome	Treatment arm	Results for prior TNFi subgroup (PBO=39, q8w=41)	Results for no Prior TNFi subgroup (PBO=87, q8w=86)
Resolution of enthesitis	q8w	[redacted]	[redacted]
Resolution of Dactylitis	Placebo	[redacted]	[redacted]
	q8w	[redacted]	[redacted]

RR relative risk, q8w every 8 weeks

3.2.2.6 Longer-term clinical effectiveness

In response to an ERG clarification question, the company stated that there were no stopping rules due to lack of efficacy in the 52-week DISCOVER studies. When looking at key outcomes, such as PsARC, these studies reported results based on observed data, using denominators based on the number of patients evaluable at the timepoint in question, rather than the number randomised (i.e. week 0). Lack of stopping rules and use of observed data are methods which are often used in longer-term studies of trial cohorts. Other methodological issues often seen in longer-term studies – but which may be difficult to avoid - are lack of appropriate control groups and lack of, or limited, blinding. Nevertheless, taken together, the use of these methods mean that results of long-term trials are often both biased and have poor applicability to clinical practice. This is the case for the 52 week DISCOVER studies. For example, in the CSR for DISCOVER-2 a PsARC response rate of [redacted] is reported at 52 weeks for the guselkumab q8w arm. Using the number randomised as the denominator the result is [redacted]. In the NHS this figure would be considerably lower since 26% of patients did not achieve a PsARC response at week 24 and so are likely to have had their treatment discontinued.

Table 25 of the CS reported the 52-week discontinuation rates for guselkumab q8w as being 7.5% (DISCOVER-1) and 5.6% (DISCOVER-2). The ERG notes considerable variation in the placebo discontinuation rates at week 24, being 20.6% in DISCOVER-1 and 5.2% in DISCOVER-2. Page 175 of the CS stated that the low rate of discontinuations for guselkumab are based on efficacy and tolerability reasons, with few patients developing antibodies to guselkumab as well, with a possible biological rationale for this being that guselkumab may regulate one of the key cytokines of the IL-23-Th17 pathway. Although it is possible that this may be a reason for low discontinuation rates, the ERG has concerns that the company has not considered other reasons, such as the very high prevalence of Eastern European trial sites in the DISCOVER trials (discussed in section 3.2.2). A further exploration of variation in discontinuation rates across trials included in the network meta-analyses is presented in Section 3.3.1.

3.2.2.7 Adverse events

Adverse events data for the two DISCOVER trials are presented in the CS for the 24-week timepoint (see Table 53 of the CS) and for the extension periods: week 60 for DISCOVER 1 and week 52 for DISCOVER 2.

As for the efficacy comparisons, the only direct comparisons were between guselkumab and placebo. At 24 weeks, both doses of guselkumab were generally well-tolerated and there were no obvious differences in adverse event rates between placebo and guselkumab for key outcomes such as serious adverse events, infections, serious infections and injection site reactions. The longer-term data showed small increases in event rates, compared to the 24 week data: in the guselkumab q8w arm the

proportion of patients who had experienced at least one serious adverse event was [redacted] at week 60 in DISCOVER-1 and [redacted] at week 52 in DISCOVER -2.

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

The methods used in the company's systematic review (reported in Appendix D of the CS) were largely appropriate and have been summarised in Section 3.1. A PRISMA flow diagram is presented in Appendix D of the CS, along with tables of the included and excluded studies and the quality assessment results. The included RCTs were judged to generally be at a low risk of bias, which the ERG concurs with although the ERG notes the lack of an appraisal of trial applicability to the NHS setting.

The company included ustekinumab as a comparator in its NMAs of biologic-naïve patients even though ustekinumab is only recommended by NICE after a patient has had treatment with one or more TNFα inhibitors. Also, in a point of clarification, the ERG questioned the inclusion of:

- The RAPID-PsA trial of certolizumab in the biologic-experienced NMAs as this trial excluded primary non-responders to a previous TNFα inhibitors so the similarity assumption appears to not have been met for inclusion in the network
- The PSUMMIT ustekinumab 90mg arms as this dose was not administered as per the marketing authorisation (i.e. all patients should weigh >100kg but they did not)

The ERG asked to company to either remove these data and re-submit new network meta-analyses or to provide justification for still including them. The company opted to justify the inclusion of these trials based on the precedent set by NICE appraisal TA537 and on the assumption that their continued inclusion would not lead to decision-altering changes in the remaining estimates in the NMA. The ERG does not agree with the precedent argument since these data were excluded from the earlier TA445 appraisal, but concurs that their inclusion is unlikely to lead to decision-altering changes in the NMA results.

3.3.1 Discontinuation rates across trials included in the NMA

Discontinuation rates are an important model parameter which show heterogeneity across trials and treatments. One-year discontinuation rates were presented in Tables 3 and 4 in Appendix W of the CS. Guselkumab had the lowest rate, at 6.9%, with rates for the remaining treatments ranging from 10.3% (tofacitinib) to 26.5% (apremilast).

Reasons for discontinuation rates varying across trials may be treatment-specific, such as differences in the proportion of patients showing lack of efficacy and adverse events, or they may not be specifically related to a treatment. Some examples of the latter may include:

- Variation in healthcare settings (e.g. level of care available outside of the trial)
- Differences in trial protocols, (e.g. efforts to follow-up and retain patients within the trial)
- Variation in staff involvement, attitude, and understanding of the study protocol[5]
- Chance effects – discontinuation rate estimates are prone to variation due to chance events because small numbers of discontinuation events are often used to calculate rates

The possible impact of these non-specific factors can perhaps best be examined by exploring variation across trials in placebo discontinuation rates, since these would be expected to be very similar with respect to adverse events, though they might be expected to fall slightly over time with respect to lack of efficacy, in line with increased placebo response rates. Table 7 presents 16-week discontinuation rates for placebo arms across trials included in the NMAs; the 16-week timepoint was chosen to maximise the amount data available. The discontinuation rates range from almost zero in DISCOVER 2, to 17.3% in PSUMMIT 2, illustrating the uncertainties involved when using trial-based estimates for discontinuation rates. A chance effect seems to drive the high placebo discontinuation rate in PSUMMIT 2, with 7 of the 18 discontinuations being due to adverse events - by comparison, no discontinuations due to adverse events were seen in the corresponding ustekinumab 45mg arm in PSUMMIT 2.[6]

Table 7 Rates of placebo discontinuation at 16 weeks in the NMA trials

Therapy and Trial	Number discontinued	Discontinuation rate (%)
INF, IMPACT	2/52	3.8
GOL, GO-REVEAL	10/113	8.8
UST, PSUMMIT 1	7/206	3.4
UST, PSUMMIT 2	18/104	17.3
CZP, RAPID-PsA	15/136	11.0
SEC, FUTURE 2	10/98	10.2
SEC, FUTURE 3	8/137	5.8
SEC, FUTURE 4	7/114	6.1
GUS, DISCOVER 1*	[redacted]	[redacted]
GUS, DISCOVER 2*	[redacted]	[redacted]
GUS, Deodhar 2018*	[redacted]	[redacted]

* Data were extracted from clinical study reports. 16-week data unavailable for all PALACE trials, FUTURE 5, OPAL-BROADEN, SPIRIT-P1, IMPACT-2 and Nash 2018. CZP certolizumab, GUS guselkumab, INF infliximab, SEC secukinumab, UST ustekinumab

Looking in more detail at the DISCOVER trials, an unexpectedly low rate of placebo discontinuations due to lack of efficacy by week 24 was seen, with 0/246 (DISCOVER 2) and 4/126 (DISCOVER 1) placebo patients discontinuing. In other words, after nearly six-months of placebo treatment, only 1.1% of patients discontinued due to lack of efficacy. Given that lack of efficacy is inherently associated with placebo, it seems plausible that the excellent patient retention seen in the DISCOVER trials is being affected by other factors, notwithstanding the company's assertion that low rates of discontinuation due to lack of efficacy of guselkumab may be linked to its mechanism of action.

As mentioned earlier in section 3.2.2, the issue of trial setting was suggested by the ERG's clinical adviser as one possible alternative explanation for lower withdrawal rates in the DISCOVER trials - limited access to primary care in Eastern European countries may result in fewer alternative treatment options for patients, which may mean better retention of patients within trials. The ERG therefore extracted data on geographical settings for studies included in the NMAs to see how this varied across trials (Table 8). These data indicate that in the guselkumab trials a much higher proportion of patients was recruited from eastern European sites when compared to other trials. The data also show that the early trials of TNF-inhibitors were undertaken primarily, or totally, in U.S. or western European settings. The data in Table 8 relate to the numbers of trial sites by location. Few studies reported the actual number of patients randomised per site; where these data were reported they indicated that a very high proportion of patients in the guselkumab trials were recruited in sites located in eastern Europe or in Asia (see Table 9); in DISCOVER-2 the figure is 97%.

Table 8 Geographic distribution of recruiting sites in trials included in the NMAs

Therapy; Trial name	Number (%) of recruiting sites					
	North America	South America	Western Europe	Eastern Europe	Asia	Other
INF; IMPACT	27 (75%)	0	9 (25%)		0	0
INF; IMPACT 2	27 (75%)	0	9 (25%)		0	0
ADA; Genovese 2007	16 (100%)	0	0	0	0	0
ETN; Mease 2000	NR	NR	NR	NR	NR	NR
ETN; Mease 2004	17 (100%)	0	0	0	0	0
CZP; RAPID-PSA	32 (35%)	14 (15%)	30 (33%)	16 (17%)	0	0
TOF; OPAL-BEYOND	43 (34%)	24 (19%)	32 (26%)	18 (14%)	5 (4%)	3 (2%)
TOF, ADA; OPAL-BROADEN	27 (22%)	13 (11%)	26 (21%)	47 (38%)	4 (3%)	6 (5%)
APR; PALACE 1	43 (47%)	0	22 (24%)	11 (12%)	0	16 (17%)
APR; PALACE 2	31 (32%)	0	30 (31%)	29 (30%)	5 (5%)	3 (3%)
APR; PALACE 3	30 (23%)	0	63 (48%)	23 (18%)	7 (5%)	8 (6%)
UST; PSUMMIT1	45 (44%)	0	32 (31%)	15 (15%)	0	10 (10%)

Therapy; Trial name	Number (%) of recruiting sites					
	North America	South America	Western Europe	Eastern Europe	Asia	Other
UST; PSUMMIT 2	42 (51%)	0	33 (40%)	7 (9%)	0	0
IXE AND ADA; SPIRIT-P1	41 (35%)	5 (4%)	29 (25%)	36 (31%)	5 (4%)	0
IXE; SPIRIT-P2	59 (54%)	0	32 (29%)	8 (7%)	7 (6%)	3 (3%)
IXE; SPIRIT H2H	5 (4%)	19 (15%)	59 (45%)	12 (9%)	13 (10%)	22 (17%)
SEC; FUTURE 2	28 (36%)	0	24 (31%)	20 (26%)	1 (1%)	4 (5%)
SEC; FUTURE 3	21 (26%)	23 (28%)	22 (27%)	7 (9%)	0	9 (11%)
SEC; FUTURE 4	23 (36%)	0	19 (30%)	18 (28%)	0	4 (6%)
SEC; FUTURE 5	28 (16%)	47 (27%)	40 (23%)	34 (20%)	17 (10%)	6 (4%)
GUS; Deodohar 2018	10 (29%)	0	7 (21%)	17 (50%)	0	0
GUS; DISCOVER-1	13 (15%)	0	17 (20%)	36 (42%)	14 (16%)	6 (7%)
GUS; DISCOVER-2	2 (2%)	0	16 (14%)	86 (73%)	6 (5%)	8 (7%)

APR apremilast, ETA etanercept, GUS guselkumab, INF infliximab, IXE Ixekizumab, SEC secukinumab, TOF tofacitinib, UST ustekinumab

Table 9 Geographic distribution of the number of patients randomised (where reported)

Therapy and trial	Number and % of participants randomised, by continent and country					
	North America	South America	Western Europe	Eastern Europe	Asia	Other
IXE, ADA; SPIRIT-P1	n=87 (21%)	n=12 (3%)	n=55 (13.2%)	n= 250 (60%)	n=12 (3%)	n=0
GUS; DISCOVER-1	USA and Spain n=67 (18%)			n=314 (82%): Poland n=107 (28%) Russia n=64 (17%) Ukraine n=70 (18%) Asia, Eastern European countries other than Poland, Russia, and Ukraine n=73 (19%)		n=0
GUS; DISCOVER-2	USA and Spain n=25 (3%)			n=714 (97%): Poland n=85 (12%) Russia n=273 (37%) Ukraine n=221 (30%) Asia, Eastern European countries other than Poland, Russia, and Ukraine n=135 (18%)		n=0
GUS; Deodhar 2018	n=17 (11%)	n=0	n=13 (9%)	n=119 (80%)	n=0	n=0
SEC; FUTURE-3	n=NR (14%)	n=0	n=NR (47%)	n=NR (31%)	n=0	n=NR (8%)

NR Not reported, SEC secukinumab, GUS guselkumab, IXE Ixekizumab, ADA adalimumab

In summary, comparisons of treatment discontinuation rate estimates which are based on data from longer-term extension studies of clinical trials are subject to considerable uncertainty. This is because discontinuation rates may be affected by factors which are not treatment-specific and which can vary across trials. These factors include healthcare settings, trial protocols, levels of staff involvement and attitude, and the play of chance. Although at least some of these factors appear to vary across the trials included in this appraisal, the company did not consider them as a cause of uncertainty when calculating their discontinuation rate estimates. Moreover, the absence of stopping rules for lack of efficacy in all but one of the trial extension studies which were used to estimate discontinuation rates for use in the economic model, means their results have limited relevance to clinical practice. For example, in the NHS most patients without a PsARC response after the relevant period of treatment (usually 12, 16 or 24 weeks) would discontinue treatment. Therefore, as an alternative to using only trial extension study data, discontinuation rate estimates might also be informed by using trial data from these earlier timepoints, which are important when making decisions about treatment discontinuation in clinical practice.

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

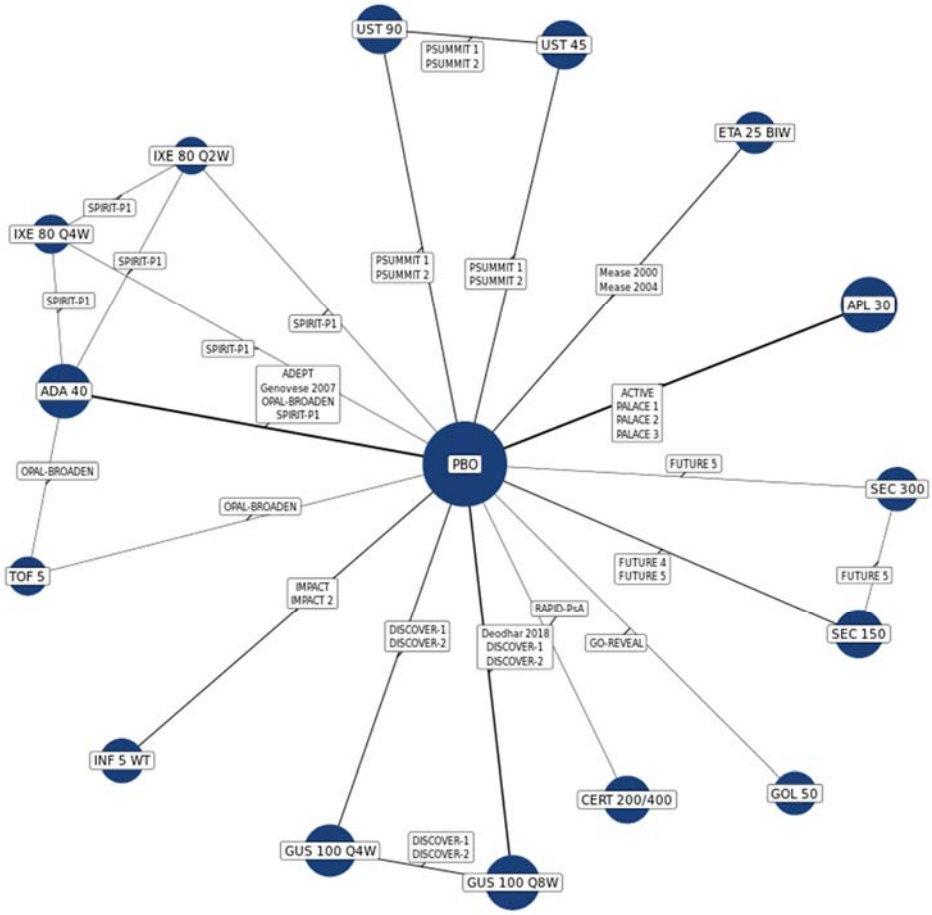
The CS reported a number of network meta-analyses (NMAs) intended to compare guselkumab to other relevant treatments, primarily TNFα inhibitors and interleukin inhibitors. NMAs were provided for all main outcomes. NMAs were provided for patients who were biologic-naïve (subpopulation 2) and biologic-experienced (subpopulation 3). TNFα inhibitor contraindicated patients (subpopulation 4) were handled by removing TNFα inhibitors from the biologic-naive analyses.

3.4.1 General critique of the network meta-analysis approach

While a network meta-analysis was performed, all included trials were primarily comparisons with placebo, with very limited head-to-head comparisons of active treatments (limited to some three-arm trials). Also, most treatments were examined either in a single trial, or a set of closely related trials from the drug manufacturer. Figure 4 reproduces the network diagram for ACR50 in biologic-naive patients (from Figure 19 in the CS) which illustrates this issue.

These limitations mean that all NMAs have little extra data above what would be obtained from indirect comparison of treatments via placebo using the Bucher method, as there are few loops in the network, nor enough trials to estimate or investigate heterogeneity. The ERG notes that the NMAs may therefore suffer from any bias common to such indirect comparisons. Comparisons will be biased if there are differences between trials in patient populations, how treatments were given, or outcomes analysed, that were not accounted for in the model. All NMA results should therefore be treated with caution, as would be applied to naïve indirect comparisons; results should be taken only to indicate how guselkumab might broadly compare with other treatments, and exact results and rankings should not be considered reliable.

Figure 4 NMA network diagram for ACR (biologic-naïve, CS Figure 19)



The CS identified evidence that the placebo response has improved over time (see CS Figure 18) and that trials with lower placebo response had generally higher estimates of effect (see CS Appendix D Figure 3 for the example for ACR 50). This would suggest that older trials might be overestimating treatment effectiveness, and so unfairly bias models against guselkumab. To address this, the CS reported NMAs with placebo-response adjustment, provided these models converged and there was evidence of a placebo-response effect (in line with NICE Decision Support Unit guidance). The ERG accepts that adjusting for placebo response is reasonable and appropriate. However, the ERG notes that the modelling assumption that older trials, if performed today, would therefore have a smaller treatment effect estimate, while plausible, is not proven. It is possible that the placebo-response association with effect estimates may have other causes. The limited size of the data set meant that NMA models also had to assume that the association between placebo response and treatment effect was the same for all treatments, which is a strong assumption, and unlikely to be true in reality.

To assess the validity and consequences of adjusting for placebo response, the ERG requested unadjusted results be provided for comparison. The ERG notes that making this request does not imply that we disagree with the principle of adjusting for placebo response.

The ERG's main concern with the NMAs reported in the CS is that they combined outcomes measured at different times. For the DISCOVER trials of guselkumab outcomes assessed at 24 weeks were used in all NMAs. Other trials and treatments had outcomes assessed at 16 weeks or earlier (in line with the authorised treatment durations for those treatments). Table 10 summarises the timepoints at which outcomes were assessed for each treatment in the NMAs.

Table 10 Timings of effect assessment by treatment in the NMAs

Treatment	Week of assessment			
	12	14	16	24
Adalimumab	X			
Apremilast			X	
Certolizumab pegol	X			
Etanercept	X			X
Infliximab		X	X	
Golimumab		X		
Guselkumab				X
Ixekizumab				X
Secukinumab			X	X
Tofacitinib	X			
Ustekinumab				X

The ERG thinks that comparing outcomes at such varying times could lead to bias in favour of those drugs (including guselkumab) which had longer treatment durations, as patients had more time to respond, or gain larger responses. However, the ERG notes that these varying times are in line with recommended treatment durations, and so may it be a fair comparison in terms of how long treatment would be given in practice (including terminating treatment for non-responders). The ERG also notes that the option of "early escape" at 16 weeks for patients in the DISCOVER trials with poor PsARC performance may also have biased 24-week results in favour of guselkumab (see Section 3.2.2.2).

The ERG therefore requested the following NMAs as sensitivity analyses:

1. Using only 24-week data (including DISCOVER 1 and 2 at 24 weeks)
2. Using only 16, 14 or 12-week data (including DISCOVER 1 and 2 at 16 weeks)

The company partly responded to this request by providing NMAs for point 2. NMAs for PsARC were provided for point 1, but these did not match the ERG request, as they appeared to include data for treatments assessed before 24 weeks. When responding to this request the company stated:

...Guselkumab is not a typical therapy as it is a first in class regulator of the 'upstream' IL-23 pathway. As such, it offers a significantly longer duration of response compared to other comparators. As described in the submission, it is hypothesised that guselkumab's novel mechanism of action is the reason for this. However – possibly because of the same mechanism of action – it takes longer than 16 weeks to reach peak effectiveness...

The ERG notes that this appears to be inconsistent with claims in the CS that guselkumab has a "rapid" effect, and with outcome data presented in the CS (for full discussion see Section 3.2.2.4). The ERG therefore considers examining the data at 16 weeks to be essential. The ERG also notes that the optimal, unbiased analysis would therefore be to use outcomes only at 24 weeks (although this data does not exist for most treatments), and regrets that the company did not provide more analyses exclusively at 24 weeks.

3.4.2 Critique and summary of the network meta-analysis results

This section provides a summary of the NMA results presented in the CS, its appendices and responses to clarification questions.

Different studies contributed to different NMAs according to outcomes reported in each trial. Table 11 summarises the contributions of each trial in the biologic-naïve analyses (adapted from CS Appendix D Table 15).

Table 11 Contribution of studies to NMAs (biologic-naive population) (from CS Appendix D Table 15)

Trial name (sample size, n)	Timepoint used (weeks)	Comparators	Included in NMAs				
			PsARC	PASI	HAQ-DI (PsARC)	AEs	SAEs
ACTIVE (219)	16	PBO; APR				✓	✓
ADEPT (313)	12	PBO; ADA	✓	✓	✓	✓	✓
FUTURE 2 (298)	16	PBO; SEC150; SEC300	✓*	✓		✓*	✓*
FUTURE 3 (414)	16	PBO; SEC150; SEC300				✓*	✓*
FUTURE 4 (341)	16	PBO; SEC150 (noLD); SEC150; SEC 300					
FUTURE 5 (996)	16	PBO; SEC150(noLD); SEC150; SEC300				✓*	✓*
GO-REVEAL (259)	14	PBO; GOL	✓	✓	✓	✓	✓
IMPACT (104)	16	PBO; INF	✓	✓	✓	✓	✓
IMPACT 2 (200)	14	PBO; INF	✓	✓	✓	✓	✓
NA (100)	12	PBO; ADA	✓		✓	✓	✓
OPAL-BROADEN (318)	12	PBO; TOF; ADA	✓	✓		✓	✓
PALACE 1 (336)	16	PBO; APR	✓*	✓*	✓		
PALACE 2 (321)	16	PBO; APR	✓*	✓*	✓		
PALACE 3 (336)	16	PBO; APR	✓*	✓*	✓		
PSUMMIT 1 (615)	24	PBO; UST45; UST90	✓	✓	✓	✓	✓
PSUMMIT 2 (312)	24	PBO; UST45; UST90	✓	✓	✓		
RAPID-PSA (409)	12	PBO; CERT200; CERT400	✓*	✓*		✓*	✓*
SPIRIT-P1 (417)	12	PBO; IXE Q2W; IXE Q4W; ADA	✓	✓		✓	✓
SPIRIT-H2H (566)	24	IXE 80 mg Q4W/ Q2W; ADA 40 mg				✓	✓
Mease 2000 (60)	12	PBO; ETA	✓	✓			✓
Mease 2004 (205)	12	PBO; ETA	✓		✓		✓
Deodhar 2018 (149)	24	PBO; GUSQ8W	✓	✓	✓	✓	✓
DISCOVER-1 (381)	24	PBO; GUSQ8W; GUSQ4W	✓	✓	✓	✓	✓
DISCOVER-2 (739)	24	PBO; GUSQ8W; GUSQ4W	✓	✓	✓	✓	✓

To aid understanding, given the large number of NMAs performed, results are presented in simplified tables giving only the rankings of the treatments in the NMAs, and comparisons with selected treatments**.** The guselkumab q8w treatment is used in all tables; where other treatments had multiple dose groups the ranking tables show only the ranking for the best-ranked dose for that treatment. This was generally, but not always, the highest dose group. In all tables, interleukin inhibitors are highlighted in yellow; TNFα inhibitor drugs in blue. Green highlighting indicates where guselkumab was found to be superior to the named treatment, and the credible interval for the difference excluded the null value of 1. Similarly, cells highlighted in red show where guselkumab was found to be inferior to the named treatment, and the credible interval for the difference excluded the null. White

cells indicate where credible intervals for differences in ranking did not exclude the null value, and so are not conclusive.

