NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

Contents:

The following documents are made available to consultees and commentators:

The **final scope and final stakeholder list** are available on the NICE website.

1. **Company submission** from **Eli Lilly & Company**

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
    a. British Association of Dermatologists
    The Royal College of Physicians endorses the statement made by the British Association of Dermatologists and will not be submitting anything further
    b. National Eczema Society
    James Pitayanukul, patient expert, contributed to the response from the National Eczema Society

4. **Expert personal perspectives** from:
    a. Dr Richard Weller– clinical expert, nominated by British Association of Dermatologists
    b. Alice Lambert – patient expert, nominated by National Eczema Society
    James Pitayanukul, patient expert, contributed to the response from the National Eczema Society (see 3b) and will not be submitting anything further

5. **Evidence Review Group report** prepared by Centre for Reviews and Dissemination and Centre for Health Economics – York

6. **Evidence Review Group report – factual accuracy check**

7. **Technical report**

8. **Technical engagement response from company**
    a. Technical engagement response from company
    b. Appendix

9. **Technical engagement responses from experts:**
    a. Dr Richard Weller – clinical expert, nominated by British Association of Dermatologists

b. Dr Andrew Pink - clinical expert, nominated by Eli Lilly

10. **Technical engagement responses from consultees and commentators:**
    a. Sanofi
    Pfizer submitted a no comment response
    National Eczema Society submitted a no comment response

11. **Evidence Review Group critique of company response to technical engagement** prepared by Centre for Reviews and Dissemination and Centre for Health Economics York

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Baricitinib for the Treatment of Moderate-to-Severe Atopic Dermatitis [ID1622]

Document B

Company evidence submission

Eli Lilly and Company

June 2020

File name	Version	Contains confidential information	Date
ID1622_Eli Lilly_Baricitinib AD_Document B_Fully Redacted	Final	Yes	23rd June 2020

Contents

Tables and figures

List of Abbreviations

Abbreviation	Definition
AAD	American Academy of Dermatology
ACFB	Absolute Change from Baseline
AD	Atopic Dermatitis
ADSS	Atopic Dermatitis Sleep Scale
AE	Adverse Event
AESI	Adverse Event of Special Interest
ANCOVA	Analysis of Covariance
ATP	Adenosine Triphosphate
BARI	Baricitinib
BMI	Body Mass Index
BNF	British National Formulary
BSA	Body Surface Area
BSC	Best Supportive Care
CHMP	Committee for Human Medicinal Products
COVID	Coronavirus Disease
CRD	Centre for Reviews and Dissemination
CSR	Clinical Study Report
DLQI	Dermatology Life Quality Index
DSA	Deterministic Sensitivity Analysis
DUPI	Dupilumab
EASI	Eczema Area and Severity Index
eGFR	estimated Glomerular Filtration Rate
EMA	European Medicines Agency
EQ5D	EuroQoL 5-Dimensions
FBC	Full Blood Count
FE	Fixed Effects
GCP	Good Clinical Practice
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
GP	General Practitioner
HADS	Hospital Anxiety Depression Scale
HIS	Health Index Score
HRQoL	Health-Related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
IGA	Investigators Global Assessment
IPD	Individual Patient Data
ITC	Indirect Treatment Comparison
ITT	Intention-to-Treat
JAK	Janus-associated kinase

LSM	Least Squares Mean
MCFB	Mean Change from Baseline
MIMS	Monthly Index of Medical Specialities
MMRM	Mixed-Effect Model Repeated Measure
NA	Not Applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network Meta-Analysis
NMB	Net Monetary Benefit
NRI	Non-Responder Imputation
NRS	Numeric Rating Scale
PAS	Patient Access Scheme
PBI	Patient Benefit Index
PBO	Placebo
PCFB	Percent Change from Baseline
PDE	Phosphodiesterase
PGI-S-AD	Patient Global Impression of Severity
POEM	Patient Orientated Eczema Measure
PPD	Purified Protein Derivative
PPS	Per-Protocol Set
PROMIS	Patient-Reported Outcome Measurement Information System
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PUVA	Psoralen-Ultraviolet B
Q2W	Every 2 Weeks
Q4W	Every 4 Weeks
Q8W	Every 8 Weeks
QALY	Quality-Adjusted Life Year
QC	Quality Control
RCT	Randomised Controlled Trial
ROW	Rest of World
SAE	Serious Adverse Event
SCORAD	SCORing Atopic Dermatitis
SD	Standard Deviation
SLR	Systematic Literature Review
SmPC	Summary of Product Characteristics
SP	Safety Population
STAT	Signal Transducers and Activators of Transcription
TCI	Topical Calcineurin Inhibitor
TCS	Topical Corticosteroid
TEAE	Treatment-Emergent Adverse Event

URTI	Upper Respiratory Tract Infection
USA	United States of America
USD	United States Dollar
UVB	Ultraviolet-B
VAS	Visual Analogue Scale
VBA	Visual Basic for Applications
WPAI	Work Productivity and Activity Impairment
WTP	Willingness-to-Pay

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 *Decision problem*

The decision problem addressed within this submission is broadly consistent with the NICE final scope for this appraisal. Any differences between the decision problem addressed within this submission and the NICE final scope are outlined in Table 1.

The full anticipated marketing authorisation for baricitinib (Olumiant®) is for the treatment of moderate-to-severe atopic dermatitis (AD) in adult patients who are candidates for systemic therapy. The indication of relevance for this submission focusses on part of the marketing authorisation for baricitinib. The expected eligible patient population for baricitinib in UK clinical practice is moderate-to-severe AD patients who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This population is in line with the clinical positioning of baricitinib in current UK practice and the eligibility criteria for the BREEZE-AD4 (JAIN) trial. In these patients, current treatment options are limited to dupilumab or BSC in patients for whom use of dupilumab is not recommended or contraindicated. This reflects the highest unmet clinical need for an effective, tolerable, easily-administrable treatment option for patients whose only therapeutic alternative is expensive injection-delivered biologics.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with moderate-to-severe atopic dermatitis (AD) who are candidates for systemic therapy that had an inadequate response or intolerance to existing topical treatments.	Adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one systemic immunosuppressant due to intolerance, contraindication or inadequate disease control.	The population considered in this submission is most relevant to UK clinical practice as it is expected that clinicians will use baricitinib after considering a systemic immunosuppressant agent. The eligibility criteria for the BREEZE-AD4 (JAIN) trial aligns with this patient population and is a subgroup of the full license population. Scenario analyses based on the full licensed population have been conducted.
Intervention	Baricitinib with and without corticosteroids	Baricitinib with and without corticosteroids	N/A – in line with the NICE final scope.
Comparator(s)	• Phototherapy including ultraviolet B (UVB) radiation or psoralen-ultraviolet A (PUVA) • Systemic immunosuppressive therapies (azathioprine, ciclosporin, methotrexate and mycophenolate mofetil) • Alitretinoin (in people with AD affecting the hands) • Dupilumab • Best supportive care (BSC)	• Dupilumab • BSC (emollients, low-to-mid potency topical corticosteroids, phototherapy, psychological support, and rescue therapy including higher potency topical or oral corticosteroids or topical calcineurin inhibitors)	The use of baricitinib within UK clinical practice is expected to be 5th line therapy following failure or contraindication of topical therapies, phototherapy and systemic immunosuppressant agents, making dupilumab and BSC the relevant comparators in UK clinical practice. Alitretinoin is not a relevant comparator based on its licenced indication and place in therapy in the treatment of severe chronic hand eczema. This is in line with the dupilumab submission (TA534) which presented a base case comparison with BSC only.[1]
Outcomes	The outcome measures to be considered include: • Measures of disease severity	The outcome measures to be included in the submission include: • Measures of disease severity and	Whilst data for time-to-relapse and disease-free period are not explicitly available, evidence for maintenance of response is available for the

	• Measures of symptom control • Disease-free period/maintenance of remission • Time to relapse/prevention of relapse • Adverse effects of treatment • Health-related quality of life	symptom control (including IGA, EASI scores, Itch NRS, Skin pain NRS) • Adverse effects of treatment (including AEs, SAEs, AESIs) • Health-related quality of life (including EQ-5D-5L, DLQI, POEM, HADS, ADSS, WPAI-AD) • Maintenance of response (including IGA, EASI scores, Itch NRS, Skin pain NRS and HRQoL outcomes)	population of interest from BREEZE-AD4 (response rate at 24 weeks).
Subgroups to be considered	If the evidence allows the following subgroups will be considered. These include: • skin colour subgroups, • people with moderate dermatitis and those with severe dermatitis, and • people who are ciclosporin naïve and those who have previously received ciclosporin.	The subgroups specified in the NICE final scope were not considered in this submission.	Data were not available to conduct subgroup analyses for skin colour subgroups. The patient population considered in the submission will be adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. As such, all patients can be considered to have moderate-to-severe AD, since systemic therapies are not considered until failure of topical treatments, phototherapy and photochemotherapy (psoralen-ultraviolet A [PUVA]). However, the clinical classification systems used to define AD severity are not consistent, with patients often receiving highly individualised treatment, and therefore defining separate subgroups of moderate AD

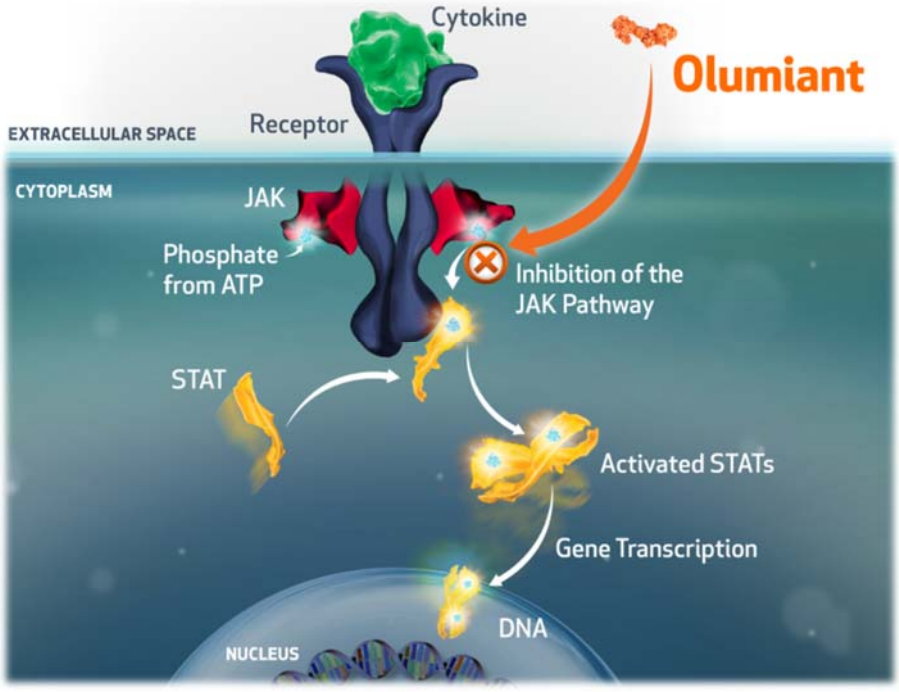
			and severe AD was not considered plausible or possible. In the patient population considered in the submission who have experienced failure with or are intolerant to or have contraindication to at least 1 systemic therapy, the vast majority of these patients will have received prior ciclosporin as ciclosporin is currently the only licensed systemic immunosuppressant for AD. Therefore, subgroup analyses based on ciclosporin-naivety was not considered relevant to the submission.
Special considerations including issues related to equity or equality	None identified	N/A – in line with the NICE final scope.	N/A – in line with the NICE final scope.

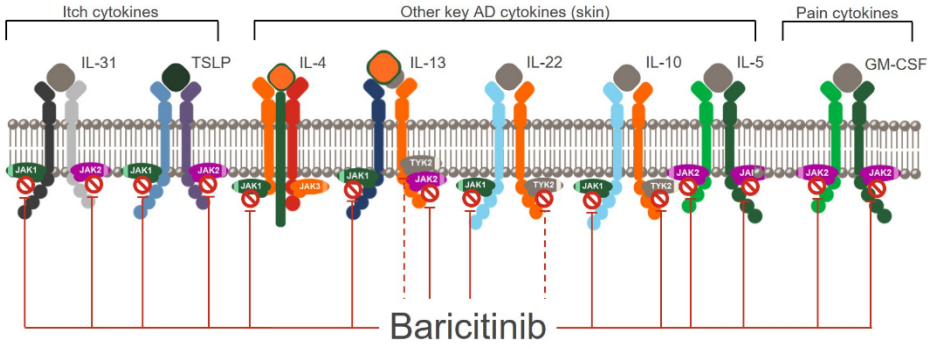
Abbreviations: AD: atopic dermatitis; ADSS: Atopic dermatitis sleep scale; AE: adverse event; AESI: adverse event of special interest; BSA: body surface area; BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: the European Quality of Life-5 Dimensions 5 Levels; HADS: Hospital Anxiety Depression Scale; IGA: Investigator's Global Assessment; NRS: numeric rating scale; POEM: Patient Orientated Eczema Measure; PUVA: psoralen-ultraviolet B; SAE: serious adverse event; UVB: ultraviolet B; WPAI-AD: work productivity and activity impairment questionnaire: atopic dermatitis.

B.1.2 *Description of the technology being appraised*

A summary of the mechanism of action, marketing authorisation status, costs and administration requirements of baricitinib in the treatment adult atopic dermatitis is presented in Table 2.