It should be noted that ranking tables may disguise differences between treatments. For example, relative risks between treatments ranked 1 and 2 may be large , despite a ranking difference of only 1. Also, rankings here are simple rankings relative to placebo, and may not precisely represent the Bayesian ranking of treatment obtained from the NMA models (which are estimated with uncertainty).

Table 12 summarises the treatment rankings in the company-preferred NMAs (i.e. those presented in the CS and its appendices) for the biologic-naïve population (Population 2). These analyse are usually adjusted for placebo response, except for HAQ-DI analyses, where unadjusted model had a better fit. The CS did not report full network results comparing treatments for HAQ-DI (and these were not supplied after request for clarification), so likely rankings have been inferred from presented results, and the results for the 16-week analysis added for comparison.

These results suggest that guselkumab may be the best treatment for improving skin symptoms (based on PASI), and it was superior to most other treatments, except for secukinumab or ixekizumab. In contrast, guselkumab produced more modest results for other outcomes (ACR, PsARC, HAQ-DI), generally being ranked inferior to anti-TNFα drugs (although differences were not usually conclusive), and of similar ranking to secukinumab or ixekizumab. Of particular concern is its relatively poor ranking for HAQ-DI. Here guselkumab appeared to be inferior to etanercept and infliximab, and of poorer ranking than either secukinumab or ixekizumab.

Table 12 Summary of treatment rankings in the company-preferred NMAs (biologic-naive patients)

	ACR	PASI	PsARC	HAQ-DI *	HAQ-DI **	HAQ (PsARC responders)	HAQ (PsARC non-responders)	Adverse events
Guselkumab	5	1	6	7	8	5	1	2
Adalimumab	8	8	9	9	9	3	5	3
Apremilast	11	11	11	11	11	4	7	7
Certolizumab	6	6	4	6	6			10
Etanercept	3	10	2	1	1	2	4	
Infliximab	**1**	5	1	2	2	1	3	11
Golimumab	7	9	3	8	7	6	8	9
Ixekizumab	4	2	7	3	3			8
Secukinumab	2	3	5	4	4			6
Tofacitinib	10	7	10	10	10			1
Ustekinumab	9	4	8	5	5			4
Placebo	12	12	12	12	12	7	9	5

** Not a formal NMA comparison between treatments (hence no green/red highlights)*

*** Results from 16-week analysis*

Table 13 summarises the company-preferred analyses for the biologic-experienced patients (Population 3). These analyses were generally not adjusted for placebo response, as unadjusted model had a better fit. Guselkumab ranks better in general in these analyses, because the anti-TNFα drugs are excluded. The more limited data meant that few comparisons (except with placebo) were conclusive. As for biologic-naïve patients, guselkumab ranks as the superior treatment for PASI outcomes (but not clearly superior to secukinumab or ixekizumab). Of concern is the poor ranking on HAQ-DI at 16 weeks, where the ranking drops from 2nd to 6th, compared to the analysis in the CS. Other treatments do not change ranking substantially between these analyses, suggesting that analysing guselkumab at 24 weeks may be leading to over-estimation of its effect on HAQ-DI.

Table 13 Summary of treatment rankings in the company-preferred NMAs (biologic-experienced patients)

	ACR	PASI	PsARC	HAQ-DI *	HAQ-DI **	HAQ (PsARC responders)	HAQ (PsARC non-responders)	Adverse events
Guselkumab	3	1	1	2	6	1	1	1
Apremilast	5	6	7	7	7	2	3	8
Certolizumab	1	5	2	4	3			4
Ixekizumab	4	2	6	3	2			6
Secukinumab	2	3	3	1	1			2
Tofacitinib	7	7	5	5	4			7
Ustekinumab	6	4	4	6	5	3	2	5
Placebo	8	8	8	8	8	4	4	3

** Not a formal NMA comparison between treatments (hence no green/red highlights)*

*** Results from 16-week analysis*

3.4.2.1 Additional analyses requested by the ERG

Adjustment for placebo response

The ERG requested the results of analyses unadjusted for placebo response. It should be noted that, although we consider the unadjusted models here, the company analyses generally found that adjusted models had better model fit (in terms of DIC) and there was evidence that placebo response was associated with effect estimates. Hence the ERG considers that adjusted models (where reported) were a more plausible fit to the data.

To illustrate the potential impact of placebo-response adjustment, Figure 5 shows the estimated response rate in the NMAs of PASI 75 by treatment, comparing the adjusted model (black line) to the unadjusted models (grey lines). Using an adjusted model substantially increases the response rate for guselkumab, more than for any other treatment. By contrast, the response is substantially reduced for infliximab and golimumab, with smaller reductions for etanercept and apremilast.

Figure 5 Predicted PASI 75 response rate by treatment in the different NMA models

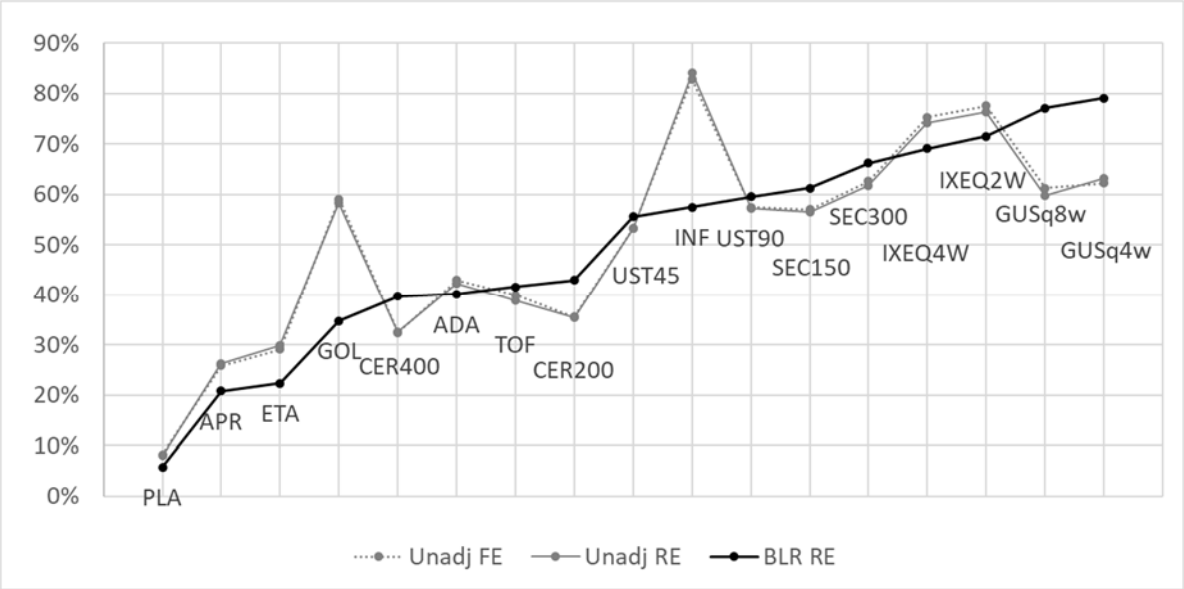


Table 14 summarises results from analyses not adjusted for placebo response in the biologic-naïve population. The ranking of guselkumab for ACR 20 and PsARC is little changed from the main analyses, although now it is inferior to most TNFα inhibitors. Guselkumab is no longer top-ranked on PASI 75. Guselkumab now ranks slightly lower than ixekizumab for both ACR 20 and PASI 75, but remains of similar ranking to secukinumab. Notable is the drop in ranking for adverse events (from 2nd to 7th). This is surprising as we might expect placebo-arm adverse event rates to be more robust over time. Results for biologic-experienced patients are not shown, as adjusted analyses were not generally performed for that population.

Table 14 Summary of treatment rankings in the unadjusted NMAs (biologic-naive patients)

	ACR 20	PASI 75	PsARC	Adverse events
Guselkumab	8	3	7	7
Adalimumab	6	7	8	1
Apremilast	11	11	10	8
Certolizumab	7	9	4	5
Etanercept	3	10	3	
Infliximab	1	1	2	11
Golimumab	2	5	1	10
Ixekizumab	4	2	6	9
Secukinumab	5	4	5	3
Tofacitinib	10	8	11	2
Ustekinumab	9	6	9	6
Placebo	12	12	12	4

Analyses using 16-week data

The ERG also requested analyses based on outcomes measured at 16-weeks only, to ensure consistency across trials. Table 15 summarises results from analyses specifically at 16 weeks in the biologic-naïve population. ACR results were not available for this case. For most outcomes, results are similar to the overall analysis. Guselkumab is no longer the top-ranking treatment based on PASI (although not conclusively different from secukinumab or ixekizumab). Ranking for PsARC and HAQ-DI remains generally lower to the TNFα inhibitors. HAQ-DI ranking for guselkumab is also rather lower than for secukinumab or ixekizumab.

Table 15 Summary of treatment rankings in the 16-week data NMAs (biologic-naive patients)

	PASI	PsARC	HAQ-DI	HAQ (PsARC responders)	HAQ (PsARC non-responders)
Guselkumab	2	6	9	7	4
Adalimumab	8	9	8	3	5
Apremilast	11	11	11	5	6
Certolizumab	6	3	6		
Etanercept	10	2	1	2	3
Infliximab	5	1	2	1	2
Golimumab	9	4	7	6	7
Ixekizumab	1	8	3		
Secukinumab	3	5	4		
Tofacitinib	7	10	10		
Ustekinumab	4	7	5	4	1
Placebo	12	12	12	8	8

Table 16 summarises results from analyses specifically at 16 weeks in the biologic-experienced population. Most notable here is that ranking for guselkumab on PASI drops lower than for secukinumab or ixekizumab (although this is not a conclusive difference). Also of note is the low ranking on HAQ-DI, as mentioned above.

Table 16 Summary of treatment rankings in the 16-week data NMAs (biologic-experienced patients)

	PASI	PsARC	HAQ-DI **	HAQ (PsARC responders)	HAQ (PsARC non-responders)
Guselkumab	3	2	6	4	3
Apremilast	6	7	7	1	2
Certolizumab	5	1	3		
Ixekizumab	2	5	2		
Secukinumab	1	3	1		
Tofacitinib	7	6	4		
Ustekinumab	4	4	5	2	1
Placebo	8	8	8	3	4

Comparison of treatment effect estimates

The tables above have summarised the ranking of treatments in different models. Here we consider the impact of the different models on the relative treatment effect between guselkumab and other treatments. We focus on comparing guselkumab (q8w) to secukinumab and ixekizumab (as these appear most similar to guselkumab), and to etanercept (as an example TNFα inhibitor). Table 17 summarises the relative treatment effects for ACR20, PASI 75 and PsARC across the various analysis models. Here pale green cells indicate where guselkumab was ranked higher, and pale pink where guselkumab was ranked lower, than the comparator. Differences in rankings, however, were not generally conclusive (most credible intervals include 1).

For PsARC the different models (unadjusted or 16-week only) have only very moderate impacts on comparative effect estimates, with small changes in relative risk estimates. For ACR 20 changes in relative risks are also small, although in all cases the size of relative risk increases, in favour of the comparator treatments, when using unadjusted models.

The largest changes come for PASI 75. Here, the relative risks in favour of guselkumab in the company-preferred models reduce substantially when switching to either unadjusted or 16-week data models. Any possible benefit over secukinumab or ixekizumab disappears entirely in unadjusted and 16-week data models, and, for biologic experienced patients with 16-week data, switches to favouring secukinumab or ixekizumab.

Table 17 Relative risks comparing guselkumab to key treatments in NMAs

Analysis	Drug	Risk ratio (95% CI) compared to guselkumab (q8w)		
		ACR 20	PASI 75	PsARC
Biologic naive	Etanercept	1.10 (0.84 - 1.43)	3.7 (1.89 - 12.5)	1.12 (0.96 - 1.33)
Company-preferred	Ixekizumab	1.06 (0.84 - 1.33)	1.1 (0.94 - 1.39)	1.1 (0.91 - 1.41)
	Secukinumab	1.12 (0.89 - 1.35)	1.18 (0.96 - 1.72)	1.01 (0.81 - 1.21)
Biologic experienced	Ixekizumab	1.03 (0.53 - 1.82)	1.06 (0.61 - 1.82)	1.11 (0.71 - 1.56)
Company-preferred	Secukinumab	1.05 (0.58 - 2.07)	1.09 (0.62 - 1.92)	1.09 (0.7 - 1.56)
Biologic naive	Etanercept	1.47 (1.10 - 1.95)	2.16 (0.85 - 9.82)	1.29 (1.02 - 1.63)
Unadjusted	Ixekizumab	1.18 (0.81 - 1.64)	1.28 (0.87 - 1.80)	1.00 (0.69 - 1.35)
	Secukinumab	1.14 (0.79 - 1.70)	1.04 (0.47 - 1.67)	1.03 (0.68 - 1.89)
Biologic naive	Etanercept		3.16 (1.60 - 9.71)	1.14 (0.97 - 1.32)
16-week	Ixekizumab		1.04 (0.87 - 1.27)	1.07 (0.89 - 1.31)
	Secukinumab		1.03 (0.82 - 1.39)	1.05 (0.86 - 1.22)
Biologic experienced	Ixekizumab		1.14 (0.64 - 2.30)	1.10 (0.71 - 1.57)
16-week	Secukinumab		1.33 (0.82 - 2.85)	1.15 (0.73 - 1.70)

3.4.2.2 Serious Adverse Events

The company performed a NMA of serious adverse events (see CS Fig 30 page 110). The ERG notes that, because there were very few SAEs in any trial, credible intervals for this NMA are extremely wide, and the results are not likely to be robust to even minor changes in the number of SAEs in any trial. The ERG concludes that no conclusions can safely be drawn about the SAE profile of guselkumab in comparison to other treatments.

3.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG requested full BUGS-style data and code in order to replicate the submitted NMA analyses. Unfortunately, the ERG could only replicate the analyses for PsARC. Models for PASI and ACR, which were analysed as ordered categorical variables, did not converge or run successfully in BUGS (the analysis package used). This may be because the complexity of these models requires very careful specification of data structures, computer code or initial value estimates (not supplied by the company). While the ERG does not consider the models supplied to be in error (the BUGS code was valid), and so does not consider the analyses invalid, it is naturally concerned about the non-replicability of the company's analyses.

This issue prevented the ERG from performing any further NMAs.

3.5.1 The EXCEED trial

The ERG identified a new trial called EXCEED[7] which compared secukinumab to adalimumab. The company noted that EXCEED would not have been eligible for their NMAs (for example because not all patients in the trial had a PASI >10). However, because it is a recent trial including western European and US patients, the ERG deems it important to consider its results.

Table 18 summarises the outcomes reported in EXCEED (ACR 20 and PASI 90, at 24 weeks) and compares them to DISCOVER and ADEPT (the largest adalimumab vs placebo trial). This shows that the secukinumab arm in EXCEED outperformed guselkumab on ACR 20 in both DISCOVER trials, and outperformed DISCOVER-1 for PASI 90. Adalimumab in EXCEED also outperformed guselkumab on ACR 20 (but not other outcomes), in contrast to the NMA in Table 12. This may be partially explained by different distributions of biologic-experienced or naive patients in the trials.

While this is obviously limited by being a naïve indirect comparison, these results raise concerns about the top-ranking of guselkumab on PASI observed in the NMAs, and also raises doubts as to the robustness of the placebo-response adjusted model.

Table 18 Comparison of the Exceed trial to the DISCOVER and ADEPT trials

Trial	Treatment	ACR 20 % response	PASI 90 % response
EXCEED	Adalimumab	64	42
	Secukinumab	71	63
DISCOVER-1	Placebo	22	12
	Guselkumab (q8w)	52	50
DISCOVER-2	Placebo	33	10
	Guselkumab (q8w)	64	69
ADEPT	Placebo	14	0
	Adalimumab	58	30

3.6 Conclusions of the clinical effectiveness section

3.6.1 DISCOVER trials

The evidence on the efficacy and safety of guselkumab is based on the results of three placebo-controlled RCTs. These showed statistically significant and clinically important benefits of guselkumab when compared to placebo across important outcomes. Guselkumab can therefore be assumed to be an effective treatment for improving symptoms of psoriatic arthritis. Guselkumab appears to have an acceptable safety profile.

Although the trials were largely conducted appropriately, the ERG was concerned about how 'early escape' patient data were analysed at 24 weeks in the two largest trials (DISCOVER 1 and DISCOVER 2); the approach used appears to have resulted in some biased effect estimates in DISCOVER 1. The ERG also has concerns about the applicability of the trial results to an NHS setting, given that most patients were recruited in Eastern European settings in patients who had mostly not received two prior synthetic DMARDs and who had high baseline PASI scores. The anticipated recommended stopping rule for guselkumab due to lack of response is expected to be 24 weeks, but the DISCOVER trial results suggest that most of the patients who are going to respond to guselkumab, respond by week 16. Although responses after week 16 may be seen, they do not appear to be specific to guselkumab. The ERG considers there to be little reliable evidence of a specific time-to-response effect suggesting that guselkumab warrants a 24-week stopping rule.

The DISCOVER-1 trial included a small subgroup of 119 patients who had previous exposure to a TNFα inhibitor. Although relative effect estimates versus placebo were similar across the two subgroups (TNF-naïve and TNF-experienced), the TNF-experienced subgroup was too small to provide any conclusive evidence about similarity of efficacy. Subgroups were also reported based on severity of psoriasis. Although there was a suggestion that efficacy might increase as psoriasis severity increases, the number of events was sometimes small in the minimal and moderate-severe subgroups resulting in uncertainty, reflected in the wide confidence intervals associated with the effect estimates.

3.6.2 Network meta-analyses

Guselkumab was compared to other treatments in network meta-analyses (NMAs). The limited data, however, means that NMAs are equivalent to indirect comparisons using the Bucher method, and may be subject to the kinds of bias typical in such analyses. The ERG agrees with the principle of adjusting for placebo response in the NMAs, but notes that, because of the limited number of trials, its robustness is uncertain.

The ERG's main concern was with using outcomes assessed at 24 weeks for guselkumab and comparing these with outcomes assessed at 16 weeks (or earlier) for other treatments. This may unfairly bias results in favour of guselkumab. The ERG does not consider the company's claim that guselkumab takes longer to reach full effectiveness to be completely justifiable, given the evidence in the DISCOVER trials. The ERG therefore thinks that the NMAs based on 16-week data are the fairest comparison of guselkumab with other treatments. The ERG stresses that this relates to the fair comparison of treatments in the NMAs, and should not be taken as the ERG advocating that guselkumab should only be given for 16 weeks.

The results of the NMAs suggest that guselkumab is one of the most effective treatments at reducing skin symptoms (as measured by PASI), but may potentially be less effective than TNFα inhibitors at reducing joint symptoms (measured by ACR or PsARC outcomes). Of some concern was guselkumab's poor ranking on the HAQ scale measuring quality of life, where it ranked lower than most other commonly used treatments. It should be noted that, due to the limited data, most differences in effectiveness across treatments were not conclusive and most credible intervals included the null value.

Guselkumab appeared to be very similar in effectiveness to other interleukin inhibitors (secukinumab and ixekizumab) for all outcomes. All three drugs were ranked higher than TNFα inhibitors on PASI outcomes, but lower on ACR and PsARC outcomes. The ERG concludes that guselkumab has a broadly similar action and effectiveness to other interleukin inhibitors, with no evidence that it is superior or inferior to them.

Switching from the company-preferred NMAs to those without placebo-response adjustment, or to those using only 16-week data generally reduced the relative effectiveness of guselkumab on all outcomes, but changes were not large enough, in most cases, to substantially alter treatment rankings, or the general conclusions.

3.6.3 Overall conclusions

Guselkumab appears to be a safe and effective drug, when compared to placebo, and improves all major outcomes associated with psoriatic arthritis. It appears to be effective both in patients who are biologic-naïve and in those who are biologic experienced.

Guselkumab appears to be of similar efficacy to other interleukin inhibitors (secukinumab and ixekizumab). Like the other interleukin inhibitors, guselkumab appears to be potentially superior to TNFα inhibitors at reducing skin symptoms (as measured by PASI) but potentially inferior for joint symptoms (ACR and PsARC outcomes).

Overall, guselkumab appears to be an effective treatment for psoriatic arthritis, but it is unlikely to be clinically preferable to existing TNFα inhibitors for most patients. Its similarity to secukinumab and ixekizumab suggest it may be a reasonable alternative to those therapies. It may therefore be best considered as a treatment when TNFα inhibitors are ineffective or unsuitable, or where patients have severe skin symptoms.

4 COST EFFECTIVENESS

4.1 ERG comment on company's review of cost-effectiveness evidence

The company's review did not identify any studies on the cost-effectiveness of guselkumab in PsA. The company identified 15 cost-effectiveness studies[8-22] and 17 HTA reports.[23-43] Nine of the cost-effectiveness studies and 7 HTA reports were considered relevant to inform decision-making in UK. The data extraction and quality assessment for each study are presented in Table 3 and 4, Appendix G of the CS, respectively. The company concluded that all recent models submitted to NICE shared a common structure, based on the York model (TA199). No further results were presented or commented by the company.

Points for critique

Overall, the literature searching for the cost-effectiveness systematic review was of a high standard and well reported. One of the two search filters was not referenced, but it appears to be the *CADTH search filter Economic Evaluations/ Cost/Economic Models – OVID Medline*. Therefore, the ERG considers that all relevant publications are likely to have been identified.

4.2 Summary and critique of the company's submitted economic evaluation by the ERG

The Markov model submitted by the company tracks PsA patients eligible for bDMARD therapy over their lifetime. At each line of active treatment, patients receive a bDMARD or BSC over an initial trial period to assess treatment response, where the duration of response depends on the active therapy. At the end of the trial period, response is assessed with responders continuing into maintenance treatment, while non-responders move to the next line of treatment. Responders move to the next line of therapy if they discontinue treatment due to loss of response or adverse events. Number of lines of therapy depend on the subpopulation under analysis. In contrast to all previous NICE TAs of bDMARDs in PsA, discontinuation rates were treatment specific.

Key scenario analyses conducted by the company explored the impact of alternative i) approach to estimate HAQ-DI conditional on PsARC, ii) HAQ-DI rebound assumptions, iii) response definitions, iv) source of utility data, v) source of disease related costs, and vi) discontinuation rates.

4.2.1 History of NICE appraisals

NICE has previously appraised the use of biologics for PsA in adults in a number of TAs, which include the following single or multiple technology appraisals: Ixekizumab (TA537) 2018; Tofacitinib (TA543) 2018; Certolizumab pegol and secukinumab (TA445) 2017; Ustekinumab (TA340) 2017; Apremilast (TA433) 2017; Golimumab (TA220) 2011; Etanercept, infliximab and adalimumab

(TA199) 2010. In addition, there is one forthcoming appraisal, Upadacitinib [ID2690], and a terminated appraisal, Abatacept [TA568] 2019.

The CS describes the decision model as following the structure of recent submissions to NICE, namely TA445, TA537 and TA543. However, the company's cost-effectiveness analysis differ from previous TAs in the following key elements:

1. Discontinuation from maintenance treatment – previously assumed to occur at an annual rate of 16.5% for all treatments based on registry data. The company's preferred approach sourced treatment specific discontinuation rates from pivotal trials extension data, which mostly had unclear treatment stopping rules and may not reflect clinical practice.
2. Heterogeneity in psoriasis severity – appraisals since TA445 have presented base-case cost-effectiveness results for three psoriasis subgroups rather than a single set of results for an 'average' PsA patient. Although the CS presents in Appendix T psoriasis subgroup cost-effectiveness results for all analyses, the results presented in the main CS refer to an 'average' PsA patient reflecting the characteristics of the patient population in the DISCOVER trials. The baseline PASI score for each psoriasis severity subgroup was sourced from the DISCOVER trials, which is generally higher than the values used in previous appraisals (TA543, T537 and TA445). The DISCOVER trials also have a high proportion of patients with mild to severe psoriasis (78.9%).
3. Assumption of class effects in NMA models – Base-case analyses of TA445 and TA543 relied on effectiveness estimates (PsARC response, HAQ-DI change conditional on PsARC response and PASI response) sourced from placebo adjusted NMA models, which assumed class effects. The CS did not consider class effects in the NMA models.
4. Categories of PASI response – the models in appraisals since TA445 have classified patients in maintenance treatment according to four categories of PASI response ([0-49], [50-74], [75-89], [90-100]), whereas the CS (and TA537) considers one additional category by splitting the PASI response ≥90 category into two: 90-99, and 100.
5. Source of arthritis related costs – appraisals since TA199 have used the Kobelt et al., 2002[44] algorithm to estimate arthritis related costs based on the linear relationship between HAQ-DI scores and health care costs in rheumatoid arthritis. The CS uses an alternative algorithm developed by McHugh et al., 2019 based on cost data in PsA patients.

The appropriateness and implications of these differences between previous appraisals and the CS are discussed in the relevant sections below.

4.2.2 *NICE reference case checklist*

The company model's compliance with the NICE reference case is checked in Table 19.

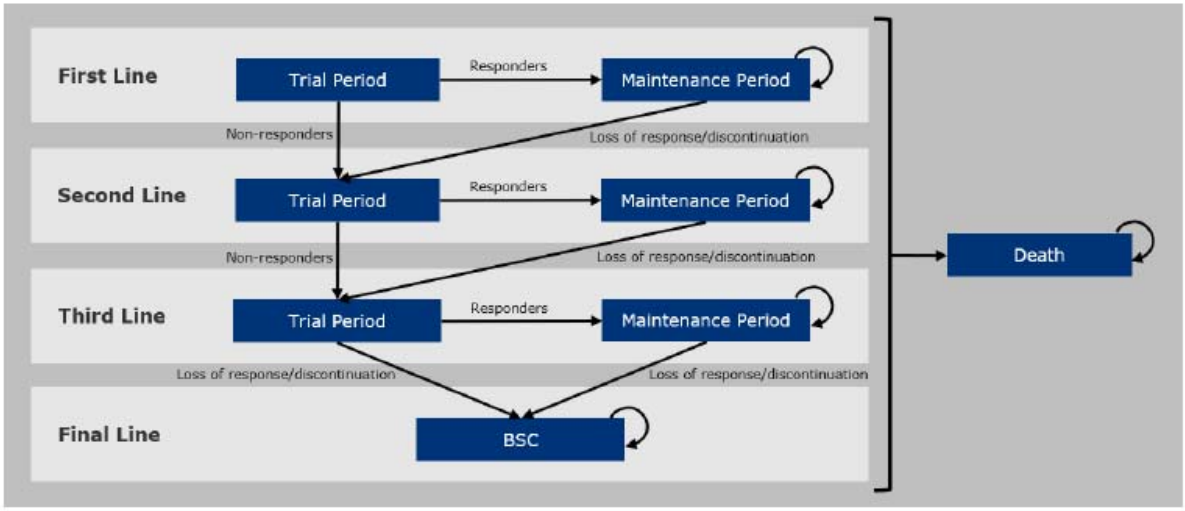
Table 19 NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	The CS is appropriate.
Perspective on costs	NHS and PSS	The CS is appropriate.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The CS is appropriate.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The CS is appropriate.
Synthesis of evidence on health effects	Based on systematic review	The CS is appropriate.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	The CS is appropriate. HRQoL is expressed in QALYs. A mapping algorithm is applied to convert HAQ-DI and PASI scores estimated by the model into EQ-5D-3L.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	The CS is appropriate.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The CS is appropriate.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The CS is appropriate.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The CS is appropriate.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	The CS is appropriate.
PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome.		

4.2.3 *Model structure*

The company developed a *de novo* Markov model based on the model used in TA445 (referred to as the updated York model), and described as being similar to the model used in TA537. The model cycle length was 28 days and no half-cycle correction was applied. Patients are tracked over their lifetime (40 years). The model structure is illustrated in Figure 6.

Figure 6 Model schematics (CS, Figure 33)



The model compares the cost-effectiveness of treatment sequences including guselkumab and comparators. It tracks changes in HAQ-DI and PASI scores to estimate health state HRQoL and costs based on these scores.

Patients enter the model and receive treatment with a bDMARD or BSC until they have completed the response assessment period, the duration of which is treatment specific (12 weeks for all bDMARDs, except i) apremilast, secukinumab and ixekizumab [16 weeks], and ii) guselkumab and ustekinumab [24 weeks]). Trial period health states were modelled as tunnel states to account for the different durations of the trial period by treatment. Treatment benefit in terms of HAQ-DI (functional capability) and PASI (psoriasis severity) scores is realised immediately upon model entrance, with the extent of benefit conditional on response to treatment for HAQ-DI score improvements (PsARC response in the base-case) and the level of PASI response (which is assumed to be correlated with PsARC response). There was no treatment discontinuation permitted during the trial period. At the end of the trial period, responders (in the base-case, PsARC responders) transition to the maintenance period health state, while non-responders move to the next line of therapy in their treatment sequence. HAQ-DI and PASI scores rebound to baseline values when non-responders leave the trial period.

Patients in the maintenance period remain on treatment until they withdraw from treatment due to either adverse events or loss of treatment response, and move to the next line of therapy. The rates of discontinuation are treatment specific. The maintenance period health state is stratified according to the level of PASI response ([0-49], [50-74], [75-89], [90-99], [100]). The model assumes that patients in the PASI 0-49, 50-74, 75-89, 90-99 and 100 categories achieve an improvement of 25%, 50%, 75%, 90% and 100% over the baseline PASI score, respectively. During maintenance treatment, reductions in HAQ-DI and PASI scores are maintained for patients in active treatment. These scores both rebound to baseline values when patients transition out of the maintenance period.

When patients reach the final line of therapy, they receive BSC, which was modelled as not having a trial period. Once patients reach this line of therapy, their HAQ-DI scores begins to deteriorate at a constant rate until reaching the maximum score of 3, while their PASI scores remain constant (and equal to baseline PASI).

Patients in all health states are at risk of adverse events, with probabilities being treatment specific, and subject to a disease specific age and sex adjusted mortality risk.

Points for critique

The model is largely consistent with the updated York model, which the ERG considers appropriate. However, it presents the following differences:

- Trial period has varying duration according to treatment, while TA445 (and TA543) had a common duration of three months for this period. This common period was assumed in TA445 to reflect response assessment around 12-16 weeks, as the authors considered that there was lack of a clinically meaningful difference in the biologics' PsARC and PASI response rates between 12 and 16 weeks. Since the model assumes that benefits in terms of HAQ-DI and PASI score reductions are accrued instantly upon entering the trial period and that all patients who enter the trial period complete it (unless they die), patients treated with biologics with longer trial periods who do not respond will accrue some clinical benefit for a longer time interval than those on biologic treatment with shorter trial periods.
- The model classifies patient in maintenance treatment according to five PASI response categories ([0-49], [50-74], [75-89], [90-99], [100]), rather than four categories of PASI response ([0-49], [50-74], [75-89], [90-100]) in TA445 (and TA543). The model in TA537 had this same structural feature as the one in the CS. This alternative assumption has implications on HRQoL, as EQ-5D scores mapped (partially) from PASI scores will be comparatively higher when the more granular categories are considered, particularly for biologics with a better PASI response.
- Treatment discontinuation rates for patients in the maintenance cycle are treatment specific
- Inclusion of treatment specific adverse events with associated costs and QALY loss. Previous appraisals did not include adverse events in their model structure.

The implications of these changes are discussed in the relevant sections below.

The company states that three lines of therapy before the final line were considered in the model (CS, p128). However, the treatment sequences considered in the model only consider at most two lines of active treatment. The ERG comments further on this in Section 4.2.5.

4.2.4 *Population*

The cost-effectiveness of guselkumab was assessed in three out of the four subpopulations defined in the NICE scope:

- Subpopulation 2: Patients whose disease has not responded adequately to at least 2 cDMARDs (biologic naïve);
- Subpopulation 3: Patients whose disease has not responded adequately to cDMARDs and ≥1 TNFi (biologic experienced); and
- Subpopulation 4: Patients in whom TNFi are contraindicated (TNFi contraindicated).

The company did not present evidence for subpopulation 1 (patients whose disease has not responded adequately to only one previous cDMARD) because of limitations in the availability of trial data by previous use of cDMARDs. This is in line with the subpopulations assessed in previous TAs (TA445, TA537 and TA543).

The model further discriminates the subpopulations based on the presence and severity of concomitant psoriasis. The severity of psoriasis was defined as: i) minimal psoriasis (BSA <3%); ii) mild to moderate psoriasis (BSA ≥3% and PASI ≤10); and iii) moderate to severe psoriasis (BSA ≥3% and PASI >10). The cost-effectiveness results for the subgroups by psoriasis severity are presented in Appendix T of the company's submission. The base case cost-effectiveness results are presented for 'All patients' based on the proportion of patients with minimal (21.1%), mild to moderate (47.9%), and moderate to severe (31.0%) psoriasis, which was informed by pooled data from the DISCOVER-1 and DISCOVER-2 trials.

Baseline characteristics of patients in the DISCOVER trials were assumed to reflect those of patients seen in UK clinical practice. Therefore, population characteristics used in the model (such as age, proportion of males, weight) and baseline PASI and HAQ-DI scores were based on pooled data from the DISCOVER trials. Table 20 presents the baseline PASI and HAQ-DI scores used in the model for each subpopulation and subgroup.

Table 20 Baseline PASI and HAQ-DI scores used in the model for each subpopulation and subgroup

	Biologic naïve subpopulation / TNFi contraindicated subpopulation*		**Biologic-experienced subpopulation**	
	Baseline PASI	**Baseline HAQ-DI**	**Baseline PASI**	**Baseline HAQ-DI**
All patients	9.43	1.23	9.78	1.38
Minimal PsO	1.79	1.23	1.50	1.24
Mild to moderate PsO	4.97	1.22	5.54	1.46
Moderate to severe PsO	21.36	1.22	23.58	1.38
* TNFi contraindicated subpopulation was informed by data from the biologic naïve subpopulation.				

PsO, psoriasis; TNFi, tumour necrosis factor inhibitor;

Points for critique

As discussed in Sections 3.3 and 3.4.2 , the ERG has concerns about how well the patient populations of the DISCOVER trials align with the population seen in UK clinical practice. This has an impact on both the baseline population characteristics used in the model and, importantly, the average baseline PASI and HAQ-DI scores by subpopulation and subgroups defined by presence and severity of psoriasis. The definition of psoriasis severity differs slightly from that used in the previous TAs of TA445 (updated York model) and TA537 (note that TA543 did not differentiate subpopulations by psoriasis severity). In TA445, minimal concomitant psoriasis was defined as BSA <3% or PASI <2.5, while TA537 used a subgroup of no concomitant psoriasis. Mild to moderate concomitant psoriasis was defined as BSA ≥3% and PASI between 2.5 and 10 in TA445, while TA537 used BSA ≥3% and PASI ≤10. Moderate to severe concomitant psoriasis was defined as BSA ≥3% and PASI >10 in TA445, while TA537 used BSA >3% and PASI >10.

The proportion of patients from the DISCOVER trials with minimal (21.1%), mild to moderate (47.9%), and moderate to severe (31.0%) psoriasis was used to provide a weighted average baseline PASI and HAQ-DI score for each of the subpopulations (all patients). These proportions are considerably different from TA445 where clinical opinion suggested that 50% of UK patients have minimal concomitant psoriasis, 25% have mild to moderate concomitant psoriasis and 25% moderate to severe concomitant psoriasis. The corresponding baseline PASI scores for these subgroups in the updated York model were much lower than those reported in the DISCOVER trials: baseline PASI=0, 7.3 and 12.5 for minimal, mild to moderate, and moderate to severe psoriasis, respectively, while the baseline HAQ-DI score was 1.22 for all subgroups and subpopulations in the York model. The ERG assessed the impact of using the proportion of patients by psoriasis severity and baseline PASI and HAQ-DI scores from the updated York model based on clinical opinion relevant to UK practice in a scenario analysis.

The ERG requested justification in their points for clarification for sourcing the baseline population characteristics from the DISCOVER trials over the average patient characteristics from the wider set of studies included in the NMA. In response, the company provided a comparison of the patient baseline characteristics in the guselkumab clinical trials and NMAs (Table 21). Overall, the baseline characteristics of the DISCOVER trials are not dissimilar to the pooled clinical trial data included in the NMA; however, there are notable differences in the proportion of psoriasis with BSA >3%, duration of PsA, number of swollen and tender joints in DISCOVER-1, and baseline PASI score. The

difference in baseline PASI score is relatively small for the subpopulations but differences by psoriasis severity are not presented.

Table 21 Comparison of patient baseline characteristics in the guselkumab clinical trials and NMAs (provided by the company in response to ERG points for clarification)

	DISCOVER-1	DISCOVER-2	Deodhar (2018) Phase II trial	bDMARD-naïve NMA	bDMARD-experienced NMA
Baseline age (mean)	48.4	45.7	46.3	48.1	49.4
% male	51.2	52.5	51	50.8	46.5
Body weight (kg)	86	84.3	85	86	86
Duration of PsA	6	7	7	7.3	7.7
No. swollen joints (mean)	9.9	12.3	11.5	12.1	11.3
No. tender joints (mean)	19.2	21.3	20.5	21.7	21.4
Baseline PASI score (mean)	8.5	9.9	11.3	9.2	9.3
PsO BSA >3% (%)	65.4	73.5	NR	63.9	56.3
Baseline HAQ-DI score	1.2	1.3	1.4	1.2	1.2

bDMARD, biologic disease modifying anti-rheumatic drug; BSA; body surface area; HAQ-DI, health assessment questionnaire-disability index: NR, not reported; PASI, psoriasis area and severity score; PsA, psoriatic arthritis: PsO, psoriasis.

item 1. Baseline PASI scores and proportion of patients by psoriasis severity in the DISCOVER trials may not match those seen in UK clinical practice.

4.2.5 *Interventions and comparators*

A wide range of comparators, including TNFi's (adalimumab, infliximab, etanercept, certolizumab and golimumab), PDE-4 inhibitor (apremilast), IL-17 modulators (secukinumab and ixekizumab), IL-12/23 modulator (ustekinumab), and JAK inhibitor (tofacitinib) were modelled in line with their licensed dose and assessment time point according to NICE guidance (see Table 61 of the company's submission). Guselkumab (IL-23 modulator) was modelled as having a stopping rule at week 24 based on the anticipated SmPC. The base-case cost-effectiveness analysis considered guselkumab Q8W as per the anticipated SmPC, while the cost-effectiveness results for guselkumab Q4W dosing were included in an appendix to the company's submission.

BSC was included as a comparator for each subpopulation. This was modelled similarly to the targeted therapies during the trial period. BSC was assumed to be representative of the placebo arm of the clinical trials included in the NMA and modelled to have a trial period of 12 weeks for the

efficacy benefits of placebo. However, unlike the targeted therapies, at the end of the initial trial period HAQ-DI for BSC deteriorated according to the natural history of progression (PASI score remained constant while on BSC in line with the assumptions used for the active therapies).

The model allowed treatment sequences to be considered. The selection of comparators for the first treatment in a sequence for each subpopulation was based on previous NICE recommendations and the NICE scope. The selection of the second and third line treatment options was based on those included in previous TAs (TA445, TA537 and TA543). As the different subpopulations are eligible for different treatment options, the number of treatment options varies by subpopulation and the length of the treatment sequence also varies by subpopulation (see Table 22). The same sequence of treatments was modelled for the presence and severity of psoriasis within a subpopulation, but the dosage for some of the comparators reflected the psoriasis severity level (see Table 10 of Appendix T of CS).

Ustekinumab was selected as the second line therapy in subpopulation 2 (biologic naïve subgroup) because NICE guidance recommends that it is only used after two cDMARDs and at least one TNFi, or where treatment with TNFi is contraindicated. When ustekinumab was considered as a first line therapy in subpopulation 4 (TNFi contraindicated subgroup), secukinumab was selected as the second line therapy. The third line of therapy for subpopulations 2 and 4 was BSC. For subpopulation 3 (biologic experienced subgroup), alternative targeted therapies were not considered at second line and patients received final line BSC. Patients receiving BSC as a comparator were assumed not to switch therapies in the model and therefore remained on the same treatment throughout the modelled time horizon.

For all subpopulations, patients switched to the next line of therapy following a lack of response to PsARC at the end of the trial period or by subsequent withdrawal due to loss of efficacy or adverse events in those patients who initially responded. For the biologic-naïve and TNFi contraindicated subpopulations, patients moving to a subsequent line of active therapy were assumed to have the same response probabilities as first line treatment of biologic-experienced patients. For the biologic-experienced subpopulation the number of prior bDMARDs failed was not modelled and no effect degradation was applied to subsequent lines of therapy. As a final line of therapy, BSC was not modelled to have a trial period. This means that HAQ-DI score deteriorated as soon as patients moved to the final line of therapy (PASI score remained constant while patients received BSC in the final line).

Table 22 Treatment sequences for each subpopulation included in the model

Subpopulation	Treatment options		
	First line	**Second line**	**Third line**
Subpopulation 2: Patients whose disease has not responded adequately to at least 2 cDMARDs (biologic naïve)	Guselkumab Adalimumab Apremilast Certolizumab[a] Etanercept Golimumab Infliximab Ixekizumab[b] Secukinumab[c] Tofacitinib BSC	Ustekinumab[d]	BSC
Subpopulation 3: Patients whose disease has not responded adequately to cDMARDs and ≥1 TNFi (biologic experienced)	Guselkumab Apremilast Certolizumab[a] Ixekizumab[b] Secukinumab 300mg Tofacitinib Ustekinumab[d] BSC	BSC	
Subpopulation 4: Patients in whom TNFi are contraindicated (TNFi contraindicated)	Guselkumab Ixekizumab[b] Secukinumab 300mg Tofacitinib Ustekinumab[d] BSC	Ustekinumab[d] Secukinumab 300mg (only when ustekinumab is first line)	BSC

[a]Mixed dose: 50% 200 mg Q2W and 50% 400 mg Q4W. [b]Mixed dose: 69% Q4W and 31% Q2W.
[c]Mixed dose: 69% 150 mg and 31% 300 mg. [d]Mixed dose: 79.9% 45 mg and 20.1% 90 mg.
bDMARD, biologic disease modifying anti-rheumatic drug; BSC, best supportive care; cDMARD, conventional disease modifying anti-rheumatic drug; Q2W, every two weeks; Q4W, every four weeks.

Points for critique

The comparators included in the CS were modelled in line with their licensed dose and assessment time point according to NICE guidance. It is worth noting that the time point of assessment varies from 12 weeks (adalimumab, certolizumab, etanercept, golimumab, infliximab, and tofacitinib) to 24 weeks (ustekinumab). These assessment time points reflect the stopping rule based on response to treatment and are typically reflective of the individual treatment marketing authorisations. The data used in the NMAs are broadly consistent with the different time points of assessment (see Table 9). For guselkumab, the time point of assessment is 24 weeks based on the DISCOVER trials. The time point of assessment is important for cost-effectiveness because it affects the total costs of treatment since longer time points imply that the treatment is used by all patients for longer. Therefore, it is important to consider the implications for cost-effectiveness of time points of assessment up to the maximum of 24 weeks as determined by the evidence base (this is discussed further in Section 4.2.8).

The choice of modelled treatment sequences are consistent with previous NICE TAs and in line with the NICE scope. However, they do not reflect the full range of sequences of treatment seen in UK clinical practice. The ERG requested additional justification for the choice of modelled sequences for each subpopulation in light of the treatment pathway and full range of treatment options used in UK clinical practice. The company justified the choice based on previous NICE recommendations and indicated that there is no standardised approach to treatment sequencing in the UK beyond appropriate adherence to NICE recommendations.

The ERG sought further clinical opinion on the most common sequence of treatments used in clinical practice and the main reason for switching among different therapies. Clinical opinion suggests that switching among different TNFi therapies represents a valid treatment strategy as well as switching to IL modulators (ustekinumab, secukinumab and ixekizumab) or tofacitinib. Growing evidence suggests that there are complex and variable treatment patterns for bDMARDs in PsA patients, with discontinuation or switching of therapies mainly due to tolerability issues or lack of efficacy, as well as loss of efficacy over time and adverse events.[45] As a result, it is unlikely that there is a standardised approach to treatment sequencing. Registry data provides some indication of the frequency of therapy switching and outcomes among PsA patients. However, most of these registries will only show the effect of switching among TNFi agents because the newer therapies have been trialled on fewer patients. In response to the ERG's points for clarification regarding the possibility of a treatment sequence that runs through several TNFi agents before switching to IL modulators or tofacitinib, the company conducted a rapid review of data on the efficacy of TNFi cycling. The limited number of studies identified suggested that there are significant implications of TNFi cycling for efficacy outcomes, often reducing the benefit of the second TNFi by around 60%. On this basis the company argued that it would be clinically inappropriate to cycle through several TNFi's given that there are now therapies with new mechanisms of action available to clinicians. This means that the empirical question about whether to treat patients with the newer agents first or save them until after trying TNFi's remains unanswered.

In response to ERG points for clarification the company indicated that data based on market share showed that the most common therapies offered in second line are secukinumab and ustekinumab, while the most common therapies in third line are secukinumab, ustekinumab and ixekizumab. This suggests that there are a number of possible sequences of treatment that are not modelled in the company's analysis but would be considered reasonable in clinical practice. More importantly, it suggests, in line with clinical opinion, that multiple lines of therapy is a valid treatment strategy instead of placing patients on BSC at second and third line; patients are unlikely to receive only two active therapies in the biologic-naïve or TNFi contraindicated populations, or only one active therapy after ≥1 TNFi's in the biologic-experienced population.

item 2. The modelled treatment sequences do not reflect the range of treatment sequences seen in UK clinical practice. Importantly, treatment with multiple lines of active therapy is a valid treatment strategy instead of placing patients on BSC at second and third line.

For subpopulation 3, certolizumab is included as a comparator in line with the NICE scope. The ERG considers certolizumab to be a relevant comparator in this population; however, the ERG notes that the efficacy evidence available from the RAPID PsA trial, which is the only trial comparing certolizumab with placebo, explicitly excluded primary non-responders to a previous TNFi. Therefore, the evidence from this trial does not support the definition of the biologic-experienced subpopulation 3. The ERG could not re-run the NMA for the biologic-experienced population by excluding this trial, but it is expected that its inclusion would not lead to changes in the ranking of the remaining estimates in the biologic-experienced NMA.

The definition of BSC as a comparator or final line of therapy in the CS is unclear. In response to the ERG's points for clarification the company indicated that BSC represented a mixture of conventional synthetic DMARDs and NSAIDs for patients who failed two prior cDMARDs but that specific assumptions regarding the proportion of patients receiving cDMARDs or type of DMARD were not applied. As the effectiveness estimates for BSC are based on the placebo group in the NMA and the use of cDMARDs varies at baseline across the clinical trials included in the NMA, no specific definition relevant to UK clinical practice is used.

Another important consideration related to the intervention and comparators is the response criterion used in the CS to assess treatment continuation. The ERG notes that at the time of writing this report, guselkumab has not yet been granted a marketing authorisation for the treatment of PsA, which provides clarity on a stopping rule depending on response assessment following a trial period of a defined duration. The ERG also notes that NICE recommendations generally adopt response based treatment continuation rules, i.e., only those treatments that demonstrate clear evidence of response continue treatment. In the CS, treatment continuation is defined based on achieving a PsARC response. However, in its recommendations, NICE has also given consideration to the possibility of continuation on a bDMARD for patients whose PsARC response does not justify continuation of treatment but who show a PASI 75 response. For TNFi's, the evidence generally shows PsARC response (around 70% of patients) to be higher than PASI response (between 20 and 40% of patients). The opposite is true for IL-17, IL12/23 and IL-23 inhibitors, where the proportion of PASI responders (between 65 and 87%) is higher than the proportion of PsARC responders (between 55 and 70%). This suggests that there is a significant proportion of individuals who will achieve a PASI 75 response but not a PsARC response. This is particularly relevant for guselkumab, which is evaluated in the CS to have a median PsARC response level relative to the other comparator bDMARDs, but has the highest PASI 75 response (see Section 3.2.2.2). A formal decision rule for treatment continuation

based on achieving a PASI 75 response but not a PsARC response was not considered in the CS, or in previous TAs for PsA. This may be an important consideration for guselkumab in relation to the 16 week versus 24 week stopping rule (also see Section 4.2.8), where the strength of guselkumab relative to the other bDMARDs appears to be in achieving a higher PASI 75 response but this is achieved at week 24 rather than week 16.

item 3. Continuation of treatment based on achievement of PASI 75 response for patients whose PsARC response does not justify continuation of treatment is an area of uncertainty.

4.2.6 *Discontinuation rates of biological therapy*

Treatment discontinuation rates for patients who initially responded to treatment and entered the therapy-specific maintenance period were based on one-year discontinuation rates from clinical trials for each treatment (see Table 64 of CS). Therefore, the model assumes differential annual discontinuation rates for the intervention and comparators, which ranges from 6.9% for guselkumab to 22.1% for infliximab. These rates were largely sourced from pivotal clinical trial extension data (see Table 3, Appendix W of CS). Patients on BSC were assumed not to discontinue treatment (discontinuation rate was set to 0%). Where multiple trials provided estimates for the same treatment, the company used a weighted average of the trial populations. Where a 'mixed' comparator was considered, based on different drug dosages, the company used a weighted average of the discontinuation rates according to the proportion of patients on each dosage or regimen. It is unclear whether discontinuation rates estimated at time points other than 52 weeks (e.g., 48 weeks for certolizumab in the RAPID PsA trial) were adjusted for in the model.

The use of treatment specific discontinuation rates is a key difference from previous TAs in PsA since TA199, which have all assumed a 16.5% annual discontinuation rate for bDMARDs that is not treatment-specific. This 16.5% estimate was derived from a meta-analysis conducted by the assessment group for TA199 based on registry data on PsA patients on first line therapy with bDMARDs (etanercept, adalimumab and infliximab) and represents the annual rate of withdrawal beyond the first 3 months of treatment (i.e., in the treatment maintenance period).