Table 2: Technology being appraised

UK approved name and brand name	Baricitinib (Olumiant®)
Mechanism of action	Baricitinib is an orally available small molecule that acts selectively and reversibly to inhibit the JAK family of protein tyrosine kinases, specifically JAK1 and JAK2. These enzymes mediate pathways involved in the inflammatory processes underlying AD (Figure 1). The Janus-associated kinase (JAK) signalling pathway mediates cellular responses to many cytokines and growth factors via a cascade of activation initiated at the cell surface: ligand-receptor interaction activates the JAKs, which in turn leads to phosphorylation and activation of signal transducers and activators of transcription (STATs), which translocate to the nucleus to mediate target gene regulation.[2] In this way, JAK-STAT signalling is involved in the dysregulated immune responses observed in AD including the exaggeration of Th2 cell response, the activation of eosinophils, the maturation of B cells, and the suppression of regulatory T cells (Tregs). By inhibiting this signalling, baricitinib modulates the intracellular signalling of multiple cytokines involved in AD (Figure 2). The JAK-STAT pathway activated by IL-4, plays a critical role in the pathogenesis of AD by upregulating epidermal chemokines, pro-inflammatory cytokines and pro-angiogenic factors and by downregulating antimicrobial peptides (AMPs) and factors responsible for skin barrier function.[2] In addition, the JAK-STAT pathway is activated by IL-31, which is thought to be the key causative factor for itch in AD.[2] **Figure 1: The JAK-STAT signalling pathway and its inhibition by baricitinib (Olumiant®)**  **Abbreviations:** ATP: adenosine triphosphate; JAK: Janus-associated kinase; STAT: signal transducers and activators of transcription.

	Source: olumiant.com[3] **Figure 2: Baricitinib modulates the intracellular signalling of multiple cytokines involved in atopic dermatitis pathogenesis**  **Abbreviations:** AD: atopic dermatitis; GM-CSF: granulocyte macrophage colony-stimulating factor; IL: interleukin.
Marketing authorisation/ CE mark status	Marketing authorisation for baricitinib in AD from the European Medicines Agency (EMA) is expected in [redacted] and positive opinion from the Committee for Human Medicinal Products (CHMP) is expected in [redacted].
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Baricitinib already holds a marketing authorisation in rheumatoid arthritis which has previously been appraised by NICE.[4] The anticipated indication for baricitinib following EMA marketing authorisation is "for the treatment of moderate to severe atopic dermatitis in adult patients who are candidates for systemic therapy". **Contraindications included in the draft Summary of Product Characteristics (SmPC) for baricitinib in AD:** • Hypersensitivity to the active substance baricitinib or the following excipients: cellulose, microcrystalline; croscarmellose sodium; magnesium stearate; mannitol; iron red oxide (E172); lecithin (soya) (E322); macrogol; poly (vinyl alcohol); talc; titanium dioxide (E171) • Pregnancy
Method of administration and dosage	Baricitinib is for oral use, taken at any time of day with or without food. It may be used with or without topical corticosteroids or topical calcineurin inhibitors. The recommended dose for AD patients is 4 mg once daily. An optional down-titration dose of 2 mg is appropriate for some patients, such as those aged 75 years or older, and may be appropriate for patients with a history of chronic or recurrent infections. The efficacy of baricitinib can be enhanced when given with TCS.
Additional tests or investigations	No additional tests or investigations are required.
List price and average cost of a course of treatment	The list price of a 28-tablet pack of 2 mg or 4 mg baricitinib is £805.56 and the average annual cost of a baricitinib treatment course is £10,508.24.[5]
Patient access scheme (if applicable)	Baricitinib currently has a Patient Access Scheme (PAS) of [redacted]% discount off the list price in the UK. With the PAS, the pack price of baricitinib is £[redacted] and the average annual cost of a baricitinib treatment course is £[redacted].

	A revised PAS discount of [redacted]% will take effect following positive recommendation in this population. With this revised PAS, the pack price of baricitinib is £[redacted] and the average annual cost of a baricitinib treatment course is £[redacted].

Abbreviations: AD: atopic dermatitis; CE: European Conformity; CHMP: Committee for Human Medicinal Products; EMA: European Medical Association; JAK: Janus-associated kinase; PAS: patient access scheme; SmPC: summary of product characteristics; TCS: topical corticosteroids.

B.1.3 *Health condition and position of the technology in the treatment pathway*

B.1.3.1 Disease overview

Atopic dermatitis (AD), also known as atopic eczema, is a chronic inflammatory skin condition that affects people of all ages, although it presents most frequently in childhood. Characterised by dry, highly-inflamed skin that appears red and blotchy, AD typically affects the hands, insides of the elbows and backs of the knees, although it can be widespread across the body.[6] This is often accompanied by pruritis, an intense, uncomfortable and unrelenting itching sensation that causes an urge to scratch.[6, 7] AD is the most common form of eczema, with other types including contact dermatitis, seborrheic eczema and varicose eczema.[6] AD is typically an episodic disease where patients experience flares (transient exacerbations of symptoms, occurring as frequently as two or three times each month) and remissions, although in severe cases disease activity may be continuous.[8] These flares can be triggered by a variety of factors including irritants, allergens, hormonal changes and skin infections.[6] The causes of AD remain unclear, but it is significantly more common in individuals with allergies such as hay fever or a family history of the disease, with monozygotic twin pairs showing significantly stronger concordance for atopic dermatitis than dizygotic twin pairs.[9, 10]

B.1.3.2 Disease burden

Epidemiology

The prevalence of AD in adults in the UK has been reported as 2.5%.[11] Assuming 69% of patients have been diagnosed and are receiving treatment, it has been estimated that around 56,187 adults in England have moderate-to-severe AD, representing 7% of diagnosed AD patients.[12] Of these patients, it is estimated that approximately 15,170 patients (27%) are eligible for systemic therapy, of whom approximately 8,040 (53%) have treatment failure or a contraindication to systemic therapies and thus would be eligible for treatment with baricitinib.[12]

Symptoms of AD

Pruritis is the primary source of morbidity in patients with AD, often worsening at night leading to over a quarter of adult AD patients reporting considerable sleep disturbances due to severe itching.[13-15] Pruritis results in scratching, with approximately 36% of AD patients reporting that they often or always scratch until their skin bleeds.[16] This can exacerbate itching via increased inflammation and allergen exposure in a common feedback loop known as the “itch-scratch cycle”.[7, 14] Intense scratching leads to skin pain, with over 40% of AD patients reporting skin pain

in the last week, and compromises the skin barrier, resulting in an increased risk of viral and bacterial infections, such as *Staphylococcus aureus*.[17, 18]

Aside from contributing to further skin pain, skin infections may become systemic, which can lead to serious complications and comprise one of the most common reasons for hospitalisation of AD patients.[18, 19] A study from Denmark suggests that AD patients who are hospitalised may have an increased risk of death as compared to the general population, with a life expectancy reduced by 8.3 years.[20] An important co-morbidity of AD is the development of other atopic conditions such as food allergy, allergic rhinitis or asthma often following AD development in a common sequence known as the atopic march.[21] Patients report the excessive skin dryness and redness as further burdensome symptoms.[13]

Health-related quality of life in patients with AD

Alongside physical comorbidities, AD patients often encounter significant psychosocial impacts. For example, pruritis is associated with impaired health-related quality of life (HRQoL), through reduced sleep quality and latency as well as stress and depression.[22, 23] Half of adults with AD report that their condition significantly limits their lifestyle, with 39% reported avoidance of social interactions due to their appearance.[13, 23] Adult AD is associated with depression, anxiety and suicide ideation.[24]

The impact of disease severity

In clinical practice, AD severity is classified as mild, moderate or severe. Mild AD affects small areas of the skin and presents sporadic itch whilst moderate-to-severe AD covers a larger body surface area and presents with more intense and persistent itch.[25] Patients with moderate to severe AD, estimated to comprise 5% of AD patients in the UK, have a higher disease burden than those with mild AD, reporting more itchiness, skin pain and comorbidities including anxiety and depression alongside worse sleep.[12, 26] Over half of adults with moderate-to-severe AD report inadequate disease control, with 75% of patients reporting that being able to control their AD effectively would be the single most important improvement in their quality of life.[27, 28]

Flares

During flares, patients experience a transient exacerbation of symptoms which has a significant effect on HRQoL. Flare represents a significant burden for adult patients with moderate-to-severe AD who experience approximately 10 flares per year, each lasting an average of over 15 days, totalling nearly 40% of days in a year affected by flare.[28] Flares have a large impact on sleep: per flare, adult patients with moderate-to-severe AD reported over 11 nights of sleep to be affected, waking an average of twice a night. During a flare, 87% of these patients reported avoiding at least one everyday activity, such as bathing or swimming, while 51% reported being unhappy or depressed and 45% reported flare to have a significant effect on self-confidence.[28]

Economic burden of AD

Poorly controlled AD is associated with a significant burden on the health care system. In a recent patient survey determining the interaction between severe dermatitis and contact with health care professionals, 70% of patients reported they interact with their general practitioner (GP), practice nurse, allergy or dermatology team at least twice a year. 40% of respondents reported contact at least four times a year, with around 1 in 5 seeing a health care professional more than nine times a year concerning their dermatitis.[29] Moderate-to-severe AD is associated

with greater healthcare resource than mild AD, which is generally manageable in the primary care setting.[30] In England in 2018-2019, atopic dermatitis accounted for 1,094 hospital admissions, of which 402 (36.7%) were emergency admissions, and 1,214 finished consultant episodes.[31] When access to secondary care is necessary for patients with moderate-to-severe AD in the UK, it can be slow: approximately a third of adult patients with severe dermatitis reported having to wait 4–7 months before being seen following a referral, and 1 in 10 had to wait more than 8 months.[29] Given the waiting times and capacity constraints in secondary care, there is an unmet need for effective therapies that have a lower monitoring burden than the current standard of care.

B.1.3.3 Clinical pathway of care

As discussed in Section B.1.3.1, the severity of AD in UK clinical practice can be classified as mild, moderate or severe based on a variety of clinical features. Classification systems are not consistent, in part due to lack of a validated biomarker for disease severity, but AD may be considered moderate-to-severe when one or more of the following criteria are met:[32]

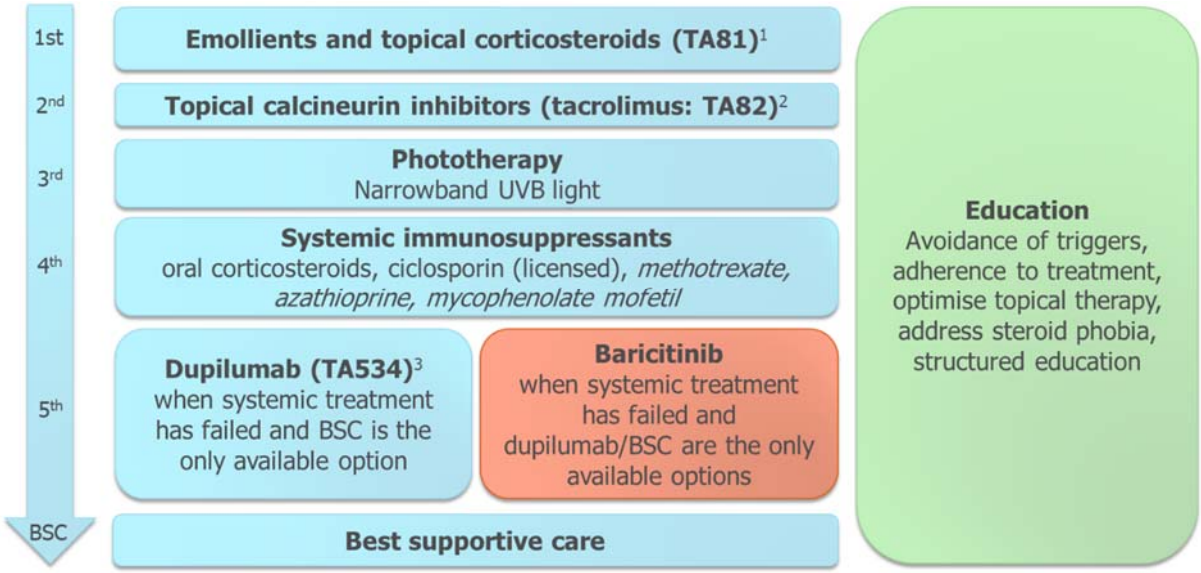
- A minimum involvement of 10% body surface area
- Presence of individual lesions with moderate-to-severe features
- Involvement of highly visible areas or those important for function
- Significantly impaired quality of life

Whilst disease severity is a key consideration, it is not the sole consideration for treatment decisions. The conclusions from the International Eczema Council Guidelines state that “the decision to start a systemic agent should include assessments of severity and quality of life, as well as individual factors such as patient preferences, impact on personal life, prior topical therapy, financial implications and comorbidities”.[33] These conclusions are reflected in current clinical practice in the UK for the management of AD in adults which is highly individualised.

The only available NICE guidelines for the treatment and management of atopic dermatitis in the UK is for patients under the age of 12 years, which contributes further to the individualisation of treatment of adult AD patients in the UK.[34, 35] In adult patients, treatment depends largely on clinician assessment of need, with over 90% of consultant-level dermatologists in a UK-based study reporting their own clinical experience influenced or strongly influenced their choice of treatment for adult patients with moderate-to-severe AD.[36]

AD is a chronic disease with no cure, and thus requires permanent treatment. AD treatment aims to relieve symptoms, prevent flares and improve QoL to maintain daily activities.[37] The current typical treatment pathway and the anticipated placement of baricitinib within it is summarised in Figure 3.

Figure 3: The anticipated positioning of baricitinib in the clinical pathway of care for in the treatment of AD



Abbreviations: AD: atopic dermatitis; BSC: best supportive care; UVB: ultraviolet B
Source: Figure adapted from Simpson EE *et al.* (2017).[33]
1. NICE TA81: Frequency of application of topical corticosteroids for atopic eczema.[38]
2. NICE TA82: Tacrolimus and pimecrolimus for atopic eczema.[39]
3. NICE TA534: Dupilumab for treating moderate to severe atopic dermatitis.[1]

The extensive use of emollients for skin hydration, epidermal repair and pruritis management and behavioural modifications such as the avoidance of environmental triggers are common features at all stages of the treatment pathway.[40] Beyond this, AD is managed in a stepwise approach based on disease severity and activity, with treatment escalation following insufficient disease control.

If symptoms persist following correct use of emollients, topical treatments including topical corticosteroids (TCS) and topical calcineurin inhibitors (TCI) are recommended. Together, emollients, TCS and TCI alongside avoidance of triggers and sufficient patient or carer education are often sufficient to manage mild-to-moderate disease.[8] However, the long-term use of TCS is associated with the risk of adverse events including skin atrophy, skin bleaching and the worsening or spreading of skin infections, making them best suited to short-term use to control disease during flares.[40-42]

The use of phototherapy, commonly narrowband ultraviolet B, or photochemotherapy (psoralen-ultraviolet A [PUVA]) can be recommended as a next-line escalation following a failure in disease control after the use of topical treatments.[36] Phototherapy can be effective in controlling AD, although it is associated with significant limitations, including the need for frequent applications by specialised staff using expensive technical equipment.[43] Alternatively, therapy can be escalated to systemic corticosteroids, but the short- and long-term side effects, including hypertension, glucose intolerance and a reduction in bone density alongside the documented propensity for disease flare following the cessation of their use often limits their clinical benefit.[36, 44]

Beyond oral corticosteroids, the only systemic immunosuppressant therapy currently licensed for AD in the UK is ciclosporin, but other systemic therapies may also be used off-label in UK clinical practice, such as methotrexate, azathioprine and mycophenolate mofetil.[36] However, these

systemic immunosuppressants have poor safety profiles. Ciclosporin use can lead to renal insufficiency, tremor, hypertension and an increased risk of malignancy, particularly of the skin, and other systemic immunosuppressants are associated with a range of common side effects including skin and other malignancies, hepatotoxicity and gastrointestinal intolerance.[33, 45] As a result, a considerable proportion of patients cannot tolerate treatment.

Finally, dupilumab has been recommended by NICE for adults with severe-to-moderate AD who experience failure with, are intolerant to or have contraindication to at least one systemic therapy.[1] Whilst dupilumab is effective in controlling the disease, there are considerable limitations to its use. Dupilumab is administered via subcutaneous injection every other week. Many patients experience injection site reactions, with over 1 in 10 patients experiencing swelling at the site of injection, and more than 1 in 100 reporting redness, pain or itch at the injection site.[46] Eye disorders such as conjunctivitis are also common adverse events of dupilumab treatment. In the CAFÉ trial, 28% patients receiving dupilumab (every other week in combination with topical corticosteroids) experienced conjunctivitis, which was severe in 0.9% and moderate in 12.1% patients.[47] These adverse events result in additional health care resource use through the need for consultant ophthalmologist visits.

If dupilumab fails to control the disease, or in patients for whom use of dupilumab is not recommended or contraindicated, no further safe and effective treatment options are available so patients are treated with best supportive care (BSC) which remains poorly defined in UK clinical practice. These patients often receive short-term oral corticosteroids, TCIs and/or systemic immunosuppressants alongside TCS as required and extensive use of emollients. Patients may also receive education, psychological support, bandages and hospitalisation.

Positioning of baricitinib in the clinical pathway of care

The expected eligible patient population for baricitinib in UK clinical practice is adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This population is in line with the clinical positioning of baricitinib in current UK practice and the eligibility criteria for the BREEZE-AD4 (JAIN) trial. In these patients, current treatment options are limited to dupilumab or BSC in patients for whom use of dupilumab is not recommended or contraindicated making these the relevant comparators considered in this appraisal. This reflects the unmet clinical need for an effective, tolerable, easily-administrable treatment option for patients whose only therapeutic alternative is expensive injection-delivered biologics.

B.1.4 *Equality considerations*

No equality issues related to the use of baricitinib in this indication have been identified or are foreseen.

B.2 Clinical effectiveness

Evidence for Baricitinib in Atopic Dermatitis

- The efficacy of baricitinib for moderate-to-severe atopic dermatitis has been investigated in a series of pivotal RCTs covering the different parts of the treatment pathway, including as monotherapy in patients who had a history of intolerance to topical therapy or of inadequate response to topical or systemic therapy (BREEZE-AD1 [JAHL] and BREEZE-AD2 [JAHM]), as combination therapy with TCS for those who have had inadequate response to topical or systemic medications (BREEZE-AD7 [JAIY]), and in combination with TCS in those with history of an inadequate response to topical medications and an inadequate response, intolerance or were contraindicated to ciclosporin (BREEZE-AD4 [JAIN])

Efficacy

- The primary endpoint of the pivotal BREEZE-AD RCTs was the proportion of patients achieving IGA ≤1 at Week 16 (BREEZE-AD1, -AD2 and -AD7) or the proportion of patients achieving EASI75 at Week 16 (BREEZE-AD4). In the BREEZE-AD3 (JAHN) long-term extension trial, IGA ≤1 was assessed at Weeks 0, 16 and 36 (overall treatment Weeks 16, 32 and 52)
- Across all pivotal BREEZE-AD RCTs (BREEZE-AD1, -AD2, -AD4 and -AD7), treatment with 4 mg baricitinib was associated with a statistically significantly higher proportion of patients achieving the primary outcome at Week 16 as compared with placebo with or without background TCS
- At Weeks 16 through to Week 52 of follow-up in the BREEZE-AD3 (JAHN) long-term extension trial, a higher proportion of patients receiving 4 mg baricitinib monotherapy had maintained this response as compared with placebo
- In the BREEZE-AD1, -AD2 and -AD7 trials, baricitinib treatment with or without background TCS was consistently associated with reduced disease burden as compared with placebo with or without background TCS as determined by secondary outcomes which assessed signs and symptoms of AD and HRQoL outcomes

Safety

- Across all BREEZE-AD trials, no clinically meaningful difference in the overall frequencies of AEs was observed between the placebo and 4 mg baricitinib groups
- Across most BREEZE-AD trials, a higher proportion of patients in the placebo arm reported SAEs than in the 4 mg baricitinib arm
- Numerically more patients in the 4 mg baricitinib group reported AEs necessitating permanent discontinuation from study treatment than in the placebo group across most trials
- No deaths occurred in the placebo or 4 mg baricitinib treatment groups across any of the trials

Indirect Treatment Comparisons

- In the absence of head-to-head clinical trial evidence for baricitinib versus dupilumab, an indirect treatment comparison (ITC) was performed. The results of the ITC indicate that both baricitinib (4 mg QD) monotherapy and in combination with TCS have similar efficacy to dupilumab (300 mg Q2W). In the majority of analyses, the results numerically favoured dupilumab, but differences were often not statistically significant

Innovation

- Baricitinib provides an easily-administered once-daily oral tablet for patients whose only alternative is biweekly subcutaneous injection of dupilumab
- As compared with dupilumab, oral administration of baricitinib will remove the limitations imposed by dupilumab, including removing the risk of injection site-related adverse events and reducing the burden on the healthcare system associated with consultant ophthalmologist visits due to the common side effect of eye disorders including conjunctivitis

Conclusion

- Baricitinib offers patients with moderate-to-severe AD a therapy option with high efficacy and good tolerability that is free from the burden of subcutaneous injection

B.2.1 *Identification and selection of relevant studies*

A systematic literature review (SLR) was conducted to identify all relevant clinical evidence on the efficacy and safety of baricitinib and relevant comparators for the treatment of adults with moderate-to-severe AD. In total, the SLR identified 62 publications reporting on 40 unique studies. Full details of the SLR search strategy, study selection process and results can be found in Appendix D.

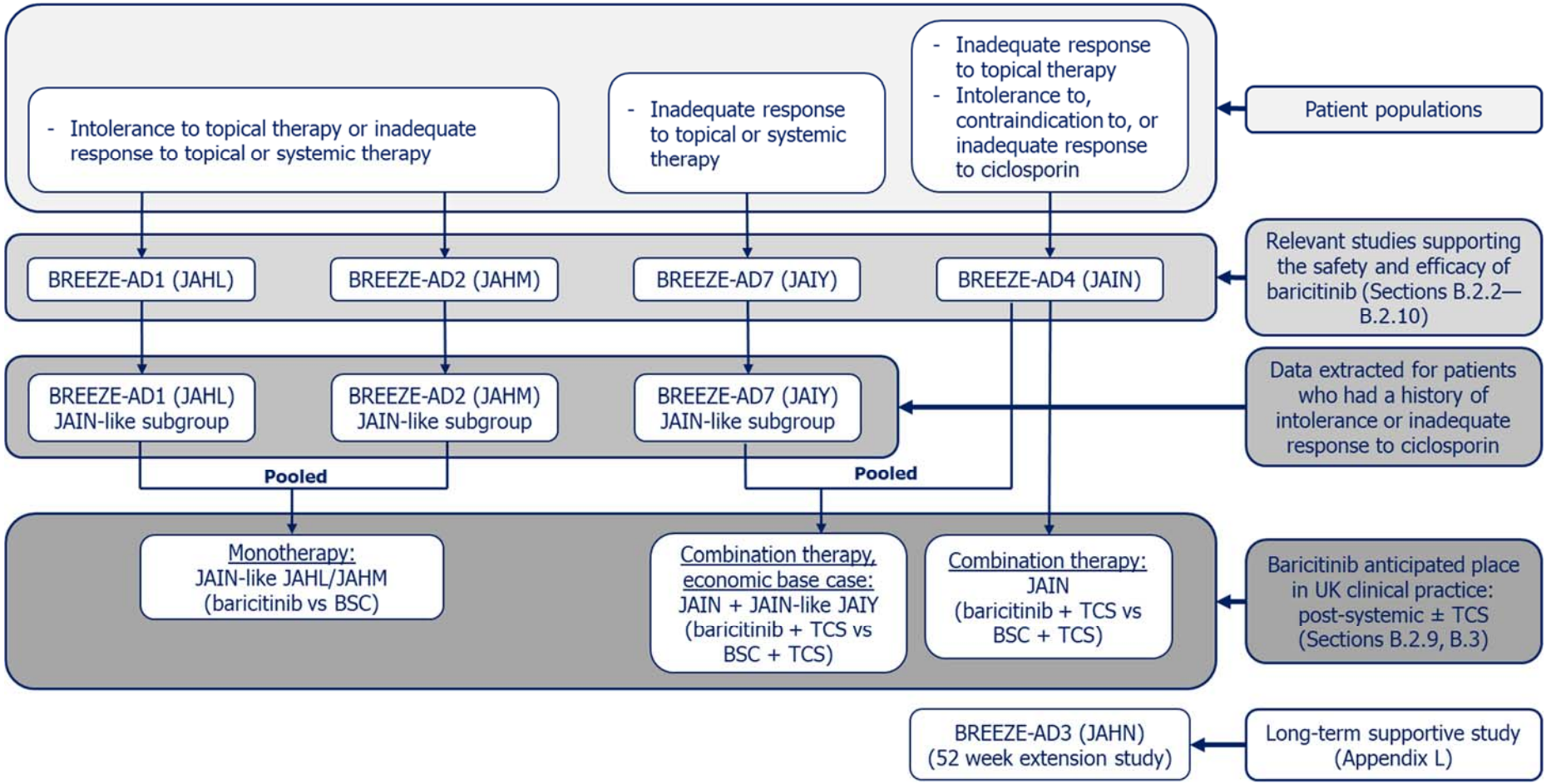
B.2.2 *List of relevant clinical effectiveness evidence*

Two publications were identified in the SLR that provide clinical evidence for the efficacy and safety of baricitinib for the treatment of moderate-to-severe AD. Guttman-Yassky *et al.* (2019)[48] reports a Phase II placebo-controlled trial conducted in the USA and Japan, and Simpson *et al* (2020) reports two randomised monotherapy Phase III RCTs, BREEZE-AD1 (NCT03334396) and BREEZE-AD2 (NCT03334422).[49-51] Evidence for the efficacy and safety of baricitinib in moderate-to-severe AD is provided by two further Phase III clinical trials (BREEZE-AD4 [NCT03428100][52] and BREEZE-AD7 [NCT03733301][53]) and one long-term extension study (BREEZE-AD3 [NCT03334435][34]), which are not yet published. Given the availability of more relevant data from the Phase III trials, the Phase II clinical trial has not been considered further within this submission.

A schematic representation of how these Phase III clinical trials and long-term extension study inform the decision problem and the sections in which they are reported is provided in Figure 4. Overviews of the RCTs and the long-term extension study are provided in Table 3 and Table 4, respectively.

The patient population in BREEZE-AD4 (JAIN) is adult patients with moderate-to-severe AD who are candidates for systemic therapy who experience failure with, are intolerant to or have contraindication to ciclosporin. A proportion of patients in BREEZE-AD7 (JAIY) are also within these criteria. This population is in line with the population of relevance for this submission: adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. For this reason, a pooled population of JAIN + JAIN-like JAIY patients inform the base case economic analysis as shown in Figure 4. Data from the other trials are used to inform scenario analyses in the economic model.

Figure 4: A schematic representation of the BREEZE-AD trials informing the decision problem



Abbreviations: TCS: topical corticosteroids.