The company explains that the assumption of equivalent discontinuation rate used in previous TAs is not supported by evidence subsequent to TA199 and it supports instead the use of treatment-specific discontinuation rates. This evidence is presented in appendix W of the CS and summarised as follows:

- Recognition of limitations of the evidence base used to inform discontinuation rates in TA199 by the NICE committee;

- Discontinuation rates for guselkumab Q8W at 52 weeks in the DISCOVER trials are considerably lower than 16.5% (8.7% DISCOVER-1 and 5.6% in DISCOVER-2);
- Observational data suggests different discontinuation rates by treatment:
    - Subgroup analysis of PsA patients in an international psoriasis registry (PSOLAR, data collected 2007-2013) suggests statistically significant differences between time on treatment for the ustekinumab curve (n=361) vs. i) adalimumab (n=402, p<0.01), infliximab (n=63, p<0.01) and etanercept (n=289, p=0.011). The ERG notes that these differences were only statistically significant when drugs were compared at the second-line of therapy, but not at first or third-line of therapy (with the exception of etanercept vs. ustekinumab at third-line therapy). The absolute annualised discontinuation rates were lower for ustekinumab compared to TNFi therapies.
    - US insurance claim database (Clinformatics, data collected 2013-2016).
      The company presents the proportion of patients with at least one claim for PsA biologics persisting on treatment for ≥24 months and mean persistence for each index biologic. Treatments included adalimumab (n=477), certolizumab (n=45), etanercept (n=344), golimumab (n=49) and ustekinumab (n=81). The differences in persistence after 2 years and duration of persistence are suggested by the company to support a differential discontinuation rate for non-TNFi agents, despite the small sample size and large standard deviation values for duration of persistence.
    - Analysis of the international PsABIO registry (unknown period of data collection) comparing the 1-year persistence of ustekinumab vs TNFi.[46]
      The company presents estimates of mean survival on treatment by line of treatment (1st, 2nd and 3rd line) with either ustekinumab (n=438) or a TNFi (n=455) at the rheumatologist discretion. The company states that the data suggests longer duration of treatment on ustekinumab compared to TNFi, even when differences suggesting longer ustekinumab persistence vs. TNFi were only statistically significant for some subgroups (bDMARDs in monotherapy and BSA>10%).
- Discontinuation rates in extensions of pivotal clinical trials used to parameterise the model support the use of treatment specific estimates.
- Guselkumab RCT data for the treatment of psoriasis (extension of VOYAGE 1 and 2) suggest a lower discontinuation rate at 48 weeks compared to adalimumab). The company also estimated the annual discontinuation rate for guselkumab in the PsA subgroup of the VOYAGE trials as approximately 8.8%, which was considered consistent with the estimate for the whole trial population (7.9%).

In response to the ERG's points for clarification (Question B7), the company presented the results of a review of PsA registers conducted as part of a targeted literature review on treatment specific discontinuation rates. The review identified 6 registry studies[47-51] [52]which reported sufficient data to inform treatment-specific 12-month discontinuation rates in PsA patients. Discontinuation rates sourced from these studies was presented and contrasted against the estimates applied in the model (Table 10, response to ERG's points for clarification). The company interprets the data as suggesting that discontinuation rates are higher in registry vs. clinical trial data and that this is probably due to a higher proportion of biologic-experienced patients in the registries. The company also notes that registry data is not available for all treatments.

Points for critique

The treatment-specific withdrawal rates used in the company's model is an important driver of cost-effectiveness and therefore the evidence supporting it must be carefully critiqued. Firstly, it is worth noting that the description in the CS of the limitations of the discontinuation rate of 16.5% used in the previous TAs for all treatments is potentially misleading. The ERG notes that these comments were made in relation to the data that could have been used to generate differential rates of discontinuation for adalimumab, etanercept and infliximab, and not the meta-analysis which estimated the discontinuation rate of 16.5%. The ERG to TA199 was concerned that the hazard ratios for treatment discontinuation for each of the bDMARDs under appraisal were based on the full time on treatment, and did not allow for separate estimates to be calculated for treatment discontinuation after the response assessment period. Furthermore, this analysis found no difference in discontinuation between etanercept and adalimumab, only between these two drugs and infliximab. There were, however, concerns of bias against infliximab, the first TNFi in the market, as it may have been used on severe patients with low expectation of maintaining drug therapy.

Secondly, the observational studies presented by the company in Appendix W in support of differential discontinuation rates have small sample sizes, and differences in treatment persistence between biologics are mostly not statistically significant. Furthermore, the data presented does not appear to distinguish between discontinuation due to lack of response in the initial biologic trial period and due to loss of response in the maintenance period.

The company does not describe the inclusion/exclusion criteria for the review of discontinuation rates in PsA registry data provided in response to ERG points for clarification (Question B7), nor does it report the data extracted from each study to derive the registry data informing the treatment discontinuation rates (with the exception of ustekinumab) in Table 10 (response to ERG points for clarification). Therefore, it is not possible for the ERG to comment on whether the review missed relevant data. The ERG notes, however, that at least one relevant UK registry study appears to have

been omitted or missed by the searches. Fagerli and colleagues (2017[53]) reported treatment-specific time to treatment discontinuation data for etanercept, infliximab and adalimumab based on data collected in PsA patients between 2001 and 2006 in the British Society for Rheumatology Biologics Register (BSRBR). In the absence of the extracted data and the exact sources used, the ERG is also unable to validate the treatment-specific estimates derived by the company from registry data.

The majority of studies used to inform treatment-specific discontinuation rates in the model do not report treatment stopping rules in the maintenance period. Therefore, it is possible that patients in these trials continued treatment beyond the loss of sustained response. Importantly, and as mentioned in Section 3.2.2.6 there were no stopping rules based on lack of efficacy in the 52-week DISCOVER trials, and, therefore, it is plausible that the discontinuation rate assumed for guselkumab does not reflect the rate at which patients in clinical practice would withdraw from treatment. The low discontinuation rate may also reflect the lack of alternative therapeutic alternatives in the health care system for patients without a sustained response to guselkumab, as highlighted in section 3.2.2.1 and 3.3.

The company suggests that better skin results with ustekinumab and guselkumab may contribute to lower discontinuation rates for these treatments, and that patients may be kept on treatment with a bDMARD despite equivocal joint response if skin response remains adequate. However, patients with PsA are primarily treated with bDMARDs to control joint disease and not skin symptoms, as psoriasis in PsA patients is usually on the less severe end of the spectrum. It is also possible that discontinuation rates based on extension trial data is double counted for some or all treatments, as the estimates do not appear to have excluded discontinuation due to lack of response in the initial trial period (i.e., time assessment for non-response). Furthermore, the clinical advisor to the ERG noted that RCT data on treatment persistence does not usually generalise directly to clinical practice. Clinical opinion also suggests that there is no biological mechanism that supports the low rate of discontinuation rate for guselkumab when compared with other bDMARDs. The ERG considers that the treatment-specific discontinuation rates used in the model relies on potentially biased evidence given the lack of information on treatment stopping rules in the extension trials and the potential for double counting of non-responders in the initial treatment period.

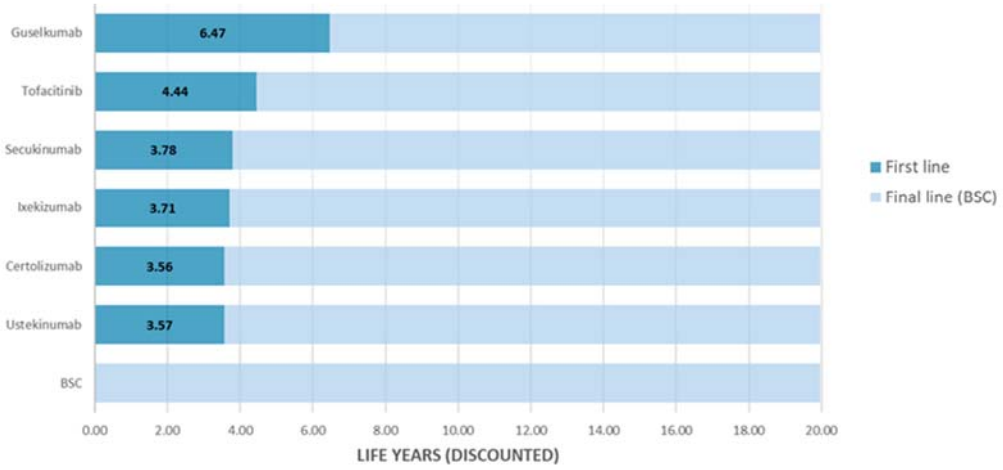
item 4. The use of treatment-specific discontinuation rates in the maintenance period in the model is a key driver of cost-effectiveness but is informed by a limited and potentially biased evidence base.

The ERG is concerned that the treatment-specific estimates of annual discontinuation rates applied in the model may not reflect the use of bDMARDs in UK clinical practice, and deviates from the approach used to assess the comparators in previous TAs. Importantly, the treatment-specific discontinuation rates have an impact on time on treatment, such that once a patient has failed to respond to treatment they are either moved to the next line of therapy or BSC. For the biologic-experienced subpopulation, a second line of active therapy is not modelled and therefore patients are moved directly to BSC. The impact of the annual discontinuation rate on time on treatment for the biologic-experienced subpopulation is illustrated in Figure 7, where discounted life years (LYG) gained for each comparator is presented for (A) treatment-specific annual discontinuation rates as used in the CS, and (B) the same annual discontinuation rate of 16.5% for all active therapies (as used in previous TAs). This figure shows that the differential discontinuation rate is effectively delaying time to BSC, which has the lowest QALY benefits and highest disease-related costs. As a consequence, treatments with a lower discontinuation rate will exhibit greater QALY benefits and lower costs by delaying the time to BSC. The difference in LYG between the treatments at first line of therapy is "erased" once the same discontinuation rate is applied across all treatments, as shown in Figure 7 (B). The same effect is shown in Figure 8 for the biologic-naïve subpopulation. In this population, there is a second-line of active therapy but the discontinuation rate is still delaying time to BSC. These findings are important in light of the fact that the modelled treatment sequences do not reflect the full range of treatment sequences seen in UK clinical practice. Importantly, treatment with multiple lines of active therapy is a valid treatment strategy. This means that in clinical practice patients are unlikely to be placed on BSC at second and third line once they discontinue treatment because other alternative treatment options are available. Therefore, if the modelled treatment duration is long enough by incorporating multiple lines of treatment, the impact of the treatment-specific discontinuation rate is less important.

> ***item 5. Treatment-specific discontinuation rates should only be modelled when the appropriate range of treatment sequences are considered that reflect the full duration of disease.***

Figure 7 Time on treatment in the biologic-experienced population for (A) treatment-specific annual discontinuation rates and (B) same annual treatment discontinuation rate of 16.5% for all active therapies.

(A) Treatment-specific annual discontinuation rate



(B) Same annual discontinuation rate of 16.5% for all treatments

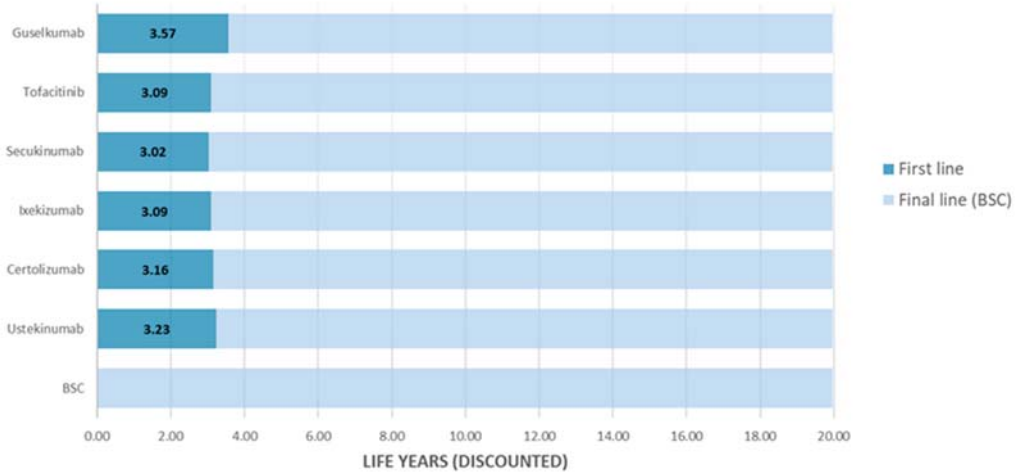
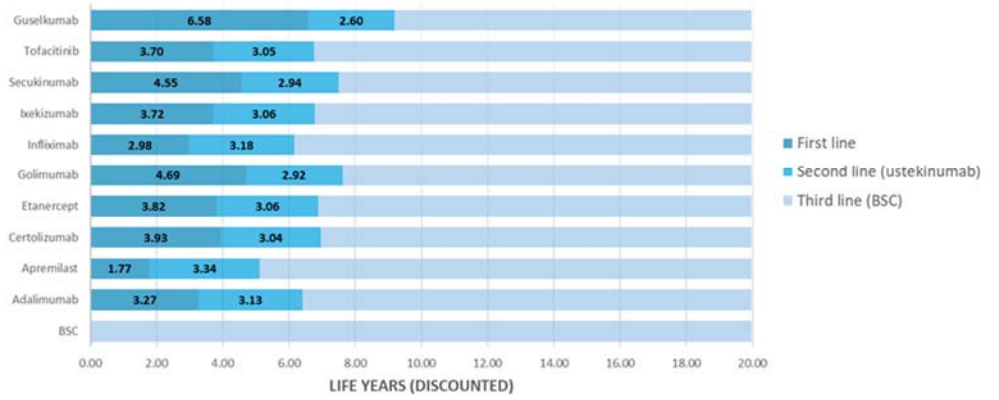
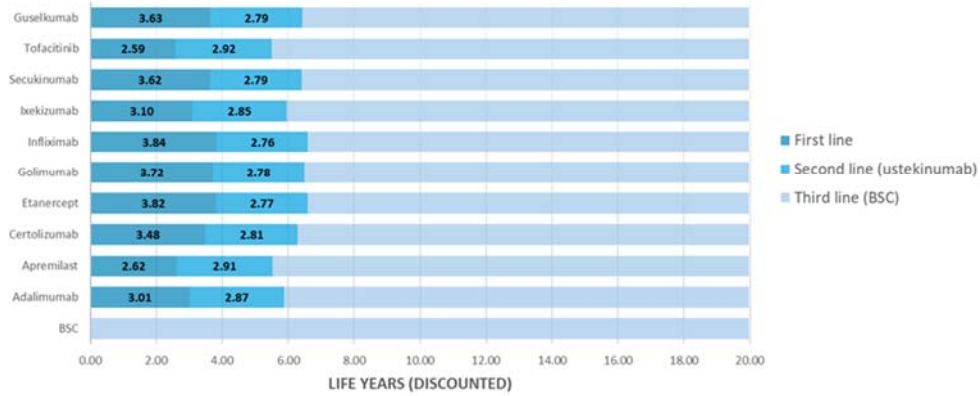


Figure 8 Time on treatment in the biologic-naïve population for (A) treatment-specific annual discontinuation rates and (B) same annual treatment discontinuation rate of 16.5% for all active therapies

(A) Treatment-specific annual discontinuation rate



(B) Same annual discontinuation rate of 16.5% for all treatments



4.2.7 *Perspective, time horizon and discounting*

The cost-effectiveness analysis takes a NHS and Personal Social Services perspective with costs and health outcomes discounted at an annual rate of 3.5%, in line with the NICE reference case.[54] The model adopts a lifetime time horizon, tracking patients for 40 years (model entrance at 47 years of age). Sensitivity analysis is conducted on the inclusion of costs related to lost productivity.

Points for critique

The ERG considers the approach used by the company as appropriate. Lost productivity costs are excluded from the company's base case results and only included in a sensitivity analysis as an exploratory assumption.

4.2.8 *Treatment effectiveness and extrapolation*

There are four main elements of treatment effectiveness included in the company's model:

- Initial treatment response, defined on achieving a PsARC response (base-case analysis);
- PASI response (PASI <50, PASI 50-74, PASI 75-89, PASI 90-99, and PASI 100);
- Change from baseline in HAQ-DI score, which was conditional on PsARC response (i.e., PsARC responders had different scores from non-responders);

PsARC response, PASI and HAQ-DI scores were informed by separate NMAs, as described in Section 3.4. The company also used separate NMAs for biologic-naïve and biologic-experienced (subpopulations 2 and 3, respectively), while evidence informing the TNFi contraindicated (subpopulation 4) was based on the biologic-naïve NMA. For each of the NMAs, the company performed unadjusted analyses (without any covariate adjustments) and meta-regression adjusted for baseline risk (placebo response), as well as random and fixed effect models. To inform the cost-effectiveness analysis, the company uses PsARC and PASI response estimates from the baseline risk-adjusted random effects model, with the exception of PsARC response in the biologic-experienced subpopulation, which was based on the unadjusted fixed effect model, as this had better model fit. HAQ-DI conditional on PsARC response was evaluated using unadjusted models.

In the CS, efficacy data for guselkumab was based on the Q8W dosing schedule and on treatment response at week 24, in anticipation of a stopping rule at week 24 in the SmPC.

Points for critique

The ERG has two main concerns with the efficacy evidence informing the cost-effectiveness model. The first is that the ERG considers that the cost-effectiveness of guselkumab using earlier timepoints for discontinuation, such as 16 weeks (the stopping rule recommended for guselkumab in plaque

psoriasis), should be analysed. The ERG considers there to be little evidence of a specific time-to-response effect to suggest that guselkumab warrants a 24-week stopping rule in PsA. Furthermore, at the time of writing this report, guselkumab has not yet been granted a marketing authorisation for the treatment of PsA. Therefore, the ERG cannot ascertain if the marketing authorisation defines a specific 24-week stopping rule. In response to the ERG's points for clarification, the company have strongly advised that they fully expect a 24-week stopping rule from the EMA and, therefore, any sensitivity analyses using a 16 week time point for responder assessment should only be exploratory in nature. The ERG considers that the cost-effectiveness of a 16-week assessment point for guselkumab should be evaluated, as the evidence suggests that extending treatment for non-responders for a further 8 weeks does not meaningfully increase PsARC response (i.e. the proportion of responders at 16 weeks is very similar to the proportion of responders at 24 weeks in the DISCOVER trials). Additionally, the DISCOVER trials informing the effectiveness of guselkumab permitted an early escape for patients at week 16 and, the ERG has concerns that the imputation conducted for the 24 weeks' data may introduce bias.

The company's updated NMAs for the three effectiveness criteria with guselkumab assessed at 16 weeks (submitted in response to A16, PfC) show when compared to the 24 weeks results:

- A small reduction in PsARC response (e.g., 66.7% vs. 67.5% in the biologic-naïve subpopulation);
- Minor changes in HAQ-DI conditional on PsARC response (e.g., -0.340 vs. -0.360 for responders, and -0.134 vs -0.139 for non-responders);
- A considerable loss of PASI response (see Table 23 and Table 26 for 16 and 24 weeks response rates in the biologic-naïve subpopulation, respectively).

Table 23 Summary of PASI response rates used in the 16 weeks model

	Biologic-naïve			
	Baseline risk-adjusted RE model			
Treatment	**PASI 50**	**PASI 75**	**PASI 90**	**PASI 100**
BSC	15.0%	5.8%	1.9%	0.5%
Etanercept	41.4%	23.2%	11.2%	4.5%
Apremilast	39.2%	21.0%	9.5%	3.5%
Golimumab	56.0%	35.2%	18.9%	8.4%
Adalimumab	61.9%	41.0%	23.3%	11.0%
Certolizumab Mixed[c]	63.0%	42.4%	24.6%	11.9%
Tofacitinib	62.9%	42.3%	24.6%	11.9%
Ustekinumab Mixed[a]	76.3%	57.6%	37.9%	21.1%
Infliximab	76.8%	58.2%	38.5%	21.5%
Secukinumab Mixed[d]	80.9%	64.0%	44.8%	26.9%
Ixekizumab Mixed[b]	85.9%	71.0%	52.3%	33.2%

Guselkumab Q8W	85.0%	69.6%	50.6%	31.7%

[a] Assumes a mix of two doses: 80% 45mg and 20% 90mg; [b] Assumes a mix of two dosing schedules: 69% Q4W and 31% Q2W; [c] Assumes a mix of two dosing schedules: 50% 200 mg Q2W and 50% 400 mg Q4W; [d] Assumes a mix of two doses: 69% 150 mg and 31% 300 mg; RE, random effects.

> ***item 6.*** Thus, the evidence synthesis results suggest that while response for guselkumab in terms of arthritis symptoms is mostly achieved by 16 weeks, psoriasis symptoms continue to improve at least until 24 weeks. The ERG notes, however, that there remains uncertainty as to whether this improvement in PASI response from 16 to 24 weeks for patients treated with guselkumab is not confounded (and to what extent) by the bias potentially introduced by allowing "early escape" in the DISCOVER trials (see Section 3.2.1, 3.2.2.2, and 3.2.2.3)***The assessment time point for response to treatment for guselkumab is based on a stopping rule at 24 weeks, in anticipation of its expected marketing authorisation for PsA, but it is unclear why a 16-week assessment time point should not also be considered in decision making.***

The second main concern with the efficacy evidence informing the cost-effectiveness model is the exclusion of analyses based on the placebo-response unadjusted models for PsARC and PASI (with the exception of PsARC response in the biologic-experienced population, where the company used the unadjusted model because it had better model fit). Although the company indicated that the adjusted models provided the best statistical model fit, a thorough exploration of the potential impact of placebo-response adjustment and the need for it was not provided in the CS. The summary statistics presented in the CS suggest that both models have comparable goodness of fit (the difference in DICs is below 5 units).

As discussed in Section 3.4.2.1, the unadjusted model results closely reflect the observed relative effectiveness in the trials. The relationships defined in baseline risk adjusted models determine that treatments that have been trialled under a high placebo-response rate (such as guselkumab) would have shown a higher treatment effect if the placebo-response rate had been lower, and vice versa. This means that the relative effectiveness of treatments trialled under a higher placebo response (such as guselkumab) are adjusted upwards, while the relative effectiveness of those trialled under a lower placebo response are adjusted downwards (see Figure 5 and Figure 9).

TA445 and TA543 considered both placebo-response adjusted and unadjusted models. As in the current appraisal, inferences based on these models showed a statistically significant effect for the placebo-response adjustment coefficient; however, both the Assessment Group and ERG of the respective appraisals, highlighted that, without a clear rationale for the placebo effect, the results must be interpreted with caution, especially in relation to trying to distinguish (and rank) treatments that

achieve fairly similar response rates (in this case, those that present more central, 'average', estimates). Both TA445 and TA543 also considered models assuming exchangeable of class effects, which were not considered in the CS.

item 7. The company does not present cost-effectiveness results for baseline unadjusted NMA models, which represents an area of uncertainty.

4.2.8.1 Probability of response to treatment

Initial response to treatment was used in the model to dichotomise patients into responders and non-responders at the end of the trial period. This was based on PsARC response in the base case analysis, with two alternative response definitions used in sensitivity analyses: i) joint PsARC and PASI 75 response, and ii) ARC response. It was assumed that the PsARC response achieved at the end of the assessment period was maintained throughout the duration of treatment for each therapy and once patients entered the maintenance period it remained constant. Table 24 provides a summary of the PsARC response probabilities used in the model.

Table 24 Summary of the PsARC response rates used in the model

	Biologic-naïve	Biologic-experienced
Intervention	Baseline risk-adjusted RE model	Unadjusted FE model
BSC	31.3%	30.4%
Apremilast	48.9%	49.5%
Tofacitinib	49.8%	60.4%
Ustekinumab Mixed[a]	56.5%	58.9%
Adalimumab	58.8%	-
Ixekizumab Mixed[b]	59.2%	59.0%
Guselkumab Q8W	67.5%	66.3%
Certolizumab Mixed[c]	68.7%	61.9%
Secukinumab[d]	70.3%	57.6%
Golimumab	73.7%	-
Etanercept	76.0%	-
Infliximab	76.3%	-

[a] Assumes a mix of two doses: 80% 45mg and 20% 90mg; [b] Assumes a mix of two dosing schedules: 69% Q4W and 31% Q2W; [c] Assumes a mix of two dosing schedules: 50% 200 mg Q2W and 50% 400 mg Q4W; [d] Assumes a mix of two doses for the biologic-naïve: 69% 150 mg and 31% 300 mg; and, for biologic-experienced 300mg; FE, fixed effects; RE, random effects.

Mean changes in HAQ-DI scores from baseline were conditional on PsARC response and were used in the model to inform utility and disease-related costs. The improvement in HAQ-DI score was applied instantaneously at the start of the trial period (alternative assumptions where explored in sensitivity analyses) and the HAQ-DI score for PsARC responders was maintained throughout the duration of the maintenance period for all treatments, except for BSC where it deteriorated over time according to the natural history of progression (see Section B.3.2.3 of CS). Table 25 provides a summary of the HAQ-DI scores conditional on PsARC response used in the model.

Table 25 Summary of the HAQ-DI scores conditional on PsARC response used in the model

	Biologic-naïve		Biologic-experienced	
Intervention	PsARC responders	PsARC non-responders	PsARC responders	PsARC non-responders
	Unadjusted RE model		Unadjusted FE model	
BSC	-0.301	0.038	-0.275	0.010
Certolizumab Mixed[c]	-0.348	-0.009	-0.349	-0.016
Golimumab	-0.348	-0.009	NA	NA
Guselkumab Q8W	-0.360	-0.134	-0.496	-0.213
Secukinumab[b]	-0.360	-0.134	-0.496	-0.213
Ixekizumab[b]	-0.360	-0.134	-0.496	-0.213
Apremilast	-0.394	-0.054	-0.367	-0.083
Tofacitinib[d]	-0.394	-0.054	-0.367	-0.083
Ustekinumab Mixed[a]	-0.479	-0.091	-0.272	-0.118
Adalimumab	-0.514	-0.094	NA	NA
Etanercept	-0.679	-0.156	NA	NA
Infliximab	-0.712	-0.159	NA	NA

[a] Assumes a mix of two doses: 80% 45mg and 20% 90mg; b Assumed the same as for guselkumab Q8W in the absence of data; d Assumed the same as for apremilast in the absence of data FE, fixed effects; RE, random effects.

Note that, because the CS did not find evidence to inform HAQ-DI conditional on PsARC for secukinumab and ixekizumab, the estimates were assumed the same as guselkumab. Similarly, the CS identified no data for tofacitinib, and assumed the same scores as for apremilast (under the argument that these two drugs have the same mode of administration).

An average PASI score was considered at each model cycle; this was calculated by applying a percentage improvement to baseline PASI. The percentage improvement for each PASI response category (PASI <50, PASI 50-74, PASI 75-89, PASI 90-99, and PASI 100) was weighted by the proportions of patients achieving each category of PASI response (as predicted by the NMA). The

improvement in PASI response from baseline score was applied instantaneously at the start of the trial period and maintained throughout the duration of the maintenance period at a constant rate for all treatments. Table 26 provides a summary of the PASI response rates used in the model.

Table 26 Summary of PASI response rates used in the model

	Biologic-naïve			
	Baseline risk-adjusted RE model			
Treatment	**PASI 50**	**PASI 75**	**PASI 90**	**PASI 100**
BSC	16.0%	6.1%	2.0%	0.6%
Etanercept	42.4%	22.9%	10.7%	4.1%
Apremilast	41.4%	22.1%	10.3%	3.9%
Golimumab	58.6%	36.9%	20.2%	9.2%
Adalimumab	64.2%	42.5%	24.6%	11.9%
Certolizumab Mixed[c]	65.2%	43.7%	25.5%	12.5%
Tofacitinib	65.5%	44.0%	25.8%	12.6%
Ustekinumab Mixed[a]	78.5%	59.4%	39.6%	22.5%
Infliximab	79.3%	60.5%	40.7%	23.4%
Secukinumab Mixed[d]	83.3%	66.1%	46.7%	28.3%
Ixekizumab Mixed[b]	87.8%	73.0%	54.5%	35.3%
Guselkumab Q8W	91.9%	80.2%	63.6%	44.3%
	Biologic-experienced			
	Baseline risk-adjusted RE model			
	PASI 50	**PASI 75**	**PASI 90**	**PASI 100**
BSC	13.4%	4.4%	1.1%	0.2%
Apremilast	34.7%	16.0%	5.9%	1.7%
Certolizumab Mixed[c]	56.3%	33.4%	15.9%	6.1%
Tofacitinib	31.6%	14.0%	5.0%	1.4%
Ustekinumab Mixed[a]	61.5%	38.7%	19.7%	8.1%
Secukinumab 300mg	74.0%	52.1%	30.7%	14.7%
Ixekizumab Mixed[b]	71.8%	49.6%	28.5%	13.3%
Guselkumab Q8W	77.3%	56.1%	34.5%	17.3%

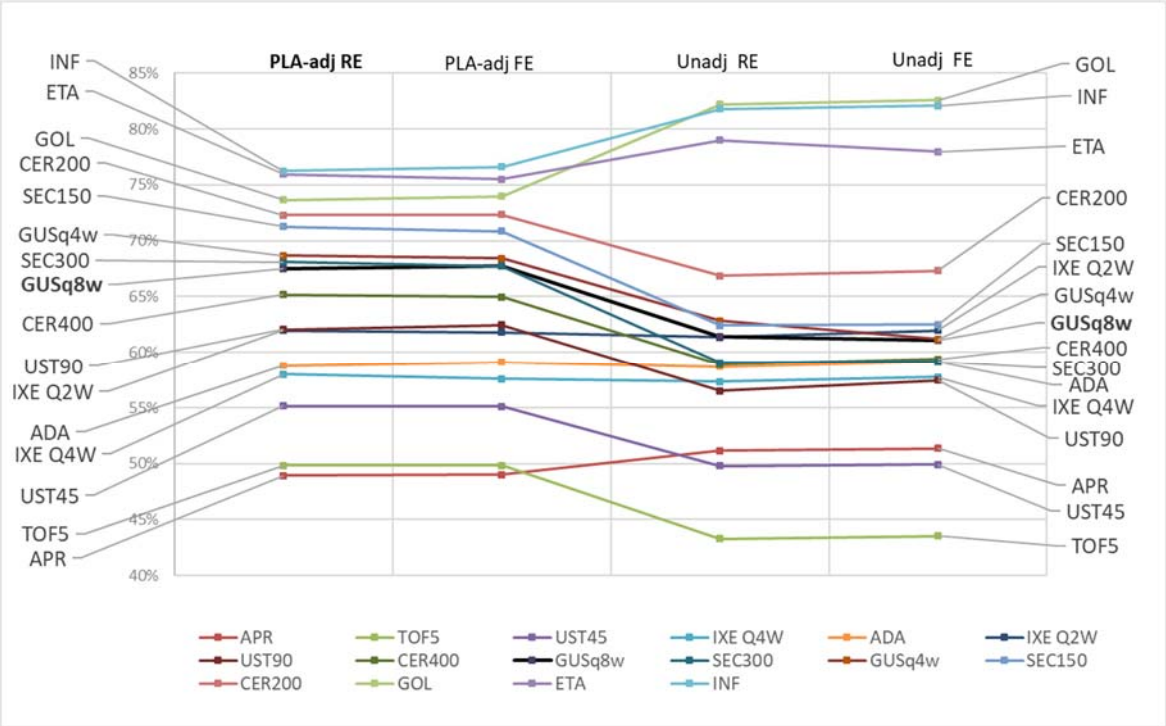
[a] Assumes a mix of two doses: 80% 45mg and 20% 90mg; [b] Assumes a mix of two dosing schedules: 69% Q4W and 31% Q2W; [c] Assumes a mix of two dosing schedules: 50% 200 mg Q2W and 50% 400 mg Q4W; [d] Assumes a mix of two doses: 69% 150 mg and 31% 300 mg; FE, fixed effects; RE, random effects.

Points for critique

The assumptions and approaches used to apply the treatment response rates within the model are generally consistent with the models used in previous TAs (TA445, TA537 and TA543). However, as noted above, the ERG's main concerns relate to the absence of cost-effectiveness results reported for

the placebo unadjusted response rates and the time point of assessment for guselkumab. The placebo unadjusted model could have important implications for both the absolute response rates and ranking of treatments, which in turn could affect the results of the fully incremental cost-effectiveness analysis. Figure 9 shows the treatment-specific probability of response (y-axis) estimated by the alternative models implemented by the CS (x-axis): placebo-response adjusted random-effects (PLA-adj RE used in the base case), placebo-response adjusted fixed-effects (PLA-adj FE), unadjusted random-effects (Unadj RE) and unadjusted fixed effects (Unadj FE). Fixed and random effects models lead to similar point estimates. However, both adjusted and unadjusted models identify three main groups of treatments ranked in terms of the level of PsARC response: those with higher probability of response (golimumab, infliximab and etanercept), those with central values (guselkumab, ixekizumab, secukinumab, certolizumab, and adalimumab), and those with lower probability of response (apremilast, ustekinumab and tofacitinib). The range of response probabilities across the different treatments reduces with placebo response-adjustment, with the relative effect of treatments in the highest and lowest value groups approximating the central value group. Importantly, placebo response-adjustment significantly changes the range of values of probability of response in the central group and the ranking between treatments. Given the uncertainty in the appropriateness of the statistical adjustment, and the fact that the confidence intervals for all treatments in this range overlap (uncertainty not presented in Figure 9), the ERG believes that the evidence available cannot substantiate any differences between the treatments in this central range.

Figure 9 Biologic-naïve - Comparison of placebo-response unadjusted models with baseline risk-adjusted models for PsARC response



PLA-adj RE, placebo response-adjusted random effects model; PLA-adj FE, placebo response-adjusted fixed effect model; Unadj RE, unadjusted random effects model; Unadj FE, unadjusted fixed effect model; ADA, adalimumab; APR, apremilast; CER, certolizumab pegol; ETA, etanercept; GOL, golimumab; GUS, guselkumab; INF, infliximab; Ixekizumab; NR, not reported; Q2W, every 2 weeks; Q4W, every 4 weeks; Q8W, every 8 weeks; SEC, secukinumab; TOF, tofacitinib; UST, ustekinumab.

A similar exploration for PASI response rates between the placebo unadjusted and placebo response-adjusted models in the biologic-naïve population demonstrated a significant difference in PASI 75 response rates for golimumab (58.3% in the unadjusted vs. 34.8% in the baseline risk-adjusted random effects model), infliximab (84.0% vs. 57.5% in the unadjusted and adjusted models, respectively) and guselkumab (59.8% vs. 77.2% in the unadjusted and adjusted models, respectively), while differences for the other treatments were smaller in magnitude. The results for the unadjusted models for each of the two response metrics (PsARC, and PASI) are presented in Appendix 9.1 for each of the subpopulations. Given the uncertainty surrounding the most appropriate model to inform response rates used in the model, the ERG assessed the impact on cost-effectiveness of considering estimates from the unadjusted models for each subpopulation.

For the biologic-naïve and TNFi contraindicated subpopulations, patients receiving a subsequent line of active therapy were assumed to have the same response probabilities as first line treatment of biologic-experienced patients. For the biologic-experienced subpopulation the number of prior bDMARDs failed was not modelled and no effect degradation was applied to subsequent lines of therapy. The company indicated that while the treatment degradation from biologic naïve using one mechanism of action to biologic experienced using a different mechanism of action is well characterised, there is a paucity of evidence on the treatment degradation caused by using the same mechanism of action in a subsequent line of therapy. For treatments with a lower PsARC response rate and, therefore, a higher likelihood of switching on to an alternative therapy before BSC, the assumption of no effect degradation may overestimate cost-effectiveness.

4.2.9 *Adverse events*

The company's model includes treatment specific adverse events with associated costs and QALY loss. Patients in the model are at risk of adverse event at every cycle in the model, with the cycle probability of suffering an adverse event being derived from the NMA described by the company in Section B.2.9.3.6 of the CS and in Appendix D (D.1.5.7). The NMA estimated the probability of patients experiencing at least one serious adverse event while on treatment for each biologic therapy and placebo.

Points for critique

The models informing previous TAs in PsA have not included adverse events, implicitly assuming that there are no differences in the safety profile of the treatments under comparison that may impact on costs and health outcomes. While the ERG acknowledges that this may not be plausible, the company's approach to capture the impact of differences in the safety profile of each treatment is unlikely to appropriately reflect such differences in the cost-effectiveness estimates. First, there is no evidence that the studies included in the NMA define serious adverse events in a consistent manner (see company's response to PfC, question A19), and therefore, it may not be considered appropriate to pool rates across the studies. Second, even if the definition of the outcome was consistent across studies, the nature of each serious adverse event could result in different costs and HRQoL. The company's approach to modelling adverse events assumes that all serious adverse events have the same impact on HRQoL as a serious infection (see Section 4.2.11) and that the distribution of serious adverse events for the purpose of costing is the same as that in the DISCOVER trials (see Section 4.2.12.6). The ERG considers that it is not reasonable or consistent to assume a disutility estimate associated with a single adverse event (infections) to estimate QALY loss associated with all adverse events, but to consider individual adverse events when estimating costs. The selection of the DISCOVER trials to inform the distribution of adverse events for costing purposes is also questionable, as it implies the same distribution of adverse events for all other treatments.

item 8. The company's approach to include adverse events in the model is unlikely to reflect the safety profile of the different treatments and is not consistent with the assumptions of previous TAs in PsA.

4.2.10 *Mortality*

Patients are subject to a mortality risk at every health state in the model. The probability of death was determined based on UK life tables (2016-18) and an excess mortality estimate for PsA sourced from the literature.

Points for critique

The CS does not describe the approach used to model mortality but further clarification was provided in response to ERG points for clarification (Question B18). The ERG considers the approach used by the company as appropriate. The only notable differ from previous TAs is the application of an excess mortality rate of 1.05, which was reported as 1.36 in TA445 from the same source (Ali et al 2007). The difference is due to the use of an alternative data cut and is not expected to have any impact on cost-effectiveness.

4.2.11 *Health related quality of life*

Utility values used in the model were dependent on mean HAQ-DI and PASI scores at each cycle and estimated separately for each line of therapy. HAQ-DI and PASI scores were mapped onto EQ-5D-3L to obtain utility values using the York algorithm, first defined in TA199. The York algorithm was originally estimated based on a confidential analysis of trial data conducted by Wyeth, who commercialised Enbrel® (etanercept). The York algorithm has been used in previous TAs in PsA. EQ-5D data were also available from the DISCOVER-2 trial and used in a sensitivity analysis. The CS states that the York algorithm was preferred for the base-case analysis to maintain consistency with previous TAs in PsA.

The improvement in PASI score for each model cycle and line of treatment was calculated based on the distribution of patients across the PASI response categories ([0-49], [50-74], [75-89], [90-99], [100]) and weighted by the number of PsARC responders and non-responders. This PASI improvement was added to the baseline PASI score to yield an absolute PASI score. The HAQ-DI score for each cycle and line of treatment was calculated by applying the HAQ-DI change conditional on PsARC response to baseline HAQ-DI score, and weighted by the proportion of responders and non-responders. In the base-case analysis, HAQ-DI and PASI improvement was applied instantaneously at the start of treatment and remained constant for patients in the maintenance period until the patient moves to the next line of therapy. When patients reach the final line of therapy, utility starts to decline over time until the maximum HAQ-DI score of 3 is reached.

The model constrains utility values so that these do not exceed those of the general population, as estimated by Ara and Brazier, 2010, and age and gender distribution in the model. The company states that this adjustment was included to be consistent with TA537.

Unlike previous appraisals in PsA, the model explicitly considers the HRQoL impact of serious adverse events. The CS assumes a single disutility estimate of -0.195, sourced from the literature for serious infections in chronic lymphocytic leukaemia, and a duration of 11.75 days, corresponding to the average duration of a serious infection in the DISCOVER trials. These estimates are combined into a per cycle disutility, which jointly with the cycle probability of a serious adverse event yields a serious adverse event-related disutility per cycle for each therapy.

Points for critique

With the exception of modelling SAEs separately, the approach used by the company to estimate HRQoL is generally consistent with that used in previous TAs in PsA. EQ-5D data were available from the DISCOVER-2 trial but the company chose to use the York algorithm in its base-case analysis to maintain consistency with previous TAs. The two algorithms share the same covariates

(HAQ-DI and PASI) and linear form. The algorithm based on the DISCOVER-2 trial has a smaller coefficient in absolute terms on HAQ-DI (-0.231 compared to -0.298 in the York algorithm) and a lower intercept (0.843 compared to 0.897 in the York algorithm). The coefficient on PASI is almost identical for the two algorithms. The impact of using the DISCOVER-2 algorithm on the cost-effectiveness results was explored in the company's sensitivity analysis, and was generally associated with less favourable ICERs for all interventions compared with BSC. Limited information is provided on the data used to inform the DISCOVER-2 algorithm making it difficult for the ERG to assess the validity of the algorithm. Given that the company has not provided a strong reason to support one algorithm over the other for informing the utility estimates used in the model for all comparator treatments, the ERG supports the use of the York algorithm for consistency with the previous TAs in PsA.

The approach used by the company to adjust the utility values to ensure that they do not exceed the utility values for the general population at a given age for the modelled gender distribution appears appropriate. However, the ERG notes that the proportion of model cycles where general population utility exceeded algorithm derived utility at first line of therapy varied considerable across treatments ranging from 0% (tofacitinib, secukinumab, ixekizumab, golimumab, certolizumab, apremilast) to over 20% (etanercept, infliximab) in the biologic-naïve subpopulation.

As mentioned in Section 4.2.9 the inclusion of SAE in the company's base-case analysis is a departure from the assumptions of previous TAs in PsA. Importantly, it relies on strong unduly justified assumptions, such as the QALY loss estimated for a serious infection capturing the impact of adverse events on HRQoL of other adverse events across all treatments.

4.2.12 *Resource use and costs*

The CS considers i) drug acquisition costs, ii) drug administration costs, iii), routine patient monitoring costs iv) disease related costs and v) adverse event costs. Disease related costs include two components: costs associated i) with HAQ-DI and ii) with psoriasis. Unit costs are informed by national published sources[55-57] and previous NICE guidance,[26, 31] uprated to 2018/19 prices where appropriate and discounted at an annual rate of 3.5%.

Points for critique

4.2.12.1 Drug acquisition costs

The company details unit costs and costs per maintenance cycle in Table 67 and 68 of the CS, respectively. Below the ERG critiques issues identified in the company's approach to the estimation and presentation of drug acquisition costs.

The company assumed the use of biosimilars whenever these were commercially available. However, the rationale for the choice of unit cost when more than one biosimilar drug is available is unclear. The unit cost assumed for adalimumab (2 x 40 mg) corresponds to that of Amgevita® (£633.60), but the cost of adalimumab biosimilars ranges between £616.25 and £633.86 for Hyrimoz® and Imraldi®, respectively.[55] The most costly biosimilar, Benepali® (£656.00) was selected for etanercept, despite the existence of a cheaper alternative (Erelzi®, £643.50).[55] In contrast, the cheapest biosimilar was chosen for infliximab (Flixabi®, £337.00).[55] Given the small cost differences between the unit costs of biosimilars for each bDMARD, this was not considered a priority item, and, therefore, does not consider any corrections for this in the ERG base-case.

Although previous TAs have not attributed a drug cost to BSC, the CS considers an acquisition cost for BSC. The company considers this necessary as their choice of algorithm to estimate HAQ-DI related costs (see CS, p147) does not include medication costs, whereas the algorithm used in previous TAs in PsA did. The acquisition cost of BSC corresponds to the cost of methotrexate 20mg/week (£1.04/week), i.e. the highest recommended dose in rheumatoid arthritis (recommended dose in PsA is 7.5mg/week).[55] The company states that this represents a conservative assumption on the cost of BSC, as opposed to using the cost of leflunomide 20 mg/week (£1.21/week[55]) (B10, response to points for clarification). The unit cost is sourced from the BNF and corresponds to the NHS indicative price of £52.01 for 100 units of methotrexate 10mg (A A H pharmaceuticals).[55] The ERG was not able to validate this cost (current version of BNF attributes a price of £53.33 to this drug), but notes that this is not the least costly formulation. The cost per week with the least costly alternative (Maxtrex® 10mg x 100) would be £0.90.[55] The ERG also notes that the cost per week of leflunomide 20 mg, when assuming eMIT costs (not available for methotrexate 10 mg) is £0.27, and so it is debatable whether the company used a conservative assumption to calculating the costs of BSC. However, these costs are generally low and unlikely to have an impact on the estimates of cost-effectiveness. Therefore, no corrections for this element of cost are considered in the ERG base-case.

The drug cost per maintenance cycle for each biologic in the model is based on the number of doses administered in a maintenance year, with the annual number of doses divided equally over each cycle. The interpretation of the cost per maintenance cycle reported in Table 68 of the CS is complicated by the need to combine it with the average number of doses per maintenance cycle to obtain the actual cost accrued at each cycle in the model. The ERG extracted the actual costs per cycle in both trial and maintenance period from the model to facilitate the comparison of drug costs. These costs are presented on Table 27. The estimates presented incorporate the confidential patient access scheme (PAS) price for guselkumab ([redacted]) and for certolizumab (free of charge for the 12 weeks response assessment period). For comparators that were modelled as

combination of doses or regimens, the “mixed” comparator cost is presented, i.e., a weighted average of the costs for each dose or regimen.

Table 27 Drug acquisition costs applied in the model

Drug cost per cycle	Response assessment period						Maintenance period
	Week 0-3	Week 4-7	Week 8-11	Week 12-15	Week 16-19	Week 20-23	
Guselkumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	£ 4.16	£ 4.16	£ 4.16	-	-	-	£ 4.16
Adalimumab	£ 633.60	£ 633.60	£ 633.60	-	-	-	£ 633.60
Apremilast	£ 540.18	£ 550.00	£ 550.00	£ 550.00	-	-	£ 550.00
Certolizumab*	£ 0.00	£ 0.00	£ 0.00	-	-	-	£ 715.00
Etanercept	£ 656.00	£ 656.00	£ 656.00	-	-	-	£ 656.00
Golimumab SC	£ 762.97	£ 762.97	£ 762.97	-	-	-	£ 701.87
Infliximab	£ 3,770.00	£ 1,885.00	£ 0.00	-	-	-	£ 942.50
Ixekizumab*	£ 2,598.75	£ 1,473.75	£ 1,473.75	£ 1,125.00	-	-	£ 1,125.00
Secukinumab*	£ 3,193.20	£ 798.30	£ 798.30	£ 798.30	-	-	£ 734.37
Secukinumab 300mg**	£4,875.12	£1,218.78	£1,218.78	£1,218.78	-	-	£1,125.00
Tofacitinib	£ 690.03	£ 690.03	£ 690.03	-	-	-	£ 690.03
Ustekinumab	£ 2,147.00	£ 2,147.00	£ 0.00	£ 0.00	£ 2,147.00	£ 0.00	£ 715.67

*Mixed dose comparator; **For biologic experienced population

The CS states that no vial wastage was assumed, but wastage is assumed for infliximab. Other biologics have a fixed dose and will use the totality of the vials, so there will be no wastage. The assumption of wastage for infliximab seems appropriate, as the dose of this drug is weight dependent (5mg/Kg) and the remaining drug in the vial is likely to be discarded. Thus, this is only a matter of accurate report, and does not require any corrections.

4.2.12.2 Drug administration

item 9. The company assumes that only infliximab has associated drug administration costs (£241.06),[56] and does not justify the exclusion of administration costs for other drugs. Previous appraisals have included a one-off cost of one hour of a specialist nurse time at the hospital applied at the first cycle of treatment for biologics administered subcutaneously. This reflects the resource use required to teach the patient how to self-administer the drug. The ERG base-case considers the cost of teaching the patient how to self-administer each new subcutaneous bDMARDs in the treatment sequence. However, the ERG notes that this cost

may only be incurred once if the injection devices are sufficiently similar across subcutaneous treatments. ***Administration costs should be included for drugs administered subcutaneously.***

4.2.12.3 Routine patient monitoring

Routine patient monitoring costs in the model include specialist visits and laboratory tests. The CS considers more intensive resource use in the trial period compared to the maintenance period, as well as less intensive resource use overall for patients on oral therapies compared to therapies administered via other routes. The resource use associated with routine patient monitoring in the CS are summarised in Table 28, alongside those for previous NICE appraisals. The company states that the source the frequency of testing for SC and IV therapies was sourced from TA537 and TA433 for oral therapies, while BSC was assumed to have the same resource use as intravenous and subcutaneous therapies. The CS does not state why TA537 and TA433 were considered a more appropriate source of resource use data.

Table 28 Routine patient monitoring resource use

Resource use	TA445/TA199/TA543		TA537				CS			
	Trial period	Maintenance period*	Trial period		Maintenance period*		Trial period		Maintenance period*	
Administration route	Any	Any	SC/IV	Oral	SC/IV	Oral	SC/IV	Oral	SC/IV	Oral
Specialist visit	1	0	2	2	0	1	2	2	2	1
Full blood count	2	2	2	2	2	0	2	1	2	0
Liver function test	2	2	2	2	2	0	2	1	2	0
Urea and electrolytes	2	2	2	2	2	0	2	1	2	0
ESR	2	2	2	2	2	0	2	1	2	0
Chest X-Ray	1	0	1	1	0	0	1	1	0	0
Tuberculosis Heaf test	1	0	1	1	0	0	1	0	0	0
ANA test	1	0	1	1	0	0	1	1	0	0
Double strand DNA test	1	0	1	1	0	0	1	1	0	0

*Resource use per annum; ANA, anti-nuclear antibody; ESR, erythrocyte sedimentation rate; SC, subcutaneous

The ERG notes that in TA433 the Committee decided to consider the same resource use for patient monitoring for oral therapies as for other therapies, even if apremilast's manufacturer had originally assumed less intensive use of resources (full details not available from published documents). TA537 considered oral therapies to be as resource use intensive as other therapies in the trial period, but only require one specialist visit per annum in the maintenance period. TA543, which took place at the same time as TA537, assumed no differences in resource use of patient monitoring across therapies. The resource use assumptions in TA537 seem to be a departure from previous appraisals, which is not justified in the appraisal documentation. In the absence of a clear rationale for these assumptions, it would be more consistent to apply the same assumptions as TA445 and TA543, and assume the same resource use associated with patient monitoring across all treatments.

item 10.Monitoring resource use should be the same across treatments to be consistent with previous TAs in PsA.