Table 3: Summary of clinical effectiveness evidence from the Phase III BREEZE-AD RCTs

Study	BREEZE-AD4 (JAIN) (NCT03428100)[54]	BREEZE-AD7 (JAIY) (NCT03733301)[55]	BREEZE-AD1 (JAHL) (NCT03334396)[56]	BREEZE-AD2 (JAHM) (NCT03334422)[57]
Study design	An international Phase III, multicentre, randomised, double-blind placebo-controlled study			
	N=463	N=[redacted]	N=624	N=615
Population	Adult patients with moderate-to-severe AD (EASI score ≥16, IGA score ≥3 and BSA involvement ≥10% at screening and at randomisation), an AD diagnosis at least 12 months prior according to the American Academy of Dermatology definition, a history of inadequate response to topical therapy and a history of intolerance to, contraindication to, or inadequate response to ciclosporin	Adult patients with moderate-to-severe AD (EASI score ≥16, IGA score ≥3 and BSA involvement ≥10% at screening and at randomisation), an AD diagnosis at least 12 months prior according to the American Academy of Dermatology definition and a history of inadequate response to topical or systemic therapy	Adult patients with moderate-to-severe AD (EASI score ≥16, IGA score ≥3 and BSA involvement ≥10% at screening and at randomisation), an AD diagnosis at least 12 months prior according to the American Academy of Dermatology definition with a history of clinically significant adverse reactions to topical therapy or a history of inadequate response to topical or systemic therapies	
Intervention	Baricitinib once daily (4 mg, 2 mg or 1 mg) in combination with TCS		Baricitinib monotherapy once daily (4 mg, 2 mg or 1 mg)	
Comparator	AD patients randomised from ITT who received TCS only		AD patients randomised from ITT who did not receive baricitinib	
	N=93	N=[redacted]	N=249	N=244
Indicate if trial supports application for marketing authorisation	Yes	Yes	Yes	Yes
Indicate if trial used in the economic model	Yes	Yes	Yes (scenario analysis only)	Yes (scenario analysis only)

Rationale for use in the model	The objective of this trial was to demonstrate efficacy, safety and tolerability of baricitinib with concomitant TCS up to Week 16 in patients with AD who have previously experienced intolerance to or failure with ciclosporin. This population is considered to be the most relevant to UK clinical practice as it is expected that clinicians will use baricitinib after considering a systemic immunosuppressant agent. When pooled with similar patients from the JAIY trial, these patients were used to inform the base case of the economic model.	The objective of this trial was to demonstrate efficacy, safety and tolerability of baricitinib with concomitant TCS up to Week 16 in patients with AD who have a history of inadequate response to topical or systemic therapy. The trial included a subgroup of patients with a history of intolerance to or failure with ciclosporin and these patients were pooled with patients from the JAIN trial to inform the base case of the economic model as they are considered to be the most relevant to UK clinical practice.	These trials included patients with AD who were not permitted to receive concomitant TCS therapy up to Week 16 and included a subgroup of patients with a prior history of immunosuppressant use. These JAIN-like populations were used in the modelling as a scenario analysis to investigate cost-effectiveness in monotherapy patients.

Reported endpoints specified in the decision problem	Measures of disease severity and symptom control: • IGA • **EASI** • SCORAD • Itch NRS • ADSS • Skin Pain NRS • BSA • POEM score • Skin infections requiring antibiotics • TCS use (days and quantity) • PGI-S-AD HRQoL: • **EQ-5D-5L** • HADS • **DLQI** • WPAI-AD **Safety outcomes**	Measures of disease severity and symptom control: • IGA • **EASI** • SCORAD • Itch NRS • ADSS • Skin Pain NRS • BSA • POEM • PROMIS Itch scores • TCS use (days and quantity) • PBI • Skin infections requiring antibiotics • PGI-S-AD HRQoL: • HADS • **DLQI** • **EQ-5D-5L** • WPAI-AD • Neuro-QoL **Safety outcomes**	Measures of disease severity and symptom control: • IGA • **EASI** • SCORAD • Itch NRS • ADSS • Skin Pain NRS • POEM • BSA • Nocturnal sleep-wake • Skin infections requiring antibiotics • PGI-S-AD HRQoL: • HADS • **DLQI** • **EQ-5D-5L** • WPAI-AD Safety outcomes

Endpoints marked in bold are used in the economic model.
Abbreviations: AD: atopic dermatitis; ADSS: atopic dermatitis sleep scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; IGA: Investigator's Global Assessment; ITT: intent-to-treat; NRS: numeric rating scale; PBI: patient benefit index; PGI-S-AD: patient global impression of severity; POEM: patient-oriented eczema measure; PROMIS: patient-reported outcome measurement information system; QoL: quality of life; RCT: randomised controlled trial; SCORAD: SCORing atopic dermatitis; SF-36: medical outcomes study 36-item short form health survey; TCS: topical corticosteroids; WPAI-AD: work productivity and activity impairment questionnaire – atopic dermatitis.
Source: BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report,[57] BREEZE-AD4 (JAIN) Clinical Study Report,[54] BREEZE-AD7 (JAIY) Clinical Study Report,[55] Simpson *et al,* 2020,[49] Bieber *et al*, 2020[58]

Table 4: Summary of clinical effectiveness evidence from the long-term extension study

Study	BREEZE-AD3 (JAHN) (NCT03334435)[59]
Study design	An international Phase III, multicentre, long-term extension study (N=[redacted][a]) Included an open-label addendum (N=[redacted])
Population	Main phase: AD patients who completed BREEZE-AD1, -AD2 or -AD7. Open-label addendum: Adult patients with moderate-to-severe AD (EASI score ≥16, IGA score ≥3 and BSA involvement ≥10% at screening and at randomisation), an AD diagnosis at least 12 months prior according to AAD definition and a history of inadequate response to, or intolerance to, topical therapy.
Intervention	Baricitinib monotherapy once daily (4 mg, 2 mg or 1 mg) Open-label addendum: 2 mg baricitinib monotherapy once daily
Comparator	NA
Indicate if trial supports application for marketing authorisation	Yes
Indicate if trial used in the economic model	No
Rationale for use/non-use in the model	BREEZE-AD3 is a non-randomised, long-term extension study for baricitinib in AD comprised of patients who completed one of the originating RCTs: BREEZE-AD1, -AD2 or -AD7. These patients have received a variety of treatments for differing durations and at the time of submission, only interim aggregated analyses are available. Therefore, it is not possible to conclude the long-term maintenance of response of patients treated with baricitinib from these data at this time so this trial is excluded from the economic analyses. For completeness, its methodology, efficacy and safety data are summarised in Appendix M.
Reported endpoints specified in the decision problem	Disease-free period/maintenance of remission: • IGA • **EASI** • Itch NRS • Time to retreatment Safety outcomes

[a] This number represents only those patients enrolled from BREEZE-AD1 and -AD2. No patients from the BREEZE-AD7 (JAIY) trial were included in the analyses presented in this submission but will be included in future data cuts of BREEZE-AD3 (JAHN).
Abbreviations: AD: atopic dermatitis; AAD: American Academy of Dermatology; BSA: body surface area; EASI: Eczema Area and Severity Index; IGA: Investigator's Global Assessment; NA: not applicable; NRS: numeric rating scale; RCT: randomised controlled trial.
Source: BREEZE-AD3 (JAHN) Clinical Study Report.[59]

B.2.3 *Summary of methodology of the relevant clinical effectiveness evidence*

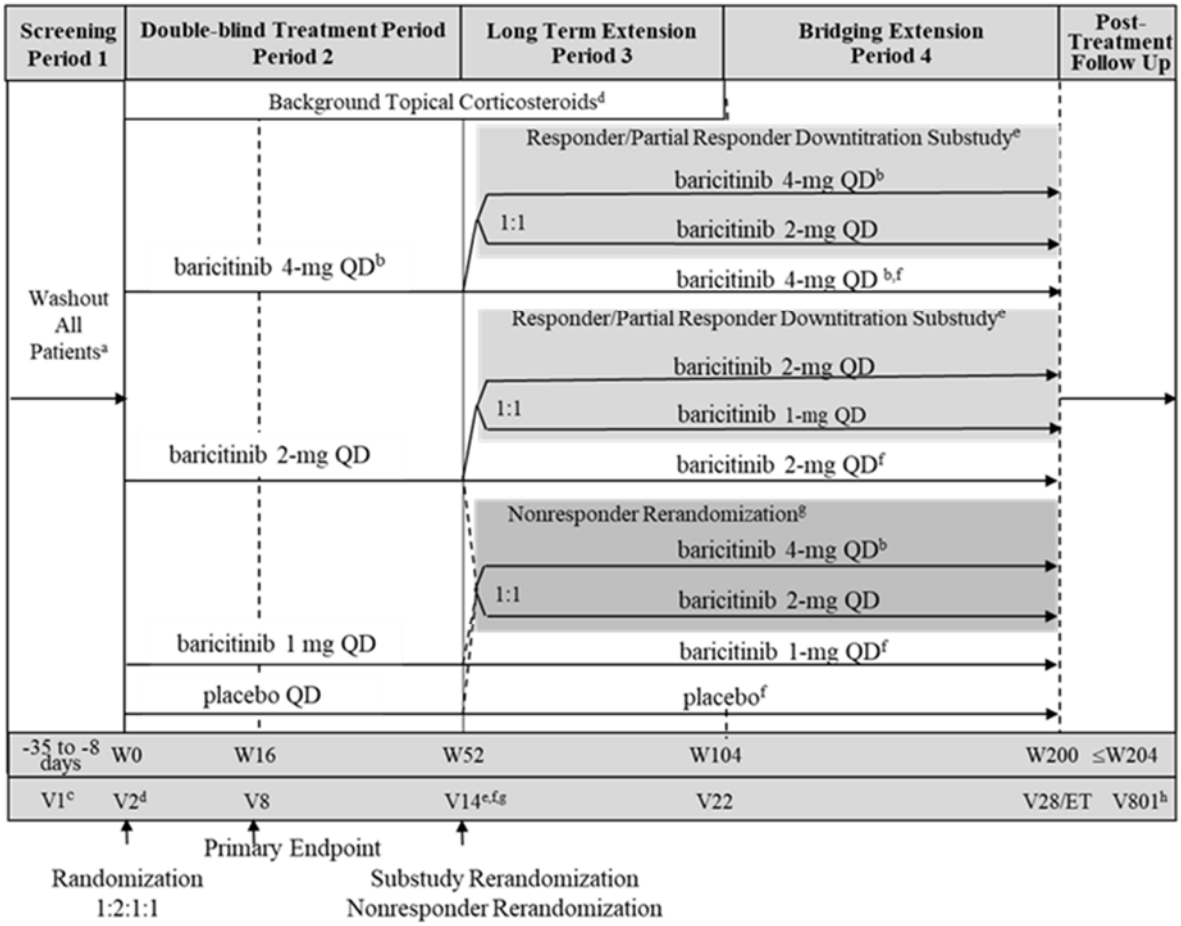
B.2.3.1 Trial design and methodology

The trial design and methodology of the BREEZE-AD RCTs are summarised below. The monotherapy trials BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) are considered together given their identical design.

BREEZE-AD4 (JAIN)

BREEZE-AD4 (JAIN) is an ongoing multicentre, double-blind, randomised, placebo-controlled Phase III study to determine the safety and efficacy of baricitinib in combination with TCS in adult patients with moderate-to-severe AD. As part of the eligibility criteria, patients were required to have a documented history of inadequate response to topical treatment and a documented history of failed ciclosporin treatment, defined as an inadequate response following its administration, or a documented contraindication, intolerance or unacceptable toxicity to its use. The use of concomitant medications for the management of AD was prohibited throughout the trial except for daily emollient use and topical treatments including TCS and topical calcineurin inhibitors which were used as background treatment throughout the trial period. Following a washout period of 5 half-lives for biologic AD treatments, 4 weeks for systemic AD treatments and 2 weeks for topical AD treatments (including TCS), excluding emollients, patients were randomised in a double-blinded fashion to once daily treatment with placebo, 1 mg baricitinib, 2 mg baricitinib or 4 mg baricitinib (1:1:2:1) alongside background TCS. The double-blind 52-week treatment period was followed by a 52-week double-blind long-term extension which included a randomised down-titration sub-study for responders and re-randomisation for non-responders. The study design of BREEZE-AD4 (JAIN) is shown in Figure 5.[54]

Figure 5: Study design of the BREEZE-AD4 (JAIN) trial



[a] Applicable to patients taking topical treatments (excluding emollients) or systemic treatments for AD at the time of screening. [b] Maximum dose of baricitinib for patients with renal impairment (defined as eGFR <60 mL/min/1.73 m^2) was 2 mg QD. [c] Patients for whom PPD skin test for the evaluation of TB infection was performed at V1 had to return and PPD test was read 48–72 hours after V1 (post-PPD). [d] At Visit 2 (Week 0), patients were supplied with mild- and moderate-potency TCS to be applied throughout the trial. [e] At Week 52, responders (IGA 0 or 1) and partial responders (IGA 2) who were assigned to baricitinib 4 mg or 2 mg at randomisation were enrolled into the down-titration study only if they did not have interrupted study drug at the time and had not used high- or ultra-high-potency TCS in the previous 14 days. If a patient in the sub-study had an IGA ≥3 at any time, they were retreated with their pre-sub-study baricitinib dose for the remainder of the study. [f] At Week 52, those who were in the baricitinib 1 mg or placebo groups and responders (IGA 0 or 1) and partial responders (IGA 2) in the baricitinib 4 mg or baricitinib 2 mg treatment groups who were not eligible for the randomised down-titration sub-study remained on their current dose of IP. If a patient had an IGA ≥3 at any time, except for patients in the baricitinib 4 mg group, they were rerandomised automatically at a 1:1 ratio to baricitinib 2 mg QD or baricitinib 4 mg QD. Re-randomisation occurred only once. Patients in the baricitinib 4 mg group remained on 4 mg. [g] Beginning at Visit 14 (Week 52), non-responders (IGA ≥3) in the placebo, baricitinib 1 mg or baricitinib 2 mg treatment groups were rerandomised at a 1:1 ratio to baricitinib 4 mg QD or baricitinib 2 mg QD. Non-responders randomised to baricitinib 4 mg at baseline remained on 4 mg. After re-randomisation, patients remained on the same dose of baricitinib for the remainder of the study. [h] Occurred approximately 28 days after the last dose of IP.

Abbreviations: AD: atopic dermatitis; eGFR: estimated glomerular filtration rate; ET: early termination; IGA: Investigator's Global Assessment; IP: investigational product; PPD: purified protein derivative; QD: once daily; TB: tuberculosis; TCS: topical corticosteroids; V: visit; W: week.

Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

The primary outcome of the trial was to evaluate whether 2 mg or 4 mg baricitinib in combination with TCS is superior to placebo in combination with TCS as measured by the proportion of patients in each treatment group achieving EASI75 at Week 16. Key secondary endpoints measured improvement in signs and symptoms of AD at Week 16 or Week 24 and included IGA,

EASI and SCORAD scores and patient-reported outcomes (ADSS Item 2 score, Itch NRS and Skin Pain NRS). Other secondary endpoints included HRQoL outcomes such as the Dermatology Life Quality Index (DLQI) and EQ-5D-5L. Safety outcomes included AEs, SAEs and TEAEs by Week 16 and by Week 24.

A summary of the methodology of the BREEZE-AD4 (JAIN) trial is presented in Table 5.