4.2.12.4 Arthritis costs – HAQ-DI costs

The company's base-case applies the algorithm developed by McHugh and colleagues to estimate annual health care costs associated with management of arthritis (as modelled by the relationship between costs and HAQ-DI score). [58]Previous appraisals have consistently used the algorithm developed by Bansback et al., 2006,[21] based on data collected by Kobelt et al., 2002,[44] to estimate these costs. The CS justifies the use of the McHugh et al., 2019, algorithm based on the fact that this was recently estimated in a PsA population, whereas the Kobelt et al., 2002, study was conducted in rheumatoid arthritis. The Kobelt et al., 2002, algorithm is applied on a sensitivity analysis. The model allows the use of a third algorithm, Poole et al., 2010, [59]but does not apply it in any analyses, due to limitations of the study described in TA537 (which were initially identified in TA445). The three algorithms are compared on Table 29, which also reports patient characteristics, categories of cost, and differences in terms of model implementation.

Table 29 Comparison of HAQ-DI costs algorithms

	McHugh et al., 2019	**Kobelt et al., 2002**	**Poole et al., 2010**	
		UK cohort	BSRBR register	THIN dataset
Disease area	PsA	Rheumatoid arthritis	PsA	
n	101	916	296	2526
Age, years (mean(SD))	57.83 (10.66)	54.8 (13.6)	46.7 (10.8)	55.5 (14.9)
% Female	57	66.6	52	51
HAQ-DI score (mean(SD))	0.84 (0.75)	1.11 (0.7)	1.786 (0.637)	-

Disease duration, years (mean/median (SD/IQR))	18.23 (11.26)	-	11 (6–18)	-
Annual costs (uprated to 2019 prices)	£1196 + £580*HAQ-DI	£1601+ £483*HAQ-DI	Exp(3.537 + 2.048*HAQ-DI + 0.026*Age – 0.012*HAQ-DI*Age)	
Categories of costs	A&E visits Primary care consultation Secondary care consultation Admitted care	Hospitalisations Surgical interventions Outpatient visits Community services Rheumatoid arthritis medication (13-15% of total costs)	A&E visits Primary care contacts Outpatient appointments Acute hospital inpatient care Investigations Prescriptions costs (38% of total costs)	
Proportion of algorithm estimated costs applied to patients treated with bDMARDs	100%	85%	62%	
Additional BSC drug costs applied in the model	Yes, £4.16 per cycle	No	No	

BSRBS, British Society of Rheumatology Biologics Register; IQR, inter-quartile range; THIN, The Health Improvement Network

The ERG notes that the CS preferred algorithm is based on a study with small sample size. Small sample sizes can be problematic when handling cost data, which is usually skewed and highly variable. The McHugh et al., 2019, algorithm is, thus, potentially more affected by uncertainty and vulnerable to bias than the Kobelt et al., 2002, algorithm. Furthermore, few patients in McHugh et al., 2019, had HAQ-DI≥2 (n=33), and the mean HAQ-DI in this study population is also lower than that of the patient population in the model (0.84 vs 1.23 and 1.38 in the biologic naïve and experienced population, respectively). The mean HAQ-DI in Kobelt et al., 2002, (1.11) is closer to that of the model population.

The Kobelt et al., 2002, algorithm is not without issues, namely the fact that it was estimated in a different (if related) disease area and being dated. Notwithstanding the issues affecting this algorithm, it would be more appropriate to use this in base-case analysis to keep consistency with previous TAs, whilst using- McHugh et al., 2019, in a scenario analysis.

item 11. The Kobelt et al., 2002, algorithm should be used to estimate arthritis related costs in the base-case analysis to ensure consistency with previous TAs in PsA.

4.2.12.5 Psoriasis related costs

The CS also considers costs associated with psoriasis, and which depend on i) achievement of PASI75 response and ii) psoriasis severity. Table 71 (CS, p149) describes annual disease costs associated with psoriasis which were sourced from TA543 and uprated to 2019 prices. The psoriasis related costs were applied consistently with previous appraisals.

The ERG notes, however, that in a scenario where a gradual (linear) increase is assumed for PASI (and HAQ-DI) the PASI related costs are not consistent with this assumption, as the costs are estimated based on the proportion of PASI75 responders regardless of whether patients are in the trial or the maintenance period. This problem does not arise for HAQ-DI costs because the cost algorithm uses the HAQ-DI score directly as an input, and the assumptions on HAQ-DI benefit are reflected on the average HAQ-DI score. The model structure does not allow applying the PASI costs in a consistent manner under this scenario, as a gradual reduction of PASI scores is not explicitly modelled. This issue is, thus, not amenable to be addressed within the current model structure. However, given the short duration of the trial period and that previous appraisals have considered the assumption of instantaneous PASI improvement acceptable, this is unlikely to have an impact on cost-effectiveness driver estimates. Therefore, the ERG does not address this issue further.

4.2.12.6 Costs of adverse events

The cost of serious adverse events were assumed to correspond to the weighted average of serious adverse events occurring in the DISCOVER trials (£3,312 per adverse event). This cost was assumed for all treatments regardless of the distribution of serious adverse events in their clinical trial data. As discussed in Section 4.2.9, this approach is unlikely to capture any differences between treatments in terms of their safety profiles, and, thus, misrepresent any differential impact on costs.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

All analyses presented in the CS include the confidential PAS discount associated with guselkumab, and include the publicly available PAS for comparators. The ERG updated the company's deterministic base-case results by incorporating the confidential PAS discounts for the comparators (as provided by the companies holding the marketing authorisation for each product) and the confidential framework prices for the biosimilars of [redacted] provided by the Department of Health and Social Care Commercial Medicines Unit (CMU). The ERG's updated analyses of the company's base-case are reported in the confidential PAS appendix.

The company's base-case deterministic and probabilistic cost-effectiveness results are presented in Table 30 and

Table 31, respectively, for the overall subpopulations considered in the CS. The deterministic results for each subpopulation by psoriasis severity subgroup are shown in Table 32. The results refer to the fully incremental cost-effectiveness analyses and only the results of non-dominated treatments are shown (i.e., after exclusion of dominated or extendedly dominated treatments).

Table 30 Company's base-case analysis –deterministic results of the fully incremental cost-effectiveness analysis

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naïve / All TNFi contra-indicated					
BSC	£83,722	5.014	-	-	-
Guselkumab Q8W	[redacted]	8.511	[redacted]	3.497	[redacted]
All biologic-experienced					
BSC	£86,772	4.254	-	-	-
Guselkumab Q8W	[redacted]	6.800	[redacted]	2.546	[redacted]

Table 31 Company's base-case analysis –probabilistic results of the fully incremental cost-effectiveness analysis

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naïve /					
BSC	£ 79,565	5.230	-	-	-
Guselkumab Q8W	[redacted]	8.428	[redacted]	3.198	[redacted]
All biologic-experienced					
BSC	£ 82,492	4.493	-	-	-
Guselkumab Q8W	[redacted]	6.838	[redacted]	2.345	[redacted]
All TNFi contra-indicated					
BSC	£ 79,569	5.229	-	-	-
Guselkumab Q8W	[redacted]	8.428	[redacted]	3.198	[redacted]

Guselkumab was the most effective treatment in all three subpopulations, generating 8.511 QALYs in the biologic-naïve and TNFi contraindicated subpopulations and 6.800 QALYs in the biologic-experienced population. Guselkumab is also more costly than the majority of comparators, but it strictly or extendedly dominates all other treatments in all three subpopulations in the base case analysis. Probabilistic results are similar to deterministic results for all subpopulations with ICERs for guselkumab vs. BSC increasing slightly when joint parameter uncertainty is considered. The differences are not sufficient to change the conclusions on the cost-effectiveness of guselkumab.

Table 32 Company's base-case subgroup analysis - deterministic results of the fully incremental cost-effectiveness analysis

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£59,955	5.406	-	-	-
Etanercept	[redacted]	8.227	[redacted]	2.820	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.546	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Etanercept	[redacted]	8.105	[redacted]	2.840	[redacted]
Guselkumab	[redacted]	8.689	[redacted]	0.583	[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£109,716	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]
bDMARD-experienced – Minimal PsO					
BSC	£61,138	5.431	-	-	-
Guselkumab	[redacted]	7.949	[redacted]	2.518	[redacted]
bDMARD- experienced – Mild to moderate PsO					

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
BSC	£81,014	4.133	-	-	-
Guselkumab	[redacted]	6.599	[redacted]	2.466	[redacted]
bDMARD- experienced – Moderate to severe PsO					
BSC	£113,227	3.493	-	-	-
Guselkumab	[redacted]	6.184	[redacted]	2.691	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£59,955	5.406	-	-	-
Secukinumab	[redacted]	8.053	[redacted]	2.647	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.719	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Guselkumab	[redacted]	8.689	[redacted]	3.423	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£109,716	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]

The company did not comment on the psoriasis subgroup results for each population, although these are presented in Appendix T of the CS for the base-case and sensitivity analysis. Psoriasis severity is an important element of heterogeneity in the cost-effectiveness of biologics for PsA. In the model, the category of psoriasis severity determines:

i. the baseline PASI and HAQ-DI scores and, thus, the magnitude of potential benefit from treatment (see Table 21) – Baseline PASI score increases from the minimal to the moderate to severe psoriasis subgroup, while baseline HAQ-DI score varies in a non-monotonic pattern across subgroups;
ii. the magnitude of disease related costs (see Table 71 of the CS, p149);
iii. the dose of secukinumab and dosing schedule of ixekizumab, which determine their effectiveness, treatment discontinuation rates and drug acquisition costs.

In the minimal psoriasis subgroup, across all subpopulations, all treatments have lower total costs on average compared to results for the corresponding overall population because no psoriasis related costs are incurred. Mean total QALYs are higher due to the lower baseline PASI, which translates into higher baseline HRQoL. However, the treatment benefit derived from PASI response will be smaller in absolute terms due to the lower baseline PASI. For the moderate-to-severe psoriasis subgroup, although mean total costs are higher and mean total QALYs are lower, treatments can yield higher benefits in absolute terms because the baseline PASI score is higher in all subpopulations (and the

baseline HAQ-DI score for biologic-naïve and TNFi-contraindicated) compared to the overall subpopulation scores.

For the company's base-case assumptions, guselkumab appears to be the most cost-effective treatment across subgroups for the majority of analyses. The only exception is for the minimal psoriasis subgroup of the biologic-naïve subpopulation, where the most cost-effective treatment at conventional NICE recommended cost-effectiveness thresholds is etanercept. Etanercept has a greater reduction in total costs ([redacted]) and increase in total QALYs ([redacted] QALYs) compared to the overall population analysis, than guselkumab ([redacted]). The ICER of guselkumab compared to etanercept is [redacted] of the NICE cost-effectiveness threshold range at [redacted] per additional QALY.

Points for critique

The ERG notes that the approach used to estimate the cost-effectiveness results for the overall population may not be appropriate (i.e., weighting inputs to reflect differences across subgroups) if the model is non-linear. A more appropriate approach would be to estimate total costs and QALY outcomes for each treatment in the three psoriasis subgroups separately and then weight these according to the distribution of psoriasis severity in the population to estimate mean total costs and QALYs for each treatment in the overall population. The incremental analysis would then use these weighted costs and QALYs to estimate the cost-effectiveness of the treatments in the overall population. The ERG applied this approach to update the company's base-case analysis for the three subpopulations and found minimal differences in the estimates of cost-effectiveness, suggesting reasonable linearity within the model. Therefore, a correction was not deemed necessary, given time constraints.

5.2 Company's sensitivity analyses

The company conducted a large number of scenario analyses (these are reported in Section B3.8.3 of the CS). The majority of scenarios had minimal impact on the ICER results of the fully incremental analysis, with nearly all other interventions either dominated or extendedly dominated in all three populations. Three scenarios that had the largest effect on the ICER were 1) equivalent annual treatment discontinuation rate of 16.5% across all comparators; 2) utility values derived from the DISCOVER-2 algorithm; and 3) ARC 20 alternative responder definition.

- In the scenario where the company used a constant annual treatment discontinuation rate of 16.5% across all comparators rather than a treatment-specific discontinuation rate, the discontinuation rate had a differential effect on ICERs across all therapies and populations. [redacted]

[redacted]

- In the scenario where the company used the utility values derived from the DISCOVER-2 algorithm rather than the York algorithm, the DISCOVER-2 utility algorithm was associated with less favourable ICERs for guselkumab compared with the York algorithm ([redacted]); however, guselkumab dominates or extendedly dominates all other treatments in all three subpopulations.
- In the scenario where the company used ARC 20 as an alternative responder definition rather than PsARC, the ARC 20 response definition was associated with less favourable ICERs for guselkumab compared with PsARC ([redacted]); however, guselkumab dominates or extendedly dominates all other treatments in all three subpopulations.

In light of ERG's item 11 in Section 4.2.12.4 above, it is also worth noting that the company conducted a scenario analysis using the Kobelt et al., 2002, algorithm as an alternative source of disease-related HAQ-DI costs. The fully incremental ICER results were slightly more favourable to guselkumab compared to the company's base case results ([redacted]).

5.3 Model validation and face validity check

The company describes the model validation process in Section B 3.10 of the CS. The ERG undertook further validation checks and identified errors in the parametrisation of two inputs in the company's model:

1. The estimate of discontinuation rates for secukinumab was not updated when different doses of secukinumab were assumed.
2. The dose of second line secukinumab – assumed to be a mix of 150mg and 300mg according to psoriasis severity, when a 300 mg dose is indicated for biologic experienced patients.

These errors were, however, corrected by the company in the updated versions of the model submitted at the PfC. No other face validity issues were identified with the model.

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

A summary of the main issues identified and critiqued in Section 4 along with the scenario where the ERG addresses each issue in its additional analyses is shown in Table 33.

Table 33 Summary of the main issues identified by the ERG in Section 4

Critique item and description The ERG considers that:		Dealt with in the ERG base case	Dealt with in the ERG scenario analyses	Area of remaining uncertainty	Significant impact on ICER
1	***Baseline PASI scores and proportion of patients by psoriasis severity in the DISCOVER trials may not match those seen in UK clinical practice.***		Sc.1	x	
2	***The modelled treatment sequences do not reflect the range of treatment sequences seen in UK clinical practice. Importantly, treatment with multiple lines of active therapy is a valid treatment strategy instead of placing patients on BSC at second and third line.***	Partly		x	x
3	***Continuation of treatment based on achievement of PASI 75 response for patients whose PsARC response does not justify continuation of treatment is an area of uncertainty.***			x	unknown
4	***The use of treatment-specific discontinuation rates in the maintenance period in the model is a key driver of cost-effectiveness but is informed by a limited and potentially biased evidence base.***	x	Sc.2	x	x
5	***Treatment-specific discontinuation rates should only be modelled when the appropriate range of treatment sequences are considered that reflect the full duration of disease.***	x	Sc.2	x	x
6	***The assessment time point for response to treatment for guselkumab is based on a stopping rule at 24 weeks, in anticipation of its expected marketing authorisation for PsA, but it is unclear why a 16-week assessment time point should not also be considered in decision making.***		Sc.3	x	Unknown, but potentially cost-saving
7	***The company does not present cost-effectiveness results for baseline unadjusted NMA models, which represents an area of uncertainty.***		Sc.4	x	
8	***The company's approach to include adverse events in the model is unlikely to reflect the safety profile of the different treatments and is not consistent with the assumptions of previous TAs in PsA.***	x		x	
9	***Administration costs should be included for drugs administered subcutaneously.***	x	Sc.5		
10	***Monitoring resource use should be the same across treatments to be consistent with previous TAs in PsA.***	x	Sc.6		
11	***The Kobelt et al., 2002, algorithm should be used to estimate arthritis related costs in the base-case analysis to ensure consistency***	x		x	

6.1 Exploratory and sensitivity analyses undertaken by the ERG

As shown in Table 33, the ERG identified a number of limitations and areas of uncertainty in the company's cost-effectiveness analysis. The elements where the ERG considered there to be a more appropriate alternative approach were modified and form part of the ERG's preferred base case assumptions. Each element is described below and the corresponding impact on the ICER shown in Section 6.2. Elements which the ERG considered as important areas of uncertainty or had a significant impact on the ICER are highlighted in Section 6.2. The effect of making changes simultaneously on elements that are considered to form part of the ERG's preferred base case assumptions is presented in Section 6.3.

The ERG did not perform any corrections to the company's model. The errors identified and referred to in Section 5.3 were corrected in the updated version of the electronic model that was submitted by the company in response to ERG points for clarification. The only modifications implemented in the model by the ERG consist of adding functionality to:

i. Apply confidential PAS discounts to the comparators;
ii. Apply estimates of PsARC and PASI response derived from the CODA for unadjusted random effects NMA models submitted by the company in response to ERG points for clarification.

6.1.1 Issues explored by the ERG in additional analyses

6.1.2 *Scenario 1. Using an alternative source for the level of psoriasis severity and baseline PASI and HAQ-DI score in the population*

This relates to *item 1* where the baseline PASI score and proportion of patients by psoriasis severity used in the company's model is based on the patient characteristics of the DISCOVER trials, which may not match those seen in UK clinical practice. The ERG assessed the impact of using the proportion of patients by psoriasis severity and baseline PASI and HAQ-DI scores from the updated York model (TA445) based on clinical opinion relevant to UK practice. In this analysis, 50% of patients have minimal concomitant psoriasis, 25% have mild to moderate concomitant psoriasis and 25% moderate to severe concomitant psoriasis. The baseline PASI scores for these subgroups were PASI=0, 7.3 and 12.5 for minimal, mild to moderate, and moderate to severe psoriasis, respectively, while the baseline HAQ-DI score was 1.22 for all subgroups and subpopulations in TA445. The corresponding weighted average baseline PASI score for the subpopulations based on the proportion of patients with different levels of psoriasis severity is 4.95, while the baseline HAQ-DI score is 1.22.

6.1.3 *Scenario 2. Using an equivalent annual treatment discontinuation rate for the interleukin therapies and a separate annual treatment discontinuation rate for TNFi therapies*

The use of treatment-specific discontinuation rates in the maintenance period of the model is a key driver of the company's cost-effectiveness results. This was demonstrated in the company's scenario analysis that used a constant annual treatment discontinuation rate of 16.5% across all comparators, resulting in a differential effect on ICERs across all therapies in all subpopulations. A 16.5% equivalent discontinuation rate for all treatments was used in all previous TAs in PsA. The company argued that this equivalent rate is not supported by evidence and is based on TA199, which only considered TNFi therapies and no alternative mechanisms of action such as those for interleukin inhibitors. Although the ERG supports the company that the use of an equivalent discontinuation rate for all treatments is likely to be highly questionable, the evidence presented by the company to support the treatment-specific discontinuation rates used in the model is based on very limited data, and is potentially biased towards guselkumab by assuming a substantially lower rate for guselkumab compared with the other interleukin inhibitors.

In this scenario the ERG assessed the impact of using an equivalent annual treatment discontinuation rate for the interleukin therapies and newer agents (i.e., guselkumab, secukinumab, ixekizumab, ustekinumab, and tofacitinib) and a separate annual treatment discontinuation rate for TNFi therapies (i.e., adalimumab, etanercept, infliximab, certolizumab, golimumab) and apremilast. For the interleukin therapies and newer agents, the annual discontinuation rate was set to 11.4%, which is the average of the treatment-specific values used in the company's model for these treatments, while the annual discontinuation rate for TNFi therapies and apremilast was set to 16.5% in line with TA199.

6.1.4 *Scenario 3. Using an alternative* 16 week stopping rule for guselkumab non-responders

In this scenario the ERG tests the impact of assuming an alternative stopping rule at 16 weeks for guselkumab non-responders in order to address *item 6*. This is explored in the updated 16 weeks model submitted by the company at response to points for clarification, and which incorporates the results of the NMAs using 16 week data requested by the ERG at points for clarification (question A16). Briefly, the guselkumab estimates of PsARC response and HAQ-DI conditional on PsARC response were broadly similar to the base-case assumption, but sizeable differences were patent for PASI response (see Table 23 and Table 26). The NMA estimated guselkumab 24 weeks PASI75 in the biologic-naïve population was 80.2% compared to 69.6% at 16 weeks, while the PASI100 was 44.3% and 31.7% at 24 and 16 weeks respectively.

However, the structure of the electronic model submitted by the company does not allow capturing the potential continued improvement in PASI response between 16 and 24 weeks for patients treated

with guselkumab. The model assumes that PASI scores remain constant for the duration of residence in the maintenance treatment health state, and so patients will not experience any further improvement in PASI scores beyond the first 16 weeks of treatment with guselkumab, which is not supported by the evidence. Furthermore, there are concerns that outcome data collected in the DISCOVER trials beyond 16 weeks are potentially affected by bias resulting from patients being allowed an "early escape" at that time point (see Section 3.2.1,3.2.2.1 and 3.2.2.2). Therefore, the model is not suitable to explore the full impact on outcomes of using a 16 weeks stopping rule for guselkumab, as it may not accurately capture the QALY gains associated with an improvement of PASI response up from 16 to 24 weeks, and over maintenance treatment. Similarly, it may not fully capture the potential benefit in terms of avoided psoriasis related costs. The 16 weeks model could, thus, misrepresent the benefits of guselkumab. Furthermore, this version of the model also assumes a 16 weeks rule for ustekinumab, which does not reflect clinical practice. This assumption impacts on the estimates of cost-effectiveness for the comparators, since this drug is used as second line of therapy for most sequences.

Given the limitations of the 16 weeks electronic model, the ERG shows the impact on cost of treatment at first line of therapy under the assumptions of this scenario analysis, and does not show results of full incremental analyses.

6.1.5 *Scenario 4. Using an alternative source of effectiveness data without a placebo response adjustment*

This scenario analysis uses an alternative source for estimates of PsARC and PASI response rates, which in the company base-case analysis were sourced from placebo response-adjusted NMA models for the biologic-naïve evidence network. In the scenario analysis, PsARC and PASI response are sourced from the company's NMA models unadjusted for placebo response to address *item 7*. PsARC and PASI response rates applied in this scenario are reported in Table 42 and Table 43, respectively.

The analysis is not conducted for the biologic-experienced subpopulation, because:

1. The impact on the estimates of cost-effectiveness for this subpopulation are expected to be smaller, as only the PASI response rates in the company's base-case were informed by a placebo response-adjusted model.
2. The direction of change in PASI response relative to the company's base-case estimates is similar to that of the biologic-naïve subpopulation.

6.1.6 *Scenario 5. Including a cost for the administration of subcutaneous drugs*

This scenario analysis applies a one-off treatment costs at the first cycle of treatment for all drugs administered subcutaneously, so as to address *item 9*. This cost corresponds to one hour of a specialist nurse time (£47)[57] to teach the patient how to self-administer the drug.

6.1.7 *Scenario 6. Applying estimates of monitoring resource use consistent with previous TAs in PsA*

The company's assumptions on the monitoring resource use for the intervention and comparators was not considered consistent with previous TAs in PsA as detailed in Section 4.2.12.3 . To address *item 10*, the ERG uses the estimates of monitoring resource use used in TA445 and TA543 (see Table 28), as an alternative to the company's assumptions on these parameters.

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

All results for the ERG scenarios are based on a deterministic analysis because the time required to run the model probabilistically across all scenarios was not feasible within the time constraints of the STA. However, the ERG did compare the results of a probabilistic and deterministic analysis across a number of scenarios and confirmed that the results were similar, suggesting reasonable linearity within the model.

The results presented in this and the subsequent section refer to the fully incremental cost-effectiveness analyses and only the results of non-dominated treatments are shown (i.e., after exclusion of dominated or extendedly dominated treatments). The only exception is for the results of guselkumab, which are always presented alongside other treatments, even when it is dominated. Results for the three subpopulations are shown throughout the section, with corresponding results for the psoriasis severity subgroups are reported in Appendix 9.2.

Similarly to the results presented in Section 5, all analyses include the confidential PAS discount associated with guselkumab, and include the publicly available PAS for comparators. Equivalent analyses reflecting the cost of comparator technologies in the NHS (i.e., with costs as per confidential commercial arrangements between the market authorisation holders and the health system) are present in the confidential PAS appendix.

6.2.1 *Scenario 1. Using an alternative source for the level of psoriasis severity and baseline PASI and HAQ-DI score in the population*

Table 34 shows the results of the fully incremental analysis for the three subpopulations using the alternative source (TA445) for the level of psoriasis severity and baseline PASI and HAQ-DI score in the population. Mean total costs decrease and mean total QALYs increase for all treatments across the three subpopulations, reflecting the higher proportion of patients with less severe psoriasis. Guselkumab remains a potentially cost-effective alternative in all subpopulations, but with higher ICERs compared to BSC in the biologic-naïve subpopulation relative to corresponding analyses for the biologic-experienced and the TNFi contraindicated subpopulations. Once a higher proportion of

patients with less severe psoriasis are included in the model, etanercept is no longer extendedly dominated by guselkumab, but the ICER of guselkumab compared to etanercept remains within conventional NICE cost-effectiveness thresholds.

Table 34 Results of scenario 1 for alternative baseline population characteristics

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naive					
BSC	£76,678	5.287	-	-	-
Etanercept	[redacted]	8.128	[redacted]	2.842	[redacted]
Guselkumab Q8W	[redacted]	8.712	[redacted]	0.584	[redacted]
All biologic-experienced					
BSC	£78,102	5.323	-	-	-
Guselkumab Q8W	[redacted]	7.883	[redacted]	2.560	[redacted]
All TNFi-contraindicated					
BSC	£76,678	5.287	-	-	-
Guselkumab Q8W	[redacted]	8.712	[redacted]	3.425	[redacted]

6.2.2 Scenario 2. Using an equivalent annual treatment discontinuation rate of 11.4% for the interleukin therapies and a separate annual treatment discontinuation rate of 16.5% for TNFi therapies

Table 35 shows the results of the fully incremental analysis for the three subpopulations using an equivalent annual treatment discontinuation rate of 11.4% for the interleukin therapies and newer agents (guselkumab, secukinumab, ixekizumab, ustekinumab, and tofacitinib) and a separate annual treatment discontinuation rate of 16.5% for TNFi therapies (adalimumab, etanercept, infliximab, certolizumab, golimumab) and apremilast. The direction of effect on total costs and QALYs is similar to the company's scenario analysis using an annual treatment discontinuation rate of 16.5% across all comparators, which resulted in a differential effect on ICERs across all treatments and populations. In the biologic-naïve subpopulation,

[redacted]

When the discontinuation rates are equivalent across interleukin inhibitors and TNFi therapies, the mean total QALYs accrued by guselkumab and the other active treatments become more similar and the differences in total costs (and hence cost-effectiveness) is largely determined by drug costs (rather

than disease-related or other costs). The discontinuation rate affects the time on treatment, such that once a patient has failed to respond to treatment they are either moved to the next line of therapy or BSC. In the company's model only one line of active therapy is considered for the biologic-experienced subpopulation, and two lines for the biologic-naïve and TNFi-contraindicated subpopulations, before moving to BSC. This means that the discontinuation rate is delaying time to BSC, which has the lowest QALY benefits and highest disease-related costs. When the discontinuation rates are equivalent across therapies, the difference in time on treatment between the alternative active therapies is reduced making the total QALY benefits and non-drug related costs more similar across all treatments. Furthermore, equivalent discontinuation rates have a differential effect on the ICERs compared to the base case analysis because a limited number of lines of therapy are considered before BSC. If the modelled treatment duration is sufficiently long, by incorporating multiple lines of active therapy (rather than BSC), the impact of the discontinuation rate is less important.

Table 35 Results of scenario 2 for alternative discontinuation rates – 16.5% (TNFi) and 11.4% other treatments

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naive					
BSC	£83,722	5.014	-	-	-
Etanercept	[redacted]	8.080	[redacted]	3.066	[redacted]
Guselkumab Q8W	[redacted]	8.097	[redacted]	0.017	[redacted]
All biologic-experienced					
BSC	£86,772	4.254	-	-	-
Guselkumab Q8W	[redacted]	6.025	[redacted]	1.771	[redacted]
All TNFi-contraindicated					
BSC	£83,722	5.014	-	-	-
Guselkumab Q8W	[redacted]	8.097	[redacted]	3.083	[redacted]
Ustekinumab	[redacted]	8.110	[redacted]	0.013	[redacted]

6.2.3 Scenario 3. Using an alternative 16 weeks stopping rule for guselkumab non-responders

Table 36 reports the cost of first line treatment with guselkumab (drug acquisition costs only) for the three subpopulations using the company's base-case assumption of a 24 weeks stopping rule for guselkumab non-responders and an alternative 16 weeks stopping rule.

Table 36 Comparison of guselkumab acquisition costs at first line of therapy with alternative stopping rules

	Guselkumab acquisition costs at first line of therapy		Difference
	24 weeks stopping rule	16 weeks stopping rule	

All bDMARD-naïve/ All TNFi-contraindicated	[redacted]	[redacted]	[redacted]
All biologic-experienced	[redacted]	[redacted]	[redacted]

The use of a 16 weeks stopping rule for patients whose arthritis symptoms do not respond to guselkumab could result in considerable savings across the three subpopulations of the model. Although the ERG could not formally model the impact on outcomes of using this alternative stopping rule for guselkumab, these savings could potentially be achieved, while maintaining a similar size of QALY gains as those generated when the 24 weeks stopping rule is used.

6.2.4 Scenario 4. Using an alternative source of effectiveness data without a placebo response adjustment

Table 37 shows the results of the fully incremental analysis for two subpopulations, biologic-naïve and TNFi contraindicated, using an alternative source for estimates of PsARC and PASI response. In the scenario analysis, these estimates are sourced from the company's NMA models unadjusted for placebo response.

As described in Section 4.2.8, the PSARC response rate for guselkumab is lower without the placebo response adjustment and similar to that of other interleukin inhibitors (e.g. secukinumab and ixekizumab), while there were clearer differences between these biologics when using placebo-response adjusted NMA (see Table 42 and Figure 9). TNFi biologics have generally higher response rates under this assumption, reflecting the lower placebo-response rates in the pivotal clinical trials. Similar patterns were observed for PASI response (see Table 43).

These changes in effectiveness result in lower mean costs for guselkumab compared to the company's base-case analysis [redacted], because fewer responders result in a reduction in costs associated with maintenance treatment for guselkumab (even if some of this cost reduction is offset by the increase on disease related costs). Mean QALYs accrued with guselkumab are also reduced compared to the company's base-case analysis (8.305 vs. 8.511QALYs), since fewer responders and a worse PASI response reduce benefits accrued in the treatment maintenance period at first line of treatment. For the biologic naïve subpopulation, the mean costs and QALYs accrued with etanercept change in the opposite direction to guselkumab, with this TNFi becoming less costly and more effective compared to the company's base-case results. In this scenario, etanercept is no longer [redacted] and is the cost-effective treatment, while the ICER of guselkumab compared to etanercept [redacted]. For the TNFi subpopulation, guselkumab remains the cost-effective treatment, albeit with a slightly higher ICER compared to BSC (cost savings are offset by the worse health outcomes).

The ERG notes that the mean total costs and QALYs of treatments not shown in Table 37 (strictly and extendedly dominated treatments), also change in the expected direction compared to the company's base-case estimates (e.g. golimumab is more costly and more effective). However, under the assumption of differential discontinuation rates the impact of changes in effectiveness and cost-effectiveness of treatments is constrained, and this constraint reduces the impact on incremental relationships between treatments.

Table 37 Results of scenario 4 for an alternative source of effectiveness data without a placebo response adjustment

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naive					
BSC	£83,705	5.023	-	-	-
Etanercept	[redacted]	7.984	[redacted]	2.962	[redacted]
Guselkumab Q8W	[redacted]	8.305	[redacted]	0.321	[redacted]
All TNFi-contraindicated					
BSC	£83,705	5.023	-	-	-
Guselkumab	[redacted]	8.305	[redacted]	3.282	[redacted]

6.2.5 Scenario 5. Including a cost for the administration of subcutaneous drugs

Table 38 shows the results of the fully incremental analysis for the three subpopulations assuming all subcutaneously administered treatments incur a one-off cost of a nurse teaching the patient how to self-administer the drug (£47).

This alternative as a modest impact on the cost-effectiveness estimates. All subcutaneously administered treatments, have an increase in costs of at least £47 and more for those analyses where treatment sequences with two lines of active therapy (i.e., in the biologic-naïve and TNFi contraindicated populations). The ICER of guselkumab vs. BSC increases marginally compared to the company's base-case analysis for all subpopulations.

Table 38 Results of scenario 5 including a cost for the administration of subcutaneous drugs

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naïve/ All TNFi-contraindicated					
BSC	£ 83,722	5.014	-	-	-
Guselkumab Q8W	[redacted]	8.511	[redacted]	3.497	[redacted]

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All biologic-experienced					
BSC	£86,772	4.254	-	-	-
Guselkumab Q8W	[redacted]	6.800	[redacted]	2.546	[redacted]

6.2.6 Scenario 6. Applying estimates of monitoring resource use consistent with previous TAs in PsA

Table 39 shows the results of the fully incremental analysis for the three subpopulations assuming the same resource use associated with monitoring patients across all interventions and comparators to be consistent with the assumptions of previous TAs in PsA. The cost of monitoring reduces to £207.63 over the trial period and to £15.56 per annual maintenance treatment cycle for all therapies (including BSC). This is in contrast with the company's base-case analysis, which assumed the cost of monitoring to be:

- Trial period: £336.97 and £354.14 over the full interval for oral therapies and all other treatments, respectively;
- Maintenance treatment: £146.51 and £308.58 per annual cycle for oral therapies and all other treatments, respectively.

Under this scenario, the costs of monitoring reduce for all treatments, but proportionally less for the oral therapies (apremilast and tofacitinib). For all other therapies, the exact size of the cost reduction will depend on how "early" patients arrive to the maintenance treatment, favouring slightly therapies with shorter trial periods. Since the cost reduction is common to all therapies and differences between therapies are small (an artefact of the different response assessment trial durations), the impact of this alternative assumption in the cost-effectiveness of guselkumab is minor across all subpopulations.

Table 39 Results of scenario 6 applying estimates of monitoring resource use consistent with previous TAs in PsA

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naïve / All TNFi-contraindicated					
BSC	£77,791	5.014	-	-	-
Guselkumab Q8W	[redacted]	8.511	[redacted]	3.497	[redacted]
All biologic-experienced					

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
BSC	£80,841	4.254	-	-	-
Guselkumab Q8W	[redacted]	6.800	[redacted]	2.546	[redacted]

6.3 ERG's preferred assumptions

The scenarios identified by the ERG as providing an alternative and more appropriate source to inform the cost-effectiveness of guselkumab are the company's scenarios i) using Kobelt et al, 2002 algorithm to estimate costs[44], ii) the 16.5% annual discontinuation rate for all biologic therapies, and 3) not including costs and disutilities associated to adverse events of treatment, and the ERG scenarios 5 and 6. These scenarios combined form part of the ERG's preferred base case assumptions.

The scenarios identified as having a significant impact on the ICER are the company's scenario using the 16.5% discontinuation rate for all biologics and the ERG scenarios 1 and 2. The company's scenario on treatment discontinuation rates forms part of the ERG's preferred base case, while scenario 1 and 2 remain an area of uncertainty that could not be addressed by the ERG due to insufficient information. Scenario 3 raises an important policy question on the use of guselkumab, but this could not be fully addressed due to constraints of the model structure.

The ERG base-case assumptions can be summarised as:

1. Annual discontinuation rate of 16.5% for all biologic therapies;
2. Arthritis related costs estimated with Kobelt et al., 2002, algorithm;
3. No difference in costs and disutilities associated with adverse events across treatments;
4. One-off cost for the administration of subcutaneous therapies;
5. Monitoring resource use consistent with previous TAs in PsA.

In addition, the ERG base-case is presented using two alternative sources of treatment evidence, so as to reflect that in the absence of a clear rationale for the placebo effect and given the uncertainty around the differences in effectiveness across classes of treatments, both placebo response-adjusted and unadjusted NMA models must be considered.

The only assumption with significant impact on results is the 16.5% discontinuation rate applied to all biologic therapies. Therefore, and given the burden of analyses, the ERG base case will be presented without the cumulative effect of each of the applied assumptions.

Table 40 and Table 41 present the results of the ERG base case analysis for all three subpopulations sourcing effectiveness for the placebo response-adjusted and unadjusted models, respectively. The ERG presents results for all treatments in the biologic-naïve subpopulation, even if these are dominated, so as to facilitate the comparison between the two analyses.

For the analyses where effectiveness (PsARC and PASI response rates) are sourced from the placebo-response adjusted NMA models, results across the three subpopulations are similar to the company's scenario assuming a 16.5% discontinuation rate for all therapies. The key differences are that mean total costs are lower across all treatments, while mean QALYs are higher. This is due to the alternative assumptions of the ERG on costing, which generally reduce costs, and the exclusion of adverse event costs and disutilities. All other things being equal, these alternative assumptions have a bigger impact in terms of cost reduction than on QALY increase.

For the biologic-experience and TNFi subpopulations, guselkumab still dominates all comparators, except BSC. The reduction in incremental costs of guselkumab vs. BSC compared to the company's scenario with same discontinuation rates results in a small reduction of the ICERs for these analysis.

Table 40 ERG base-case results sourcing effectiveness from the placebo response-adjusted model

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naive					
BSC	£76,360	5.022	-	-	-
Etanercept	[redacted]	7.776	[redacted]	2.753	[redacted]
Infliximab	[redacted]	7.857	[redacted]	0.082	[redacted]
Apremilast	[redacted]	6.973	-	-	[redacted]
Tofacitinib	[redacted]	6.980	-	-	[redacted]
Adalimumab	[redacted]	7.254	-	-	[redacted]
Guselkumab	[redacted]	7.393	-	-	[redacted]
Certolizumab	[redacted]	7.291	-	-	[redacted]
Golimumab SC	[redacted]	7.374	-	-	[redacted]
Secukinumab	[redacted]	7.380	-	-	[redacted]
Ixekizumab	[redacted]	7.179	-	-	[redacted]
All biologic-experienced					
BSC	£77,963	4.265	-	-	-
Guselkumab	[redacted]	5.524	[redacted]	1.259	[redacted]
All TNFi-contraindicated					
BSC	£76,360	5.022	-	-	-
Guselkumab	[redacted]	7.393	[redacted]	2.370	[redacted]

ADA, adalimumab: APR, apremilast; CTZ, certolizumab; ETA, etanercept; INF, infliximab; IXE, ixekizumab; SEC, secukinumab; TF, tofacitinib; UST, ustekinumab

The two ERG alternative base-case analyses appear to produce similar cost-effectiveness results with the exception of the ICER for infliximab vs. etanercept in the biologic-naïve subpopulation, which decreases substantially due to the increased effectiveness of infliximab (see Figure 9). However, the ERG notes that the total mean QALYs accrued by guselkumab, ixekizumab, secukinumab, certolizumab and adalimumab are much closer when using unadjusted estimates of effectiveness then when the analysis controls for a different placebo effect across studies. Importantly, the mean QALY gains for guselkumab are considerably lower when using the effectiveness estimates unadjusted for placebo response (-0.102 QALYs) in the biologic-naïve and TNFi contraindicated subpopulations.

Table 41 ERG base-case results sourcing effectiveness from the unadjusted model

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
All bDMARD-naive					
BSC	£76,346	5.031	-	-	-
Etanercept	[redacted]	7.883	[redacted]	2.852	[redacted]
Infliximab	[redacted]	8.044	[redacted]	0.161	[redacted]
Guselkumab	[redacted]	7.291	-	-	[redacted]
Apremilast	[redacted]	7.049	-	-	[redacted]
Tofacitinib	[redacted]	6.884	-	-	[redacted]
Adalimumab	[redacted]	7.283	-	-	[redacted]
Certolizumab	[redacted]	7.206	-	-	[redacted]
Guselkumab	[redacted]	7.291	-	-	[redacted]
Secukinumab	[redacted]	7.239	-	-	[redacted]
Golimumab SC	[redacted]	7.581	-	-	[redacted]
Ixekizumab	[redacted]	7.199	-	-	[redacted]

All biologic-experienced					
BSC	£77,498	4.278	-	-	-
Guselkumab	[redacted]	5.538	[redacted]	1.261	[redacted]
All TNFi-contraindicated					
BSC	£76,346	5.031	-	-	-
Guselkumab	[redacted]	7.291	[redacted]	2.260	[redacted]

ADA, adalimumab: APR, apremilast; CTZ, certolizumab; ETA, etanercept; GOL, golimumab; GUS, guselkumab INF, infliximab; IXE, ixekizumab; SEC, secukinumab; TF, tofacitinib; UST, ustekinumab

6.4 Conclusions of the cost effectiveness section

The company submitted a de novo Markov model to assess the cost-effectiveness of guselkumab versus other treatments previously recommended by NICE for the treatment of PsA. The model followed the structure of the revised York model from TA445 and was largely based on the assumptions used in previous models in the most recent TAs (TA445, TA537 and TA543). Where different assumptions and data was used the company outlines these in their submission. The ERG considers the approach used by the company as appropriate and accurately reflects the decision problem defined in the final NICE scope. However, the ERG has identified a number of concerns where the company has deviated from the assumptions used in previous TAs and the ERG believes there is no compelling case to suggest that the company's alternative approach is more appropriate than the assumptions used in previous TAs..

The major difference between previous TAs in PsA and the key driver of the cost-effectiveness results is the assumption of a differential treatment discontinuation rate across all treatments in the maintenance period of the model (after initial response to treatment). The treatment-specific withdrawal rates used in the company's model suggests that guselkumab has a much lower rate of discontinuation (6.9% per annum) compared to other interleukin modulators (ranging from 10.2% to 14.9% per annum) and tofacitinib (10.3% per annum), as well as TNFi's (ranging from 11.7% to 22.1% per annum) and apremilast (26.5% per annum). The evidence supporting this is based on discontinuation rates in extensions of pivotal clinical trials, which may not be reflective of what is expected to be seen in UK clinical practice (noting, for example, that the DISCOVER trials have a very high prevalence of Eastern European trial sites where the available of other alternative treatment options may be limited). Therefore, comparisons of treatment-specific discontinuation rate estimates based on clinical trial data are subject to considerable uncertainty. This is because, in addition to being affected by treatment-specific effects, such as lack of efficacy and adverse events, discontinuation rates may also be affected by non-specific factors such as healthcare settings, trial protocols, levels of staff involvement and attitude, as well as the play of chance. Without compelling evidence to suggest appropriate treatment-specific discontinuation rates, the ERG considers it more

appropriate to use an equivalent annual treatment discontinuation rate across all therapies to ensure consistency with previous TAs in PsA.

Furthermore, treatment-specific discontinuation rates would only be considered appropriate when the full duration of disease is modelled with the range of treatments available at different lines of therapy. The modelled treatment sequences do not reflect the range of treatment sequences seen in UK clinical practice. Clinical opinion suggests that switching among different TNFi therapies represents a valid treatment strategy as well as switching to different IL modulators (ustekinumab, secukinumab and ixekizumab) or tofacitinib. In other words, treatment with multiple lines of active therapy, instead of placing patients on best supportive care (BSC) at second and third line is considered valid (i.e., patients are unlikely to receive only two active therapies in the biologic-naïve or TNFi contraindicated populations, or only one active therapy after ≥1 TNFi's in the biologic-experienced population, as modelled in the CS). The ERG recognises that there is an absence of a standardised approach to treatment sequencing and lack of evidence to inform the effectiveness of switching between active therapies. However, the ERG is emphasising this issue in relation to the treatment-specific discontinuation rates used in the model; these rates have a material impact on the ICERs because the modelled treatment duration may not be sufficiently long enough with multiple lines of active therapy before reaching final line BSC.

The ERG is concerned that the company's cost-effectiveness results may not be reflecting the appropriate assessment time point for response to treatment for guselkumab. The company has based this on a stopping rule at 24 weeks, in anticipation of its expected marketing authorisation for PsA, but it is unclear why a 16-week assessment time point should not also be considered in decision making. Although the ERG could not formally model the impact on outcomes of using an alternative stopping rule of 16 weeks for guselkumab, the effectiveness data suggests that potential cost savings could be achieved (i.e., avoiding the need to extend treatment for non-responders by a further 8 weeks), while maintaining a similar size of QALY gains as those generated when the 24-week stopping rule is used (i.e., the proportion of responders at 16 weeks is very similar to the proportion of responders at 24 weeks in the DISCOVER trials). This remains an area of uncertainty.

The ERG is also concerned that the efficacy evidence informing the cost-effectiveness of guselkumab compared with the other active therapies is based on baseline risk-adjusted models for the biologic-naïve and TNFi-contraindicated populations without giving adequate consideration to the potential impact of placebo-response adjustment, which can adjust the relative differences in response probabilities (and ranking) without a clear rationale. In terms of the presentation of fully incremental cost-effectiveness analysis, treatments that are very similar in total costs and QALYs may be dominated or extendedly dominated by guselkumab when there is no robust evidence to support meaningful differences between the therapies. Therefore, the ERG considers it important to present

the cost-effectiveness results of both baseline risk-adjusted and unadjusted models to support decision making.

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are: (i) the use of an equivalent annual treatment discontinuation rate across all therapies; (ii) the exclusion of differences in adverse event costs and disutility between the treatments; (iii) the use of an alternative source for arthritis related costs; (iv) inclusion of a one-off administration cost for subcutaneous therapies; and (v) equivalent monitoring resource use across all treatments. The ERG's preferred assumptions are aimed at ensuring consistency with previous TAs in PsA. Where the company has not presented compelling evidence to support a change from previous TAs (namely, TA445, TA537 and TA543) the ERG's preferred base case is in line with the assumptions used in these TAs. In addition, the ERG's preferred base-case is presented for both placebo response-adjusted and unadjusted NMA models.

The company's base case fully incremental cost-effectiveness results for guselkumab indicate that all other interventions are either dominated or extendedly dominated in all three populations, making the ICER comparison of guselkumab with BSC. The ICER for guselkumab versus BSC was [redacted] per additional QALY in all three populations (biologic-naïve, biologic-experienced, and TNFi-contraindicated). The conclusions of the ERG's results (without incorporating the confidential PAS discounts for the comparators, except those publically available in the CS) are similar for the biologic-experienced and TNFi-contraindicated populations. For the biologic-naïve population, the most cost-effective treatment option is etanercept. The ERG results with the confidential PAS for the comparators and the confidential framework prices for the biosimilars of [redacted] provided by the Department of Health and Social Care Commercial Medicines Unit (CMU) are reported in the confidential PAS appendix.

7 END OF LIFE

As psoriatic arthritis is not a life-threatening condition, and life expectancy exceeds 24 months, end-of-life considerations do not apply.

8 REFERENCES

1. Deodhar A, Gottlieb AB, Boehncke W-H, Dong B, Wang Y, Zhuang Y*, et al.* Efficacy and safety of guselkumab in patients with active psoriatic arthritis: a randomised, double-blind, placebo-controlled, phase 2 study. *Lancet* 2018;**391**:2213-24. 10.1016/s0140-6736(18)30952-8

2. Mease PJ, Rahman P, Gottlieb AB, Kollmeier AP, Hsia EC, Xu XL*, et al.* Guselkumab in biologic-naive patients with active psoriatic arthritis (DISCOVER-2): a double-blind, randomised, placebo-controlled phase 3 trial. *Lancet* 2020;**395**:1126-36. 10.1016/s0140-6736(20)30263-4

3. Blauvelt A, Papp KA, Griffiths CE, Randazzo B, Wasfi Y, Shen YK*, et al.* Efficacy and safety of guselkumab, an anti-interleukin-23 monoclonal antibody, compared with adalimumab for the continuous treatment of patients with moderate to severe psoriasis: Results from the phase III, double-blinded, placebo- and active comparator-controlled VOYAGE 1 trial. *J Am Acad Dermatol* 2017;**76**:405-17. S0190-9622(16)31157-4 [pii]10.1016/j.jaad.2016.11.041

4. Reich K, Armstrong AW, Foley P, Song M, Wasfi Y, Randazzo B*, et al.* Efficacy and safety of guselkumab, an anti-interleukin-23 monoclonal antibody, compared with adalimumab for the treatment of patients with moderate to severe psoriasis with randomized withdrawal and retreatment: Results from the phase III, double-blind, placebo- and active comparator-controlled VOYAGE 2 trial. *J Am Acad Dermatol* 2017;**76**:418-31. S0190-9622(16)31158-6 [pii]10.1016/j.jaad.2016.11.042

5. Desai M. Recruitment and retention of participants in clinical studies: Critical issues and challenges. *Perspect Clin Res* 2020;**11**:51-3.