Table 5: Summary of BREEZE-AD4 (JAIN) trial methodology

Trial name	**BREEZE-AD4 (JAIN)[60, 61]**
Location	International: patients recruited from 103 sites across 14 countries (5 sites in Austria, 4 sites in Belgium, 9 sites in Brazil, 3 sites in Finland, 10 sites in France, 17 sites in Germany, 10 sites in Italy, 11 sites in Japan, 3 sites in the Netherlands, 10 sites in Poland, 5 sites in the Russian Federation, 7 sites in Spain, 3 sites in Switzerland, 6 sites in UK)
Trial design	Phase III, multicentre, double-blind, randomised, placebo-controlled study
Eligibility criteria for participants	A summary of the criteria for baseline inclusion in the main study are provided below. Full details of the eligibility criteria are presented in Appendix D. Key inclusion criteria: • Aged 18 years or older • Diagnosis of AD at least 12 months prior to screening as defined by the American Academy of Dermatology • Moderate-to-severe AD, defined as having all of the following at the screening visit *and* at randomisation: o EASI score greater than or equal to 16 o IGA score greater than or equal to 3, and o BSA involvement greater than or equal to 10% • Documented history of inadequate response to topical medication within 6 months prior to screening defined as an inability to achieve good disease control after use of a medium potency TCS for at least 4 weeks, or for the maximum duration recommended by the product prescribing information, whichever is shorter • Documented history of: o A medical contraindication to ciclosporin (due to hypersensitivity, a medication condition, use of prohibited concomitant medications or increased susceptibility to ciclosporin-induced renal or liver damage, or increased risk of serious infection) o Intolerance to and/or unacceptable toxicity to ciclosporin o Inadequate response to ciclosporin, defined as a failure to obtain good disease control within 6 weeks of treatment with 2.5-5 mg/kg/day ciclosporin, requirement for >5 mg/kg/day ciclosporin, or a requirement for ciclosporin use for more than 1 year Key exclusion criteria: • Currently experiencing, or have a history of, other concomitant skin conditions, including psoriasis or lupus erythematosus, which would interfere with evaluation of the effect of the study medication on AD, or which requires frequent hospitalisation and/or intravenous treatment for skin infections • Have an important side effect to TCS (e.g. intolerance to treatment or hypersensitivity reactions) which would prevent further use • Eczema herpeticum within 12 months prior to screening or more than twice in the past

	• Any serious concomitant illness anticipated to require the use of systemic corticosteroids or require active frequent monitoring At the end of the main study period (Week 52), patients were re-evaluated for suitability for inclusion in a randomised down-titration sub-study. To be eligible for treatment down-titration, patients must have met the following criteria: • IGA 0 or 1 (responder) or 2 (partial responder) at Week 52 • No use of high- or ultra-high potency TCS in the last 14 days • No study drug interruptions • Assigned to 2 mg or 4 mg baricitinib treatment group at baseline
Method of study drug administration	Administered orally once daily as 3 tablets: 2 placebo tablets with 1 treatment tablet for treatment groups, or 3 placebo tablets for the placebo group.
Permitted and disallowed concomitant medication	A summary of the key concomitant medications permitted and disallowed during the study period is provided below. Full details are provided in Appendix D. All concomitant therapies for AD were prohibited throughout the trial except for: • Daily use of emollients, excluding additives like antipruritics or antiseptics • Background TCS therapy with moderate-potency and/or low-potency TCS (e.g. triamcinolone 0.1% cream and/or hydrocortisone 2.5% ointment) used on active lesions. High- or ultra-high potency TCS permitted only as rescue therapy • TCIs (e.g. tacrolimus and pimecrolimus), or topical PDE-4 inhibitor (i.e., crisaborole, where approved) are permitted in place of TCS on areas where application of TCS is considered inappropriate by the investigator; use should be limited to problem areas (e.g., face, neck, skin folds, genital areas, etc.) • Prescription sleep medications used at entry • A single intra-articular or soft tissue corticosteroid injection until the Week 16 primary endpoints, after which these injections are permitted. • Intranasal or inhaled steroids • Topical anaesthetics and topical and systemic anti-infective medications • Non-live seasonal vaccines and/or emergency vaccinations • Antihistamine ophthalmic preparations • Non-sedating antihistamines only during Phase 2 and all antihistamines in Phases 3 and 4 were permitted
Primary outcome	Proportion of patients in the ITT population achieving EASI75 at Week 16 of treatment.
Secondary and exploratory outcomes	A summary of the key secondary outcomes is provided below. Full details of all the secondary outcomes can be found in Appendix D. • Improvement in signs and symptoms at Week 16: o EASI75 o EASI90 o Percent change in EASI score o SCORAD75 • Improvement in signs and symptoms at Week 24: o IGA of 0 or 1 with a ≥2-point improvement o EASI75 • Patient-reported outcome measures at Week 16: o 4-point improvement in Itch NRS at Week 1, 2, 4 and 16 of treatment

	○ Mean change in Item 2 of ADSS score at Week 1 or 16 of treatment ○ Mean change from baseline in Skin Pain NRS at Week 16 of treatment • HRQoL outcomes at Week 16: ○ DLQI ○ EQ-5D-5L
Pre-planned subgroups	• Gender • Age group (<65, ≥65, ≥65 to <75, ≥75 to <85, ≥85 years old) • Baseline weight (<60, ≥60 to <100, ≥100 kg) • Baseline BMI (<25, ≥25 to <30, ≥30 kg/m^2) • Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple) • Ethnicity (Hispanic, non-Hispanic) • Baseline renal function status: impaired (eGFR <60 mL/min/1.73m^2) or not impaired (eGFR ≥60 mL/min/1.73m^2) • Region (Europe, Japan, rest of world) • Specific regions (Europe, other) • Specific country (Japan, other) • Prior TCI use (Yes/No) • Prior systemic therapy use (Yes/No) • Baseline disease severity (IGA 3 or 4)
Duration of study and follow-up	The total study duration was 113 weeks, with a 5-week screening period, a 52-week treatment period, a 52-week long-term extension period, and 4-week post-treatment follow-up period.

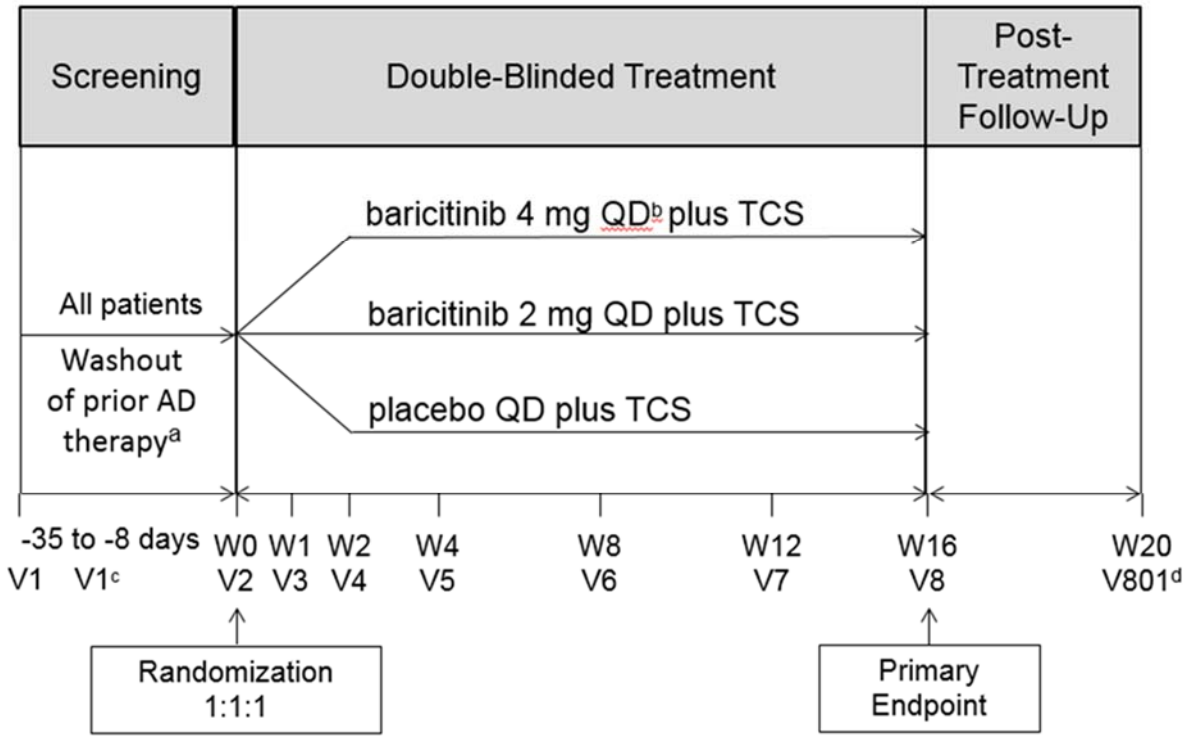
Abbreviations: AD: atopic dermatitis; ADSS: atopic dermatitis sleep scale; BMI: body mass index; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HRQoL: health-related quality of life; IGA: Investigator's Global Assessment; ITT: intent-to-treat; NRS: numeric rating score; PDE-4: phosphodiesterase-4; SCORAD: SCORing Atopic Dermatitis; TCI: topical calcineurin inhibitor; TCS: topical corticosteroids.
Sources: BREEZE-AD4 (JAIN) Clinical Protocol,[60] BREEZE-AD4 (JAIN) Statistical Analysis Plan.[61]

BREEZE-AD7 (JAIY)

BREEZE-AD7 (JAIY) was a multicentre, randomised, double-blind, placebo-controlled Phase III trial to determine the efficacy and safety of baricitinib in combination with TCS in adult patients with moderate-to-severe AD. As part of the eligibility criteria, patients were required to have moderate-to-severe AD with a documented history of an inadequate response to, or intolerance to, topical medication. As for BREEZE-AD1 and -AD2, these eligibility criteria are broader than the population of relevance for this submission. However, for completeness, these trials have been summarised in full within this submission.

The use of concomitant medications for the management of AD was prohibited throughout the trial except for daily emollient use and background TCS therapy, or topical calcineurin inhibitors (TCIs) or topical phosphodiesterase-4 (PDE-4) inhibitors where TCS therapy is considered inappropriate by the investigator. Following a washout period of 4 weeks for systemic AD treatments and 2 weeks for topical AD treatments, excluding emollients, all patients were randomised 1:1:1 in a double-blinded fashion to once daily treatment with placebo, 2 mg baricitinib or 4 mg baricitinib. The study design of BREEZE-AD7 (JAIY) is shown in Figure 6.

Figure 6: Study design of the BREEZE-AD7 (JAIY) trial



[a] Applicable to patients taking topical treatments (excluding emollients) or systemic treatments for AD at the time of screening. [b] For patients randomised to the 4 mg QD dose who had renal impairment (defined as eGFR <60 mL/min/1.73 m^2), the baricitinib dose was 2 mg QD. [c] Patients for whom PPD skin test for the evaluation of tuberculosis infection was performed at V1 had to return and PPD test was read 48–72 hours after V1 (post-PPD). [d] Occurred approximately 28 days after the last dose of the study treatment (was not required for those patients entering the long-term extension Study JAHN).
Abbreviations: AD: atopic dermatitis; eGFR: estimated glomerular filtration rate; PPD: purified protein derivative; QD: once daily; V: visit; W: week.
Sources: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

The primary outcome of the trial was to evaluate the efficacy of baricitinib in combination with TCS as measured by the proportion of patients in each treatment group achieving IGA ≤1 with a ≥2-point improvement at Week 16. Key secondary endpoints measured improvement in signs and symptoms of AD by Week 16 and included EASI and SCORAD scores and patient-reported outcomes (ADSS Item 2 score, Itch NRS and Skin Pain NRS). Other secondary endpoints included HRQoL outcomes such as the Dermatology Life Quality Index (DLQI) and EQ-5D-5L. Safety outcomes included AEs, SAEs and TEAEs by Week 16.[55]

A summary of the methodology of the BREEZE-AD7 (JAIY) trial is presented in Table 6.