6. Ritchlin C, Rahman P, Kavanaugh A, McInnes IB, Puig L, Li S*, et al.* Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo-controlled, randomised PSUMMIT 2 trial. *Ann Rheum Dis* 2014;**73**:990-9. annrheumdis-2013-204655 [pii]10.1136/annrheumdis-2013-204655

7. McInnes IB, Behrens F, Mease PJ, Kavanaugh A, Ritchlin C, Nash P*, et al.* Secukinumab versus adalimumab for treatment of active psoriatic arthritis (EXCEED): a double-blind, parallel-group, randomised, active-controlled, phase 3b trial. *Lancet* 2020;**395**:1496-505. 10.1016/s0140-6736(20)30564-x

8. Gandjour A, Ostwald DA. Cost Effectiveness of Secukinumab Versus Other Biologics and Apremilast in the Treatment of Active Psoriatic Arthritis in Germany. *Appl Health Econ Health Policy* 2020;**18**:109-25. 10.1007/s40258-019-00523-110.1007/s40258-019-00523-1 [pii]

9. Aiello E, Bianculli PM, Bhattacharyya D, Gunda P, Citera G. Cost-Effectiveness of Secukinumab Versus Other Biologics in the Treatment of Psoriatic Arthritis: An Argentinean Perspective. *Value Health Reg Issues* 2019;**20**:86-94. S2212-1099(19)30048-2 [pii]10.1016/j.vhri.2019.03.002

10. Capri S, Migliore A, Loconsole F, Barbieri M. A cost-effectiveness and budget impact analysis of apremilast in patients with psoriatic arthritis in Italy. *J Med Econ* 2020;**23**:353-61. 10.1080/13696998.2019.1707208

11. Coates LC, Moverley AR, McParland L, Brown S, Navarro-Coy N, O'Dwyer JL*, et al.* Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. *Lancet* 2015;**386**:2489-98. 10.1016/s0140-6736(15)00347-5

12. Purmonen T, Puolakka K, Bhattacharyya D, Jain M, Martikainen J. Cost-effectiveness analysis of secukinumab versus other biologics and apremilast in the treatment of active Psoriatic arthritis: a Finnish perspective. *Cost Eff Resour Alloc* 2018;**16**:56. 10.1186/s12962-018-0162-3

13. Goeree R, Chiva-Razavi S, Gunda P, Graham CN, Miles L, Nikoglou E*, et al.* Cost-effectiveness analysis of secukinumab for the treatment of active psoriatic arthritis: a Canadian perspective. *J Med Econ* 2018;**21**:163-73. 10.1080/13696998.2017.1384737

14. Cawson MR, Mitchell SA, Knight C, Wildey H, Spurden D, Bird A*, et al.* Systematic review, network meta-analysis and economic evaluation of biological therapy for the management of active psoriatic arthritis. *BMC Musculoskelet Disord* 2014;**15**.

15. Cummins E, Asseburg C, Prasad M, Buchanan J, Punekar YS. Cost effectiveness of golimumab for the treatment of active psoriatic arthritis. *Eur J Health Econ* 2012;**13**:801-9. 10.1007/s10198-011-0335-x

16. Bojke L, Epstein D, Craig D, Rodgers M, Woolacott N, Yang H*, et al.* Modelling the cost-effectiveness of biologic treatments for psoriatic arthritis. *Rheumatology (Oxford)* 2011;**50 Suppl 4**:iv39-iv47. ker245 [pii]10.1093/rheumatology/ker245

17. Cummins E, Asseburg C, Punekar YS, Shore E, Morris J, Briggs A*, et al.* Cost-effectiveness of infliximab for the treatment of active and progressive psoriatic arthritis. *Value Health* 2011;**14**:15-23. S1098-3015(10)00017-3 [pii]10.1016/j.jval.2010.10.016

18. Olivieri I, de Portu S, Salvarani C, Cauli A, Lubrano E, Spadaro A*, et al.* The psoriatic arthritis cost evaluation study: a cost-of-illness study on tumour necrosis factor inhibitors in psoriatic arthritis patients with inadequate response to conventional therapy. *Rheumatology (Oxford)* 2008;**47**:1664-70. ken320 [pii]10.1093/rheumatology/ken320

19. Woolacott N, Bravo Vergel Y, Hawkins N, Kainth A, Khadjesari Z, Misso K*, et al.* Etanercept and infliximab for the treatment of psoriatic arthritis: a systematic review and economic evaluation. *Health Technol Assess* 2006;**10**.

20. Bravo Vergel Y, Hawkins NS, Claxton K, Asseburg C, Palmer S, Woolacott N*, et al.* The cost-effectiveness of etanercept and infliximab for the treatment of patients with psoriatic arthritis. *Rheumatology (Oxford)* 2007;**46**:1729-35. 46/11/1729 [pii]10.1093/rheumatology/kem221

21. Bansback NJ, Ara R, Barkham N, Brennan A, Fraser AD, Conway P*, et al.* Estimating the cost and health status consequences of treatment with TNF antagonists in patients with psoriatic arthritis. *Rheumatology (Oxford)* 2006;**45**:1029-38. kel147 [pii]10.1093/rheumatology/kel147

22. Buchanan V, Sullivan W, Graham C, Miles L, Jugl SM, Gunda P*, et al.* Cost effectiveness of secukinumab for the treatment of active psoriatic arthritis in the UK. *Pharmacoeconomics* 2018;**36**:867-78.

23. All Wales Medicines Strategy Group. *Certolizumab pegol (CimziaReg.)*. Penarth: All Wales Therapaeutics and Toxicology Centre (AWTTC); 2014.

24. Yang H, Epstein D, Bojke L, Craig D, Light K, Bruce I*, et al.* Golimumab for the treatment of psoriatic arthritis. *Health Technol Assess* 2011;**15 Suppl 1**:87-95. 10.3310/hta15suppl1/10

25. Rodgers M, Epstein D, Bojke L, Yang H, Craig D, Fonseca T*, et al.* Etanercept, infliximab and adalimumab for the treatment of psoriatic arthritis: a systematic review and economic evaluation. *Health Technol Assess* 2011;**15**.

26. National Institute for Health and Care Excellence. *Etanercept, infliximab and adalimumab for the treatment of psoriatic arthritis*. London: NICE; 2010.

27. National Institute for Health and Care Excellence. *Golimumab for the treatment of psoriatic arthritis*. London: NICE; 2011.

28. National Institute for Health and Care Excellence. *Ustekinumab for treating active psoriatic arthritis*. London: NICE; 2017.

29. National Institute for Health and Care Excellence. *Apremilast for treating active psoriatic arthritis*. London: NICE; 2017.

30. Corbett M, Chehadah F, Biswas M, Moe-Byrne T, Palmer S, Soares M*, et al.* Certolizumab pegol and secukinumab for treating active psoriatic arthritis following inadequate response to disease-modifying antirheumatic drugs: a systematic review and economic evaluation. *Health Technol Assess* 2017;**21**:1-326. 10.3310/hta21560

31. National Institute for Health and Care Excellence. *Certolizumab pegol and secukinumab for treating active psoriatic arthritis after inadequate response to DMARDs* London: NICE; 2017.

32. National Institute for Health and Care Excellence. *Ixekizumab for treating active psoriatic arthritis after inadequate response to DMARDs*. London: NICE; 2018.

33. National Institute for Health and Care Excellence. *Tofacitinib for treating active psoriatic arthritis after inadequate response to DMARDs*. London: NICE; 2018.

34. Canadian Agency for Drugs and Technologies in Health. *apremilast (Otezla). Pharmacoeconomic review report*: CADTH; 2015.

35. Canadian Agency for Drugs and Technologies in Health. *Ustekinumab (Stelara). Pharmacoeconomic review report* CADTH; 2016.

36. Canadian Agency for Drugs and Technologies in Health. *Ixekizumab (Taltz). (Eli Lilly Canada Inc.) Pharmacoeconomic review report*: CADTH; 2018.

37. National Centre for Pharmacoeconomics Ireland. *Cost-effectiveness of apremilast (Otezla®) alone or in combination with Disease Modifying Antirheumatic Drugs (DMARDs) for the treatment of active psoriatic arthritis in adult patients who have had an inadequate response or who have been intolerant to a prior DMARD therap*: NCPE; 2016.

38. Scottish Medicines Consortium. *certolizumab pegol (Cimzia®)*: SMC; 2014.

39. Scottish Medicines Consortium. *ustekinumab (Stelara®)* SMC; 2014.

40. Scottish Medicines Consortium. *Apremilast (Otezla) psoriatic arthritis*: SMC; 2015.

41. Scottish Medicines Consortium. *tofacitinib citrate (Xeljanz®)* smc; 2018.

42. Sideris E, Corbett M, Palmer S, Woolacott N, Bojke L. The clinical and cost effectiveness of apremilast for treating active psoriatic arthritis: a critique of the evidence. *Pharmacoeconomics* 2016;**34**:1101-10. 10.1007/s40273-016-0419-7

43. O'Connor J, Rice S, Smith A, Rodgers M, Lopez RR, Craig D*, et al.* The clinical and cost effectiveness of ustekinumab for the treatment of psoriatic arthritis: a critique of the evidence. *Pharmacoeconomics* 2016;**34**:337-48. 10.1007/s40273-015-0350-3

44. Kobelt G, Jonsson L, Lindgren P, Young A, Eberhardt K. Modeling the progression of rheumatoid arthritis: a two-country model to estimate costs and consequences of rheumatoid arthritis. *Arthritis Rheum* 2002;**46**:2310-9. 10.1002/art.10471

45. Costa L, Perricone C, Chimenti MS, Del Puente A, Caso P, Peluso R*, et al.* Switching between biological treatments in psoriatic arthritis: a review of the evidence. *Drugs R D* 2017;**17**:509-22. 10.1007/s40268-017-0215-7

46. Gossec L, Siebert S, Bergmans P, De Vlam K, Gremese E, Joven-Ibáñez B*, et al.* Sat0398 persistence of ustekinumab (ust) or tnf inhibitor (tnfi) treatment in psoriatic arthritis (psa): Insights from the large, prospective, multinational, real-world psabio cohort. *Ann Rheum Dis* 2020;**79**:1149. 10.1136/annrheumdis-2020-eular.2127

47. Manara M, Caporali R, Favalli EG, Grosso V, Atzeni F, Sarzi-Puttini P*, et al.* Two-year retention rate of golimumab in rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis: data from the LORHEN registry. *Clin Exp Rheumatol* 2017;**35**:804-9.

48. Vieira-Sousa E, Eusébio M, Ávila-Ribeiro P, Khmelinskii N, Cruz-Machado R, Rocha TM, *et al.* Real-world longterm effectiveness of tumor necrosis factor inhibitors in psoriatic arthritis patients from the Rheumatic Diseases Portuguese Register. *The Journal of Rheumatology* 2020;**47**:690-700. 10.3899/jrheum.181272

49. Rotar Ž, Tomšič M, Praprotnik S. The persistence of golimumab compared to other tumour necrosis factor-α inhibitors in daily clinical practice for the treatment of rheumatoid arthritis, ankylosing spondylitis and psoriatic arthritis: observations from the Slovenian nation-wide longitudinal registry of patients treated with biologic disease-modifying antirheumatic drugs—BioRx.si. *Clin Rheumatol* 2019;**38**:297-305. 10.1007/s10067-018-4324-7

50. Letarouilly LG, Salmon JH, Coquerelle P, Goeb B, Guyot MH, Houvenagel E, *et al.* Biologic disease-modifying antirheumatic drugs in psoriatic arthritis: a real-world cohort of 439 patients [THU0310]. *Ann Rheum Dis* 2018;**77**:373. 10.1136/annrheumdis-2018-eular.4792

51. Glintborg B, Gudbjornsson B, Steen Krogh N, Omerovic E, Manilo N, Holland-Fischer M, *et al.* Impact of different infliximab dose regimens on treatment response and drug survival in 462 patients with psoriatic arthritis: results from the nationwide registries DANBIO and ICEBIO. *Rheumatology* 2014;**53**:2100-9. 10.1093/rheumatology/keu252

52. Korotaeva T, Loginova E, Koltakova A, Gubar E, Korsakova Y, Sedunova M, *et al.* Persistence and drug survival rates of the first-line biological (b) DMARDs in patients (pts) with psoriatic arthritis (PsA) from the Russian psoriatic arthritis registry (RU-PsART). *Ann Rheum Dis* 2019;**78**:1846. 10.1136/annrheumdis-2019-eular.1572

53. Fagerli KM, Kearsley-Fleet L, Watson KD, Packham J, BSRBR-RA Contributors Group, Symmons DPM, *et al.* Long-term persistence of TNF-inhibitor treatment in patients with psoriatic arthritis. Data from the British Society for Rheumatology Biologics Register. *RMD Open* 2018;**4**. 10.1136/rmdopen-2017-000596

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

55. Joint Formulary Committee. *British National Formulary (online)*. BMJ Group and Pharmaceutical Press; 2020. URL: https://bnf.nice.org.uk/ (accessed 14th October 2020).

56. NHS England. *National Schedule of NHS costs 2018/2019*. URL: https://www.england.nhs.uk/national-cost-collection/ (accessed 14th October 2020).

57. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: Personal Social Services Research Unit, University of Kent; 2019. 10.22024/UniKent%2F01.02.79286

58. McHugh N, Maguire A, Handel I, Tillett W, Morris J, Hawkins N, *et al.* Evaluation of the Economic Burden of Psoriatic Arthritis and the Relationship Between Functional Status and Healthcare Costs. *J Rheumatol* 2020;**47**:701-7. jrheum.190083 [pii]10.3899/jrheum.190083

59. Poole CD, Lebmeier M, Ara R, Rafia R, Currie CJ. Estimation of health care costs as a function of disease severity in people with psoriatic arthritis in the UK. *Rheumatology (Oxford)* 2010;**49**:1949-56. keq182 [pii]10.1093/rheumatology/keq182

9 APPENDICES

9.1 Effectiveness estimates from the NMA models unadjusted for placebo response

Table 42 and Table 43 present PsARC response rates and absolute probabilities of PASI response, respectively, as derived from the company's CODA for the unadjusted random effects NMAs provide in response to points for clarification. These estimates were applied in the ERG scenario analyses presented in Section 6.2 and Section 6.3.

Table 42 Summary of the alternative PsARC response rates used in the model, unadjusted random effects model

	Biologic-naïve
Intervention	Unadjusted RE model
BSC	31.8%
Apremilast	51.2%
Tofacitinib	43.3%
Ustekinumab Mixed[a]	51.2%
Adalimumab	58.7%
Ixekizumab Mixed[b]	58.6%
Guselkumab Q8W	61.4%
Certolizumab Mixed[c]	62.9%
Secukinumab[d]	61.4%
Golimumab	82.2%
Etanercept	79.0%
Infliximab	81.8%

[a] Assumes a mix of two doses: 80% 45mg and 20% 90mg; [b] Assumes a mix of two dosing schedules: 69% Q4W and 31% Q2W; [c] Assumes a mix of two dosing schedules: 50% 200 mg Q2W and 50% 400 mg Q4W; [d] Assumes a mix of two doses: 69% 150 mg and 31% 300 mg; RE, random effects.

Table 43 Summary of alternative PASI response rates used in the model by the ERG

	Biologic naive			
	Unadjusted RE			
	PASI 50	**PASI 75**	**PASI 90**	**PASI 100**
BSC	19.7%	8.1%	2.9%	0.8%
Etanercept	48.5%	29.9%	16.4%	7.8%
Apremilast	45.9%	26.2%	12.9%	5.4%
Golimumab	76.6%	58.3%	39.5%	23.2%
Adalimumab	63.3%	42.2%	24.6%	12.1%
Certolizumab Mixed[c]	54.3%	34.0%	18.6%	8.7%
Tofacitinib	59.5%	38.9%	22.3%	10.9%
Ustekinumab Mixed[a]	73.6%	54.0%	35.0%	19.4%
Infliximab	93.6%	84.0%	69.7%	51.7%
Secukinumab Mixed[d]	76.0%	58.2%	40.0%	24.1%
Ixekizumab Mixed[b]	88.3%	74.9%	57.6%	39.0%
Guselkumab Q8W	78.4%	59.8%	40.3%	23.3%
	Biologic experienced			
	Unadjusted RE			
	PASI 50	**PASI 75**	**PASI 90**	**PASI 100**
BSC	17.6%	6.3%	1.8%	0.4%
Apremilast	43.4%	22.5%	9.5%	3.3%
Certolizumab Mixed[c]	53.0%	31.1%	15.2%	6.2%
Tofacitinib	27.3%	12.2%	4.5%	1.4%
Ustekinumab Mixed[a]	84.2%	68.1%	48.8%	30.4%
Secukinumab 300mg	84.0%	66.2%	45.1%	25.9%
Ixekizumab Mixed[b]	67.0%	45.1%	25.7%	12.2%
Guselkumab Q8W	78.7%	60.1%	39.9%	22.7%

[a] Assumes a mix of two doses: 80% 45mg and 20% 90mg; [b] Assumes a mix of two dosing schedules: 69% Q4W and 31% Q2W; [c] Assumes a mix of two dosing schedules: 50% 200 mg Q2W and 50% 400 mg Q4W; [d] Assumes a mix of two doses: 69% 150 mg and 31% 300 mg; RE, random effects.

9.2 Subgroup results of the ERG scenario analyses

Table 44 Subgroup results of scenario 2 for alternative discontinuation rates – 16.5% (TNFi) and 11.4% (other treatments)

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£59,955	5.406	-	-	-
Etanercept	[redacted]	8.227	[redacted]	2.820	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.546	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Etanercept	[redacted]	8.105	[redacted]	2.840	[redacted]
Guselkumab	[redacted]	8.689	[redacted]	0.583	[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£109,716	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]
bDMARD-experienced – Minimal PsO					
BSC	£61,138	5.431	-	-	-
Guselkumab	[redacted]	7.949	[redacted]	2.518	[redacted]
bDMARD- experienced – Mild to moderate PsO					
BSC	£81,014	4.133	-	-	-
Guselkumab	[redacted]	6.599	[redacted]	2.466	[redacted]
bDMARD- experienced – Moderate to severe PsO					
BSC	£113,227	3.493	-	-	-
Guselkumab	[redacted]	6.184	[redacted]	2.691	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£59,955	5.406	-	-	-
Secukinumab	[redacted]	8.053	[redacted]	2.647	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.719	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Guselkumab	[redacted]	8.689	[redacted]	3.423	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£109,716	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]

Table 45 Subgroup results of scenario 4 for an alternative source of effectiveness data without a placebo response adjustment

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£59,938	5.415	-	-	-
Etanercept	[redacted]	8.303	[redacted]	2.888	[redacted]
Guselkumab	[redacted]	8.575	[redacted]	0.272	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Infliximab	[redacted]	8.369	[redacted]	3.103	[redacted]
Guselkumab	[redacted]	8.288			[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£109,700	4.369	-	-	-
Guselkumab	[redacted]	7.838	[redacted]	3.469	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£59,938	5.415	-	-	-
Guselkumab	[redacted]	8.575	[redacted]	3.160	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£77,361	5.274	-	-	-
Guselkumab	[redacted]	8.488	[redacted]	3.213	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£109,700	4.369	-	-	-
Guselkumab	[redacted]	7.838	[redacted]	3.469	[redacted]

Table 46 Subgroup results of scenario 5 including a cost for the administration of subcutaneous drugs

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£59,955	5.406	-	-	-
Etanercept	[redacted]	8.227	[redacted]	2.820	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.546	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Etanercept	[redacted]	8.105	[redacted]	2.840	[redacted]
Guselkumab	[redacted]	8.689	[redacted]	0.583	[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£109,716	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-experienced – Minimal PsO					
BSC	£61,138	5.431	-	-	-
Guselkumab	£112,389	7.949	[redacted]	2.518	[redacted]
bDMARD- experienced – Mild to moderate PsO					
BSC	£81,014	4.133	-	-	-
Guselkumab	[redacted]	6.599	[redacted]	2.466	[redacted]
bDMARD- experienced – Moderate to severe PsO					
BSC	£113,227	3.493	-	-	-
Guselkumab	[redacted]	6.184	[redacted]	2.691	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£59,955	5.406	-	-	-
Secukinumab	[redacted]	8.053	[redacted]	2.647	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.719	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£77,378	5.266	-	-	-
Guselkumab	[redacted]	8.689	[redacted]	3.423	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£109,716	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]

Table 47 Subgroup results of scenario 6 applying estimates of monitoring resource use consistent with previous TAs in PsA

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£54,024	5.406	-	-	-
Etanercept	[redacted]	8.227	[redacted]	2.820	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.546	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£71,447	5.266	-	-	-
Etanercept	[redacted]	8.105	[redacted]	2.840	[redacted]
Guselkumab	[redacted]	8.689	[redacted]	0.583	[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£103,785	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]
bDMARD-experienced – Minimal PsO					
BSC	£55,207	5.431	-	-	-
Guselkumab	[redacted]	7.949	[redacted]	2.518	[redacted]
bDMARD- experienced – Mild to moderate PsO					
BSC	£75,083	4.133	-	-	-

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
Guselkumab	[redacted]	6.599	[redacted]	2.466	[redacted]
bDMARD- experienced – Moderate to severe PsO					
BSC	£107,296	3.493	-	-	-
Guselkumab	[redacted]	6.184	[redacted]	2.691	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£54,024	5.406	-	-	-
Secukinumab	[redacted]	8.053	[redacted]	2.647	[redacted]
Guselkumab	[redacted]	8.772	[redacted]	0.719	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£71,447	5.266	-	-	-
Guselkumab	[redacted]	8.689	[redacted]	3.423	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£103,785	4.361	-	-	-
Guselkumab	[redacted]	8.058	[redacted]	3.697	[redacted]

Table 48 Subgroup results of the ERG base-case sourcing effectiveness from the placebo response-adjusted model

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£52,583	5.414	-	-	-
Etanercept	[redacted]	8.129	[redacted]	[redacted]	[redacted]
Infliximab	[redacted]	8.181	[redacted]	[redacted]	[redacted]
Guselkumab	[redacted]	7.703	-	-	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£70,018	5.274	-	-	-
Etanercept	[redacted]	8.007	[redacted]	2.733	[redacted]
Infliximab	[redacted]	8.071	[redacted]	0.064	[redacted]
Guselkumab	[redacted]	7.599	-	-	[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£102,356	4.369	-	-	-
Etanercept	[redacted]	7.179	[redacted]	2.810	[redacted]
Infliximab	[redacted]	7.308	[redacted]	0.129	[redacted]
Guselkumab	[redacted]	6.863	-	-	[redacted]
bDMARD-experienced – Minimal PsO					
BSC	£52,562	5.442	-	-	-
Guselkumab	[redacted]	6.685	[redacted]	1.243	[redacted]

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD- experienced – Mild to moderate PsO					
BSC	£72,088	4.143	-	-	-
Guselkumab	£98,421	5.360	£26,333	1.216	£21,648
bDMARD- experienced – Moderate to severe PsO					
BSC	£104,423	3.503	-	-	-
Guselkumab	[redacted]	4.842	[redacted]	1.339	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£52,583	5.414	-	-	-
Secukinumab	[redacted]	7.719	[redacted]	2.304	[redacted]
Guselkumab	[redacted]	7.703	-	-	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£70,018	5.274	-	-	-
Secukinumab	[redacted]	7.609	[redacted]	2.336	[redacted]
Guselkumab	[redacted]	7.599	-	-	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£102,356	4.369	-	-	-
Guselkumab	[redacted]	6.863	[redacted]	2.494	[redacted]

Table 49 Subgroup results of the ERG base-case sourcing effectiveness from the unadjusted model

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
bDMARD-naïve – Minimal PsO					
BSC	£52,569	5.423	-	-	-
Etanercept	[redacted]	8.206	[redacted]	2.783	[redacted]
Infliximab	[redacted]	8.320	[redacted]	0.115	[redacted]
Guselkumab	[redacted]	7.595	-	-	[redacted]
bDMARD-naïve – Mild to moderate PsO					
BSC	£70,004	5.283	-	-	-
Etanercept	[redacted]	8.096	[redacted]	2.814	[redacted]
Infliximab	[redacted]	8.230	[redacted]	0.134	[redacted]
Guselkumab	[redacted]	7.493	-	-	[redacted]
bDMARD-naïve – Moderate to severe PsO					
BSC	£102,342	4.377	-	-	-
Etanercept	[redacted]	7.334	[redacted]	2.957	[redacted]
Infliximab	[redacted]	7.566	[redacted]	0.232	[redacted]
Guselkumab	[redacted]	6.772	-	-	[redacted]

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Fully incremental ICER
					[redacted]
bDMARD-experienced – Minimal PsO					
BSC	£52,562	5.444	-	-	-
Guselkumab	[redacted]	6.687	[redacted]	1.243	[redacted]
bDMARD- experienced – Mild to moderate PsO					
BSC	£71,759	4.150	-	-	-
Guselkumab	[redacted]	5.368	[redacted]	1.217	[redacted]
bDMARD- experienced – Moderate to severe PsO					
BSC	£103,430	3.534	-	-	-
Guselkumab	[redacted]	4.877	[redacted]	1.343	[redacted]
TNFi contraindicated – Minimal PsO					
BSC	£52,569	5.423	-	-	-
Secukinumab	[redacted]	7.563	[redacted]	2.140	[redacted]
Guselkumab	[redacted]	7.595	[redacted]	0.032	[redacted]
Ustekinumab	[redacted]	7.603	[redacted]	0.008	[redacted]
TNFi contraindicated – Mild to moderate PsO					
BSC	£ 70,004	5.283	-	-	-
Secukinumab	[redacted]	7.460	[redacted]	2.177	[redacted]
Guselkumab	[redacted]	7.493	[redacted]	0.033	[redacted]
Ustekinumab	[redacted]	7.496	[redacted]	0.003	[redacted]
TNFi contraindicated – Moderate to severe PsO					
BSC	£102,342	4.377	-	-	-
Guselkumab	[redacted]	6.772	[redacted]	2.395	[redacted]