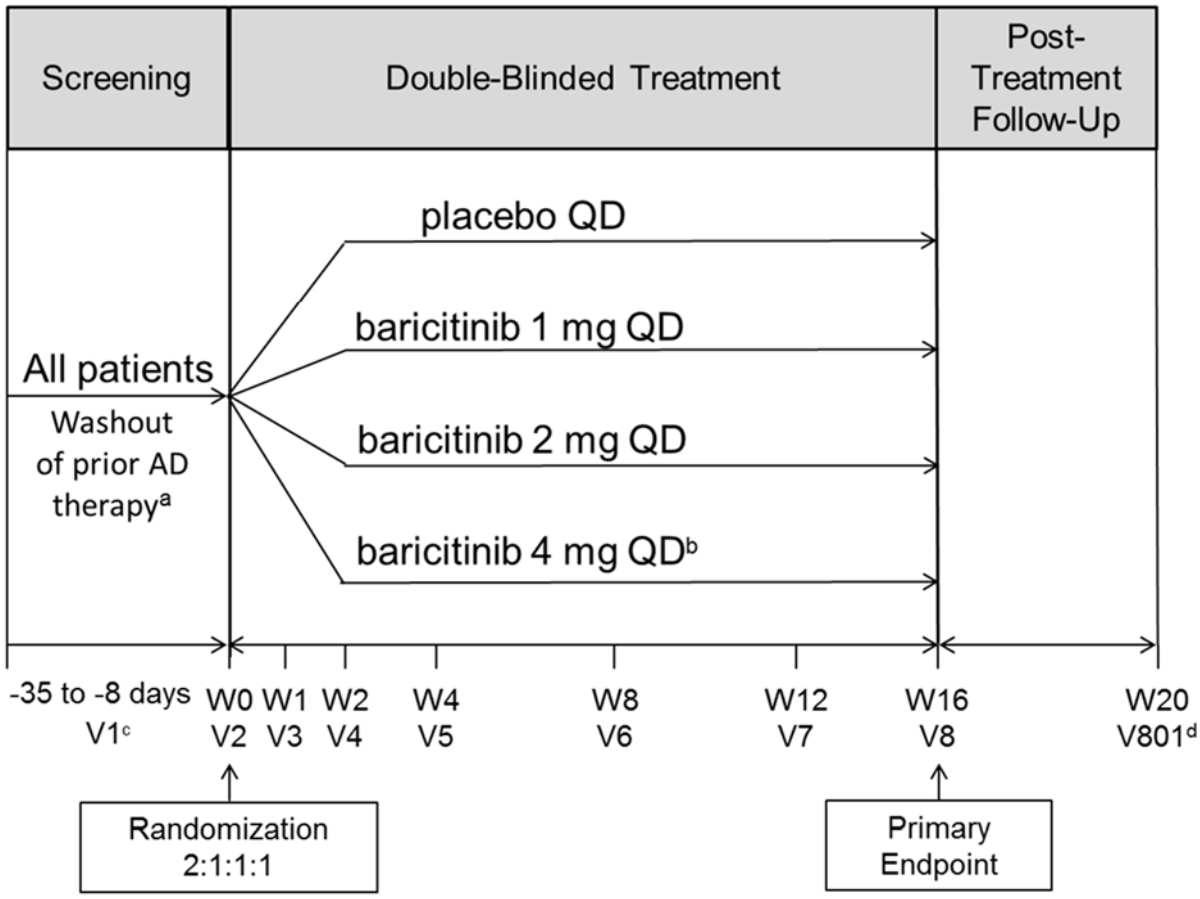
BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM)

BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) were concurrent multicentre, randomised, double-blind, placebo-controlled, parallel-group Phase III studies to determine the efficacy and safety of baricitinib in adult patients with moderate-to-severe AD. As part of the eligibility criteria, patients were required to have moderate-to-severe AD with a documented history of an inadequate response to, or intolerance to, topical medication. These eligibility criteria are broader than the population of relevance for this submission: adult patients with moderate-to-severe AD who are candidates for systemic therapy who experienced failure with, are intolerant to or have

contraindication to at least one systemic therapy. However, for completeness, these trials have been summarised in full within this submission.

The use of concomitant medications for the management of AD, including topical corticosteroids (TCS), was prohibited throughout the trial except for daily emollient use. Following a washout period of 4 weeks for systemic AD treatments and 2 weeks for topical AD treatments, excluding emollients, all patients were randomised 2:1:1:1 in a double-blinded fashion to once daily treatment with placebo, 1 mg baricitinib, 2 mg baricitinib or 4 mg baricitinib. The study design of BREEZE-AD1 and -AD2 is shown in Figure 7.[56, 57]

Figure 7: Study design of the BREEZE-AD1 and -AD2 trials



[a] Applicable to patients taking topical treatments (excluding emollients) or systemic treatments for AD at the time of screening. [b] For patients randomised to the 4 mg QD dose who had renal impairment (defined as eGFR <60 mL/min/1.73 m^2), the baricitinib dose was 2 mg QD. [c] Patients for whom PPD skin test for the evaluation of tuberculosis infection was performed at V1 had to return and PPD test was read 48–72 hours after V1 (post-PPD). [d] Occurred approximately 28 days after the last dose of the study treatment (was not required for those patients entering the long-term extension Study JAHN).
Abbreviations: AD: atopic dermatitis; eGFR: estimated glomerular filtration rate; PPD: purified protein derivative; QD: once daily; V: visit; W: week.
Sources: BREEZE-AD1 (JAHL) Clinical Study Report[56] and BREEZE-AD2 (JAHM) Clinical Study Report.[57]

The primary outcome of both trials was to evaluate the efficacy of baricitinib in adult AD patients as measured by the proportion of patients in each treatment group achieving IGA ≤1 with a ≥2-point improvement at the end of the treatment period (Week 16). Key secondary endpoints measured improvement in signs and symptoms of AD at Week 16 and included Eczema Area and Severity Index (EASI) and SCORing Atopic Dermatitis (SCORAD) scores and patient-

reported outcomes (Atopic Dermatitis Sleep Scale [ADSS] Item 2 score, Itch Numeric Rating Scale [NRS] and Skin Pain NRS). Other secondary endpoints included health-related quality of life (HRQoL) outcomes such as the Dermatology Life Quality Index (DLQI) and the 5-level EuroQol 5 Dimensions (EQ-5D-5L) questionnaires. Safety outcomes included adverse events (AEs), serious AEs (SAEs) and treatment-emergent AEs (TEAEs) by Week 16.[56, 57]

A summary of the methodology of the BREEZE-AD1 and -AD2 trials is presented in Table 6.

Table 6: Summary of BREEZE-AD7, -AD1 and -AD2 trial methodology

Trial name	BREEZE-AD7 (JAIY)[62, 63]	BREEZE-AD1 (JAHL)[64, 65]	BREEZE-AD2 (JAHM)[66, 67]
Location	International: patients recruited from 68 sites across 10 countries (5 sites in Argentina, 7 sites in Australia, 1 site in Austria, 10 sites in Germany, 5 sites in Italy, 17 sites in Japan, 8 sites in the Republic of Korea, 4 sites in Poland, 5 sites in Spain, 6 sites in Taiwan). No patients were enrolled in the UK.	International: patients recruited from 93 sites across 10 countries (9 sites in Czechia, 1 site in Denmark, 8 sites in France, 21 sites in Germany, 12 sites in India, 5 sites in Italy, 16 sites in Japan, 8 sites in Mexico, 6 sites in the Russian Federation, 7 sites in Taiwan). No patients were enrolled in the UK.	International: patients recruited from 80 sites across 10 countries (8 sites in Argentina, 6 sites in Australia, 5 sites in Austria, 9 sites in Hungary, 6 sites in Israel, 16 sites in Japan, 7 sites in the Republic of Korea, 10 sites in Poland, 9 sites in Spain, 4 sites in Switzerland). No patients were enrolled in the UK.
Trial design	Phase III multicentre, randomised, double-blind, placebo-controlled, parallel-group study		
Eligibility criteria for participants	A summary of the inclusion and exclusion criteria is provided below. Full details of the eligibility criteria are presented in Appendix D. Key inclusion criteria: • Aged 18 years or older • Diagnosis of AD at least 12 months prior to screening as defined by the American Academy of Dermatology • Moderate-to-severe AD, defined as having all of the following at the screening visit *and* at randomisation: o EASI score greater than or equal to 16 o IGA score greater than or equal to 3, and o BSA involvement greater than or equal to 10% • Documented history of inadequate response to topical medication *or* a history of intolerance to topical medication defined as having at least one of the following: o An inability to achieve good disease control after use of at least a medium potency TCS for at least 4 weeks, or for the maximum duration recommended by the product prescribing information, whichever is shorter o Patients who failed systemic therapies intended to treat AD, such as ciclosporin, methotrexate, azathioprine, and mycophenolate mofetil, within 6 months preceding screening o BREEZE-AD1 and -AD2 only: Documented history of clinically significant adverse reactions with the use of TCS such as skin atrophy, allergic reactions, or systemic effects that in the opinion of the investigator outweighed the benefits of retreatment Key exclusion criteria: • Currently experiencing, or have a history of, other concomitant skin conditions, including psoriasis or lupus erythematosus, which would interfere with evaluation of the effect of the study medication on AD, or which requires frequent hospitalisation and/or intravenous treatment for skin infections		

	• Eczema herpeticum within 12 months prior to screening or more than twice in the past • Any serious concomitant illness anticipated to require the use of systemic corticosteroids or require active frequent monitoring • BREEZE-AD7 (JAIY) only: Have an important side effect to TCS (e.g. intolerance to treatment or hypersensitivity reactions) which would prevent further use	
Method of study drug administration	Administered orally once daily as 2 tablets: a placebo tablet with a treatment tablet for treatment groups, or 2 placebo tablets for the placebo group	Administered orally once daily as 3 tablets: 2 placebo tablets with 1 treatment tablet for treatment groups, or 3 placebo tablets for the placebo group
Permitted and disallowed concomitant medication	A summary of the key concomitant medications permitted and disallowed during the study period is provided below. Full details are provided in Appendix D. All concomitant therapies for AD were prohibited except for: • Daily use of emollients, excluding additives like antipruritics or antiseptics • Prescription sleep medications used at entry • Non-sedating antihistamines • A single intra-articular or soft tissue corticosteroid injection • Intranasal or inhaled steroids • Topical anaesthetics and topical and systemic anti-infective medications • Non-live seasonal vaccines and/or emergency vaccinations Permitted in BREEZE-AD7 (JAIY) only: • Background TCS therapy with moderate-potency and/or low-potency TCS (e.g. triamcinolone 0.1% cream and/or hydrocortisone 2.5% ointment) used on active lesions. High- or ultra-high potency TCS permitted only as rescue therapy. • TCIs (e.g. tacrolimus and pimecrolimus), or topical PDE-4 inhibitor (i.e., crisaborole, where approved) are permitted in place of TCS on areas where application of TCS is considered inappropriate by the investigator; use should be limited to problem areas (e.g., face, neck, skin folds, genital areas, etc.) • Ophthalmic drugs containing antihistamines, corticosteroids or other immunosuppressants	
Primary outcome	Proportion of patients in the ITT population achieving IGA of 0 or 1 with a ≥2-point improvement at Week 16 of treatment	
Secondary and exploratory outcomes	A summary of the key secondary outcomes is provided below. Full details of all the secondary and exploratory outcomes can be found in Appendix D. • Improvement in signs and symptoms at Week 16: o EASI75 o EASI90	

	○ Percent change in EASI score ○ SCORAD75 • Patient-reported outcome measures at Week 16: ○ 4-point improvement in Itch NRS at Week 1, 2, 4 and 16 of treatment ○ Mean change in Item 2 of ADSS score at Week 1 or 16 of treatment ○ Mean change from baseline in Skin Pain NRS at Week 16 of treatment • HRQoL outcomes at Week 16: ○ DLQI ○ EQ-5D-5L	
Pre-planned subgroups	• Gender • Age group (<65, ≥65, ≥65 to <75, ≥75 to <85, ≥85 years old) • Baseline weight (<60, ≥60 to <100, ≥100 kg) • Baseline BMI (<25, ≥25 to <30, ≥30 kg/m^2) • Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple) • Baseline renal function status: impaired (eGFR <60 mL/min/1.73m^2) or not impaired (eGFR ≥60 mL/min/1.73m^2) • Region (Europe, Japan, rest of world) • Specific regions (Europe, other) • Specific country (Japan, other) • Prior systemic therapy use (Yes/No) • Baseline disease severity (IGA 3 or 4)	
Duration of study and follow-up	The total study duration was 25 weeks, with a 5-week screening period, a 16-week treatment period and 4-week post-treatment follow-up period.	The total study duration was 24 weeks, with a 27-day screening period, a 16-week treatment period and 4-week post-treatment follow-up period

Abbreviations: AD: atopic dermatitis; ADSS: atopic dermatitis sleep scale; BMI: body mass index; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HRQoL: health-related quality of life; IGA: Investigator's Global Assessment; ITT: intent-to-treat; NRS: numeric rating score; PDE: phosphodiesterase; SCORAD: SCORing Atopic Dermatitis; TCI: topical calcineurin inhibitors; TCS: topical corticosteroids.
Sources: BREEZE-AD7 (JAIY) Clinical Protocol,[62] BREEZE-AD7 (JAIY) Statistical Analysis Plan, [63] BREEZE-AD1 (JAHL) Clinical Protocol,[64] BREEZE-AD1 (JAHL) Statistical Analysis Plan,[65] BREEZE-AD2 (JAHM) Clinical Protocol,[66] BREEZE-AD2 (JAHM) Statistical Analysis Plan.[67]

B.2.3.2 Outcome definitions

All outcome definitions were consistent across the BREEZE-AD trials and are summarised in Table 7.

Table 7: Definitions of outcomes used in the BREEZE-AD trials

Outcome	Definition
EASI	The EASI assesses disease extent at four body regions and measures four clinical signs (erythema, induration/papulation excoriation, and lichenification) each on a scale of 1–3. It confers a maximum score of 72.
IGA	The IGA measured the investigator's global assessment of the patient's overall AD severity on a 5-point scale based on an overall assessment of the degree of erythema, papulation/induration, oozing/crusting, and lichenification: 0 (clear), 1 (almost clear), 2 (mild), 3 (moderate) or 4 (severe).
SCORAD	The SCORAD index assesses six clinical characteristics to determine disease severity: erythema, oedema/papulation, oozing/crusts, excoriation, lichenification, and dryness. It also assesses subjective symptoms (pruritis and sleep loss). It confers a maximum score of 103.
Itch NRS	The Itch NRS is a patient-administered 11-point scale in which 0 represents 'no itch' and 10 represents 'worst itch imaginable'. Overall severity of patient itch is indicated by selection of the number that most closely describes the worst level of itching experienced in the last 24 hours.
Skin Pain NRS	The Skin Pain NRS is a patient-administered 11-point scale in which 0 represents 'no pain' and 10 represents 'worst pain imaginable'. Overall severity of patient skin pain is indicated by selection of the number that most closely describes the worst level of pain experienced in the last 24 hours.
ADSS	The ADSS is a three-item patient-reported outcome that assesses the impact of itch on sleep including difficulty falling asleep (Item 1), frequency of waking due to itch (Item 2), and difficulty getting back to sleep last night (Item 3). Each day, patients rate Item 2 by selecting the number of times they woke the previous night due to itch (0–29).
DLQI	The DLQI is a patient-administered 10-item validated quality-of-life questionnaire that covers six domains: symptoms and feelings, daily activities, leisure, work and school, personal relationships and treatment. Each category is scored to indicate impairment in that area over the last week: 1 for "not at all," 2 for "a lot," and 3 for "very much," and unanswered ("not relevant") responses scored as 0. Scores range from 0–30 with higher scores indicating greater impairment.
EQ-5D-5L	The EQ-5D-5L is a standardised measure of health status that consists of two components: a descriptive system of the respondent's health (Health Index Score) and a rating of his or her current health state using a 0 to 100 mm VAS. The descriptive system examines mobility, self-care, usual activities, pain/discomfort and anxiety/depression, each of which is assessed on 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The VAS records the respondent's self-rated health on a vertical VAS where the endpoints are labelled as "best imaginable health state" and "worst imaginable health state."

Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; DLQI: Dermatitis Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; IGA: Investigator's Global Assessment; NRS: Numeric Rating System; SCORAD: SCORing Atopic Dermatitis; VAS: visual analogue score.

B.2.3.3 Baseline characteristics of study participants

The baseline characteristics of enrolled patients were well-balanced between treatment arms in the BREEZE-AD4 (Table 8), -AD7 (Table 9), AD1 (Table 10) and -AD2 (Table 11) trials; this

consistency was broadly maintained across trials. Patients included in the BREEZE-AD4 (JAIN) trial had experienced inadequate response to, were intolerant to or had a contraindication to ciclosporin representing a narrower population than the BREEZE-AD7, -AD1 and -AD2 trials and in line with the population of interest for this submission. Despite this, [redacted]% in the BREEZE-AD7, -AD1 and -AD2 trials had received a prior systemic therapy, of which [redacted]% had received prior ciclosporin, aligning the with population of interest.

Patients enrolled in BREEZE-AD4, -AD7, -AD1 and -AD2 trials had a mean age of [redacted] years old with women representing approximately [redacted]% of the study populations. BREEZE-AD4 (JAIN) included patients from the UK (N=[redacted]). Despite the lack of UK patients in the other trials, all four BREEZE-AD trials included a high percentage of European patients: approximately [redacted]%, [redacted]%, [redacted]% and [redacted]% in BREEZE-AD4, -AD7, -AD1 and -AD2, respectively, and the majority of patients in BREEZE-AD4, -AD1 and -AD2 were Caucasian ([redacted]%). Patients had an average duration since diagnosis of approximately [redacted] years, a mean weight of [redacted] kg and a mean BMI of [redacted].

Table 8: Baseline characteristics of patients in the BREEZE-AD4 (JAIN) trial

Characteristic	BREEZE-AD4 (JAIN)			
	PBO (N=93)	1 mg (N=93)	2 mg (N=185)	4 mg (N=92)
Age (years), mean (SD)	39 (14)	39 (14)	37 (14)	39 (13)
Female, %	47	38	28	38
Race				
Caucasian, %	80	75	78	77
Asian, %	[redacted]	[redacted]	[redacted]	[redacted]
Other, %	[redacted]	[redacted]	[redacted]	[redacted]
Duration since AD diagnosis (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Weight (kg), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Body mass index (kg/m²), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Geographic region				
Europe, %	[redacted]	[redacted]	[redacted]	[redacted]
Japan, %	[redacted]	[redacted]	[redacted]	[redacted]
Rest of world, %	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 at screening Visit 1, %	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 Visit 2, %	54	51	51	51
EASI, mean (SD)	31 (11.6)	34 (13.5)	31 (12.4)	33 (13.7)
SCORAD, mean (SD)	69 (13.0)	71 (14.1)	68 (13.4)	69 (13.4)
BSA affected, mean (SD)	48 (21.3)	57 (23.8)	50 (22.2)	54 (23.8)
POEM, mean (SD)	21 (5.7)	21 (6.0)	21 (5.9)	21 (6.0)
ADSS Item 2, mean (SD)	1.6 (1.6)	2.2 (2.7)	1.9 (3.1)	2.1 (1.8)
DLQI, mean (SD)	14.5 (6.9)	14.3 (8.3)	13.6 (7.4)	14.0 (8.1)
Itch NRS, mean (SD)	7.1 (1.9)	6.7 (2.3)	6.7 (1.9)	6.7 (2.3)

Skin Pain NRS, mean (SD)	6.5 (2.3)	6.3 (2.7)	6.1 (2.4)	6.1 (2.6)
PGI-S-AD, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS depression, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety and depression combined, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L VAS score, mean	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L HIS, mean	[redacted]	[redacted]	[redacted]	[redacted]
Prior TCS therapy, n (%)[a]	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical calcineurin inhibitor use, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Prior systemic therapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Systemic corticosteroid use	[redacted]	[redacted]	[redacted]	[redacted]
Systemic immunosuppressant use	[redacted]	[redacted]	[redacted]	[redacted]
Ciclosporin use	[redacted]	[redacted]	[redacted]	[redacted]
Biologic use, n (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]
Phototherapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]

[a] Only TCS use in the 12 months preceding screening was recorded. [b] Biologics use included 1 patient on adalimumab, 33 patients on dupilumab, 1 patient on etanercept, 1 patient on lebrikizumab, 2 patients on nemolizumab, 6 patients on omalizumab, 4 patients on tralokinumab, and 1 patient on ustekinumab.
Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; HIS: health index score; IGA: Investigator's Global Assessment; NRS: Numeric Rating Scale; PBO: placebo; PGI-S-AD: Patient Global Impression of Severity–Atopic Dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.; SD: standard deviation; TCS: topical corticosteroids; VAS: visual analogue score.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Table 9: Baseline characteristics of patients in the BREEZE-AD7 (JAIY) trial

Characteristic	**BREEZE-AD7 (JAIY)**		
	PBO + TCS (N=[redacted])	**2 mg +TCS (N=[redacted])**	**4 mg +TCS (N=[redacted])**
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]
Female, %	[redacted]	[redacted]	[redacted]
Race			
Caucasian, %	[redacted]	[redacted]	[redacted]
Asian, %	[redacted]	[redacted]	[redacted]
Other, %	[redacted]	[redacted]	[redacted]

Duration since AD diagnosis (years), mean (SD)	[redacted]	[redacted]	[redacted]
Weight (kg), mean (SD)	[redacted]	[redacted]	[redacted]
Body mass index (kg/m^2), mean (SD)	[redacted]	[redacted]	[redacted]
Geographic region			
Europe, %	[redacted]	[redacted]	[redacted]
Japan, %	[redacted]	[redacted]	[redacted]
Rest of world, %	[redacted]	[redacted]	[redacted]
IGA of 4 at screening, Visit 1, %	[redacted]	[redacted]	[redacted]
IGA of 4 at baseline, Visit 2, %	[redacted]	[redacted]	[redacted]
EASI, mean (SD)	[redacted]	[redacted]	[redacted]
SCORAD, mean (SD)	[redacted]	[redacted]	[redacted]
BSA affected, mean (SD)	[redacted]	[redacted]	[redacted]
POEM, mean (SD)	[redacted]	[redacted]	[redacted]
ADSS item 2, mean (SD)	[redacted]	[redacted]	[redacted]
DLQI, mean (SD)	[redacted]	[redacted]	[redacted]
Itch NRS, mean (SD)	[redacted]	[redacted]	[redacted]
Skin Pain NRS, mean (SD)	[redacted]	[redacted]	[redacted]
PGI-S-AD, mean (SD)	[redacted]	[redacted]	[redacted]
HADS anxiety, mean (SD)	[redacted]	[redacted]	[redacted]
HADS depression, mean (SD)	[redacted]	[redacted]	[redacted]
HADS anxiety and depression combined, mean (SD)	[redacted]	[redacted]	[redacted]
EQ-5D-5L VAS score, mean	[redacted]	[redacted]	[redacted]
EQ-5D-5L HIS, mean	[redacted]	[redacted]	[redacted]
Prior topical corticosteroid therapy, n (%)[a]	[redacted]	[redacted]	[redacted]
Prior topical calcineurin inhibitor use, n (%)	[redacted]	[redacted]	[redacted]
Prior systemic therapy, n (%)	[redacted]	[redacted]	[redacted]
Systemic corticosteroid use	[redacted]	[redacted]	[redacted]
Systemic immunosuppressant use	[redacted]	[redacted]	[redacted]
Ciclosporin use	[redacted]	[redacted]	[redacted]
Biologic use, n (%)[b]	[redacted]	[redacted]	[redacted]

[a] Only TCS use in the 12 months preceding screening was recorded. [b] Biologics use included 10 patients on dupilumab, 1 patient on lebrikizumab, 4 patients on nemolizumab, 1 patient on omalizumab, and 7 patients on tralokinumab.
Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; HIS: health index score; IGA: Investigator's Global Assessment; NRS: Numeric Rating Scale; PBO: placebo; PGI-S-AD: Patient Global Impression of Severity–Atopic Dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.; SD: standard deviation; TCS: topical corticosteroids; VAS: visual analogue score.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Table 10: Baseline characteristics of patients in the BREEZE-AD1 (JAHL) trial

Characteristic	BREEZE-AD1 (JAHL)			
	PBO (N=249)	1mg (N=127)	2 mg (N=123)	4 mg (N=125)
Age (years), mean (SD)	35 (12.6)	36 (12.4)	35 (13.7)	37 (12.9)
Female, n (%)	101 (40.6)	49 (38.6)	41 (33.3)	42 (33.6)
Race				
Caucasian, n (%)	147 (59.5)[a]	74 (58.3)	75 (61.0)	70 (56.5)
Asian, n (%)	73 (29.6)[a]	40 (31.5)	35 (28.5)	41 (33.1)
Other, n (%)	27 (10.9)[a]	13 (10.2)	13 (10.6)	14 (11.2)
Duration since AD diagnosis (years), mean (SD)	26 (15.5)	27 (14.9)	25 (14.6)	25 (14.9)
Weight (kg), mean (SD)	73 (15.7)	74 (17.2)	75 (17.7)	74 (17.2)
Body mass index (kg/m²), mean (SD)	25 (4.5)	25 (4.6)	25 (5.1)	25 (4.3)
Geographic region				
Europe, n (%)	135 (54.2)	67 (52.8)	67 (54.5)	68 (54.5)
Japan, n (%)	45 (18.1)	23 (18.1)	21 (17.1)	22 (17.6)
Other,[b] n (%)	69 (27.7)	37 (29.1)	35 (28.5)	35 (28.0)
IGA of 4 at screening, Visit 1, %	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 at baseline, Visit 2, n (%)	105 (42.2)	53 (41.7)	52 (42.3)	51 (40.8)
EASI, mean (SD)	32 (13.0)	29 (11.8)	31 (11.7)	32 (12.7)
SCORAD, mean (SD)	68 (14.0)	66 (14.3)	68 (13.0)	68 (13.0)
BSA affected, mean (SD)	53 (23.1)	47 (21.2)	50 (22.1)	52 (21.8)
POEM, mean (SD)	21 (5.6)	20 (5.6)	21 (5.6)	21 (5.6)
ADSS Item 2, mean (SD)	3.4 (5.2)	2.5 (3.4)	2.3 (4.1)	3.3 (5.2)
DLQI, mean (SD)	14 (7.4)	13 (6.9)	13 (7.7)	14 (7.1)
Itch NRS, mean (SD)	7 (2.0)	6 (2.1)	6 (2.2)	6 (2.0)
Skin Pain NRS, mean (SD)	6 (2.5)	5 (2.4)	6 (2.6)	6 (2.4)
PGI-S-AD, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]

HADS anxiety, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS depression, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety and depression combined, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L VAS score, mean	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L HIS, mean	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical corticosteroid therapy, n (%)[c,d]	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical calcineurin inhibitor use, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Prior systemic therapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Systemic corticosteroid use	[redacted]	[redacted]	[redacted]	[redacted]
Systemic immunosuppressant use	[redacted]	[redacted]	[redacted]	[redacted]
Ciclosporin use	[redacted]	[redacted]	[redacted]	[redacted]
Biologic use[e]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Two patients were excluded due to missing data and percentages have been calculated from the new denominator (N=247). [b] Other represents India, Mexico and Taiwan. [c] Patients with documented systemic treatment for AD in the past 6 months were also considered inadequate responders to topical treatments and were eligible to enrol. [d] Only topical corticosteroid use in the 12 months preceding screening was recorded. [e] Biologics use included 33 patients on dupilumab, 1 patient on etanercept, 3 patients on interleukin inhibitors, 9 patients on lebrikizumab, 1 patient on nemolizumab, 2 patients on omalizumab, 1 patient on reslizumab, 3 patients on tralokinumab, and 1 patient on ustekinumab.
Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; HIS: health index score; IGA: Investigator's Global Assessment; NRS: Numeric Rating Scale; PBO: placebo; PGI-S-AD: Patient Global Impression of Severity–Atopic Dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.; SD: standard deviation; TCS: topical corticosteroids; VAS: visual analogue score.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Table 11: Baseline characteristics of patients in the BREEZE-AD2 (JAHM) trial

Characteristic	**BREEZE-AD2 (JAHM)**			
	PBO (N=[redacted])	**1 mg (N=[redacted])**	**2 mg (N=[redacted])**	**4 mg (N=[redacted])**
Age (years), mean (SD)	35 (13.0)	33 (10.0)	36 (13.2)	34 (14.1)
Female, n (%)	90 (36.9)	45 (36.0)	58 (47.2)	41 (33.3)
Race				
Caucasian, n (%)	169 (69.3)	85 (68.0)	85 (69.1)	82 (66.7)
Asian, n (%)	72 (29.5)	36 (28.8)	37 (30.1)	38 (30.9)
Other, n (%)	3 (1.2)	4 (3.2)	1 (0.8)	2 (2.4)

Duration since AD diagnosis (years), mean (SD)	25 (14)	24 (13)	24 (14)	23 (15)
Weight (kg), mean (SD)	72 (16)	75 (17)	72 (15)	73 (15)
Body mass index (kg/m²), mean (SD)	25 (4.3)	26 (5.2)	25 (5.0)	25 (4.2)
Geographic region				
Europe, n (%)	111 (45.5)	57 (45.6)	56 (45.5)	56 (45.5)
Japan, n (%)	45 (18.4)	22 (17.6)	22 (17.9)	23 (18.7)
Other,[a] n (%)	88 (36.1)	46 (36.8)	45 (36.6)	44 (35.8)
IGA of 4 at screening Visit 1, %	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 Visit 2, %	121 (49.6)	63 (50.8)	62 (50.4)	63 (51.2)
EASI, mean (SD)	33 (12.8)	33 (12.7)	35 (16.0)	33 (12.7)
SCORAD, mean (SD)	68 (12.7)	67 (12.9)	69 (13.3)	68 (13.6)
BSA affected, mean (SD)	52 (21.7)	55 (21.9)	55 (26.1)	54 (21.5)
POEM, mean (SD)	21 (6.3)	20 (6.5)	21 (6.0)	20 (6.3)
ADSS Item 2, mean (SD)	1.8 (2.1)	1.6 (1.8)	2.1 (2.9)	1.9 (2.5)
DLQI, mean (SD)	15 (8.1)	15 (8.1)	14 (7.7)	14 (8.4)
Itch NRS, mean (SD)	6.8 (2.2)	6.4 (2.2)	6.6 (2.2)	6.6 (2.2)
Skin Pain NRS, mean (SD)	6.2 (2.5)	5.7 (2.7)	6.2 (2.5)	6.0 (2.6)
PGI-S-AD, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS depression, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety and depression combined score (SD)	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L VAS score, mean	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L HIS, mean	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical corticosteroid therapy, n (%)[b,c]	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical calcineurin inhibitor use, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Prior systemic therapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Systemic corticosteroid use	[redacted]	[redacted]	[redacted]	[redacted]
Systemic immunosuppressant use	[redacted]	[redacted]	[redacted]	[redacted]

Ciclosporin use	[redacted]	[redacted]	[redacted]	[redacted]
Biologic use, n (%)[d]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Other represents Argentina, Australia, Israel and South Korea. [b] Patients with documented systemic treatment for AD in the past 6 months were also considered inadequate responders to topical treatments and were eligible to enroll. [c] Only topical corticosteroid use in the 12 months preceding screening was recorded. [d] Biologics use included 12 patients on dupilumab, 1 patient on lebrikizumab, 1 patient on nemolizumab, 7 patients on omalizumab, 1 patient on tralokinumab, and 2 patients on ustekinumab.
Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; HIS: health index score; IGA: Investigator's Global Assessment; NRS: Numeric Rating Scale; PBO: placebo; PGI-S-AD: Patient Global Impression of Severity–Atopic Dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.; SD: standard deviation; TCS: topical corticosteroids; VAS: visual analogue score.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

B.2.4 *Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence*

The analysis sets used in the analysis of the BREEZE-AD1, -AD2, -AD4 and -AD7 trials are presented in Table 12.

Across all BREEZE-AD RCTs the main analysis method of categorical variables was logistic regression analysis with Fisher's exact test used as a secondary analysis method and the primary analysis for continuous variables MMRM analysis, with an ANCOVA model used as a secondary analysis method. Full details of the statistical methods for the primary analysis of the BREEZE-AD trials are presented in Appendix L.

Table 12: Trial populations used for the analysis of outcomes in the BREEZE-AD1, -AD2, -AD4 and -AD7 trials

Analysis set		BREEZE-AD4 (JAIN)[54, 61]	BREEZE-AD7 (JAIY)[55, 63]	BREEZE-AD1 (JAHL)[56, 65]	BREEZE-AD2 (JAHM)[57, 67]
Intent-to-treat (ITT) population	Description	• Comprises all patients who were randomised at baseline • Analyses of the primary and secondary efficacy and health outcomes were performed on the ITT			
	N	463	[redacted]	624	615
Per-protocol Set (PPS)	Description	• Comprises patients in the ITT deemed compliant with treatment in Period 2 (up to Week 52) as determined before unblinding and database lock, who do not have significant protocol variations and whose investigator site does not have significant GCP issues that require a report to the regulatory agencies prior to Week 16 (Visit 8) • Sensitivity analyses of the primary and secondary efficacy and health outcomes were performed in the PPS	• Comprises patients in the ITT without major protocol deviations as determined before unblinding and database lock • Sensitivity analyses of the primary and secondary efficacy and health outcomes were performed in the PPS		
	N	[redacted]	[redacted]	[redacted]	[redacted]
Follow-up population	Description	N/A	• Comprises patients who entered the follow-up period		
	N		[redacted]	[redacted]	[redacted]
Safety population (SP)	Description	• Comprises all randomised patients who received study medication (baricitinib or placebo) who did not discontinue for the reason 'Lost to Follow-up' at the first postbaseline visit excluding patients with no safety assessments postbaseline All safety analyses were performed on the safety population	• Comprises all randomised patients who received ≥1 dose of study medication (baricitinib or placebo) who did not discontinue for the reason 'Lost to Follow-up' at the first postbaseline visit • All safety analyses were performed on the safety population		
	N	[redacted]	[redacted]	[redacted]	[redacted]

[a] One patient in the BREEZE-AD7 (JAIY) trial failed screening and was randomised to placebo in error but did not receive study treatment. [b] One patient in the BREEZE-AD2 (JAHM) trial was randomised but did not receive study treatment due to an inability to collect laboratory samples.
Abbreviations: GCP: Good Clinical Practice; ITT: intent-to-treat; N/A: not applicable; PPS: per-protocol set; SP: safety population.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] BREEZE-AD4 (JAIN) Statistical Analysis Plan,[61] BREEZE-AD7 (JAIY) Clinical Study Report,[55] BREEZE-AD7 (JAIY) Statistical Analysis Plan,[63] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD1 (JAHL) Statistical Analysis Plan,[65] BREEZE-AD2 (JAHM) Clinical Study Report,[57] BREEZE-AD2 (JAHM) Statistical Analysis Plan,[67] Simpson *et al,* 2020,[49] Bieber *et al*, 2020[58]

B.2.4.1 Censoring

In the BREEZE-AD trial programmes data have been analysed according to two prespecified censoring rules for all efficacy endpoints:

- **Primary censoring rule:** data were censored as missing or non-responder imputation after permanent study drug discontinuation or following initiation of rescue therapy with TCS (any potency in the monotherapy trials or high or ultra-high potency in the combination trials) or systemic therapy. This censoring rule was applied to all continuous and categorical efficacy and health outcome endpoints. Alternatively, this censoring rule is equivalent to the using all the data up to rescue.
- **Secondary censoring rule**: data were censored as missing or non-responder imputation after permanent study drug discontinuation or after initiation of systemic rescue therapies, but were not considered as missing after rescue with TCS. As a consequence, data for patients rescued with high or ultra-high potency TCS or with phototherapy were not censored at the time of rescue as they could continue or only temporarily interrupt the study drug. Patients who were rescued to systemic corticosteroids were required to permanently discontinue the study drug, thus had post-rescue observations censored. The secondary censoring rule was applied to primary and key secondary efficacy and health outcome endpoints.

Results for all efficacy endpoints have been presented in Section B.2.6 using the primary censoring rule, as this informs the base case economic analysis (see Section B.3). Additionally, given that it is reasonable to expect the concomitant use of rescue medication with baricitinib when required, results using the secondary censoring rule are presented alongside the primary censoring data for IGA, EASI and DLQI outcomes for BREEZE-AD4 and -AD7 since these data inform a scenario analysis explored within the economic model in B.3.8.3.

Non-responder imputation (for categorical variables) and MMRM (for continuous variables) were the primary methods used to handle missing data. The detailed statistical analyses used to calculate the primary endpoint and key secondary endpoints in all BREEZE-AD trials, alongside sample size calculations and methods for handling missing data, are presented in detail in Appendix L.

B.2.4.2 Participant disposition

CONSORT diagrams of patient disposition for all studies are presented in Appendix D.

BREEZE-AD4 (JAIN)

A total of [redacted] patients were screened for eligibility into the BREEZE-AD4 (JAIN) trial, of whom [redacted] were randomised in a 1:1:2:1 ratio, although [redacted] assigned to the 2 mg baricitinib group did not receive treatment: [redacted] received placebo, [redacted] received 1 mg baricitinib, [redacted] received 2 mg baricitinib and [redacted] received 4 mg baricitinib. Overall at Week 16, [redacted] patients ([redacted]) had discontinued: [redacted] from the placebo group, [redacted] from the 1 mg baricitinib group, [redacted] from the 2 mg baricitinib group, and [redacted] from the 4 mg baricitinib group. By Week 24, [redacted] patients ([redacted]%) had discontinued: [redacted] from the placebo group, [redacted] from the 1 mg baricitinib group, [redacted] from the 2 mg baricitinib group, and [redacted] from the 4 mg baricitinib group.[54] A summary of reasons for discontinuation from the BREEZE-AD4 (JAIN) trial at Week 16 and Week 24 is presented in Table 13.

BREEZE-AD7 (JAIY)

A total of [redacted] patients were screened for eligibility into the BREEZE-AD7 (JAIY) trial, of whom [redacted] were randomised in a 1:1:1 ratio: [redacted] received placebo, [redacted] received 2 mg baricitinib and [redacted] received 4 mg baricitinib. Overall, [redacted] patients ([redacted]%) discontinued: [redacted] from the placebo group, [redacted] from the 2 mg baricitinib group and [redacted] from the 4 mg baricitinib group.[55] A summary of reasons for discontinuation from the BREEZE-AD7 (JAIY) trial is presented in Table 14.

BREEZE-AD1 (JAHL)

A total of 757 patients were screened for eligibility into the BREEZE-AD1 (JAHL) trial, of whom 624 were randomised in a 2:1:1:1 ratio: 249 received placebo, 127 received 1 mg baricitinib, 123 received 2 mg baricitinib and 125 received 4 mg baricitinib. Overall, 49 patients (6.5%) discontinued: 23 from the placebo group, 11 from the 1 mg baricitinib group, 10 from the 2 mg baricitinib group and 5 from the 4 mg baricitinib group.[56] A summary of reasons for discontinuation from the BREEZE-AD1 (JAHL) trial is presented in Table 14.

BREEZE-AD2 (JAHM)

A total of 728 patients were screened for eligibility into the BREEZE-AD2 (JAHM) trial, of whom 624 were randomised in a 2:1:1:1 ratio: 244 received placebo, 125 received 1 mg baricitinib, 123 received 2 mg baricitinib and 123 received 4 mg baricitinib. Overall, 45 patients (6.2%) discontinued: 19 from the placebo group, 10 from the 1 mg baricitinib group, 10 from the 2 mg baricitinib group and 6 from the 4 mg baricitinib group.[57] A summary of reasons for discontinuation from the BREEZE-AD2 (JAHM) trial is presented in Table 14.

Table 13: Reasons for study discontinuation at Week 16 and Week 24 in the BREEZE-AD4 (JAIN) trial

Characteristic, N (%)	Week 16				Week 24			
	PBO + TCS (N=[redacted])	1 mg + TCS (N=[redacted])	2 mg + TCS (N=[redacted])	4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	1 mg + TCS (N=[redacted])	2 mg + TCS (N=[redacted])	4 mg + TCS (N=[redacted])
Completed study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Reason for discontinuation, n (%)								
Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawal by patient	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The number of discontinuations at Week 24 include discontinuations at Week 16. [a] For the 1 mg baricitinib group at Week 16, other reasons were protocol deviation (N=2) and physician decision (N=1). No further other discontinuation reasons were recorded by Week 24. [b] For the 2 mg baricitinib group at Week 16, other reasons were protocol deviation (N=1). [c] For the 2 mg baricitinib group at Week 24, the additional other reason was protocol deviation (N=1).
Abbreviations: AE: adverse event; PBO: placebo; TCS: topical corticosteroids.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

Table 14: Reasons for study discontinuation in the BREEZE-AD7, -AD1 and -AD2 trials

Characteristic, N (%)	BREEZE-AD7[55]			BREEZE-AD1[56]				BREEZE-AD2[57]			
	PBO + TCS (N=[redacted])	2 mg + TCS (N=[redacted])	4 mg + TCS (N=[redacted])	PBO (N=[redacted])	1 mg (N=[redacted])	2 mg (N=[redacted])	4 mg (N=[redacted])	PBO (N=[redacted])	1 mg (N=[redacted])	2 mg (N=[redacted])	4 mg (N=[redacted])
Completed study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Reason for discontinuation, n (%)											
Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawal by patient	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] For the BREEZE-AD7 placebo + TCS group, other reasons included patient failed screening and was randomised in error (N=1) and patient was noncompliant with study visits (N=1). [b] For the BREEZE-AD1 placebo group, other reasons included treatment required for am medical history event (N=1) and new job (N=1). [c] For the BREEZE-AD1 1 mg baricitinib group, other reasons included pregnancy (N=1) and positive tuberculosis test at screening (N=1). [d] For BREEZE-AD2 1 mg baricitinib group, other reasons included inability to obtain laboratory samples (N=1) and pregnancy (N=1).
Abbreviations: AE: adverse event; PBO: placebo; TCS: topical corticosteroids.
Sources: BREEZE-AD7 (JAIY) Clinical Study Report,[55] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

B.2.5 *Quality assessment of the relevant clinical effectiveness evidence*

The BREEZE-AD RCTs were assessed for risk of risk of bias and generalisability in line with NICE requirements. Overall, the results of the BREEZE-AD trials may be considered at low risk of bias, as summarised in Table 15. Randomisation, concealment of treatment allocation and blinding of the participants and care providers were adequate. Baseline characteristics were well balanced between the treatment groups. All randomised patients were included in the ITT analysis for primary and secondary efficacy outcomes. There were no unexpected differences in the rates of treatment discontinuation between treatment arms.

Table 15: Quality assessment results for the BREEZE-AD RCTs

	BREEZE-AD1	BREEZE-AD2	BREEZE-AD4	BREEZE-AD7
Was randomisation carried out appropriately?	Y	Y	Y	Y
Was the concealment of treatment allocation adequate?	Y	Y	Y	Y
Were the groups similar at the outset of the study in terms of prognostic factors?	Y	Y	Y	Y
Were the care providers, participants and outcome assessors blind to treatment allocation?	Y	Y	Y	Y
Were there any unexpected imbalances in dropouts between groups?	N	N	N	N
Is there any evidence to suggest that the authors measured more outcomes than they reported?	N	N	N	N
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Y	Y	Y	Y

Adapted from Systematic Reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination).
Abbreviations: N: no; NA: not applicable; RCT: randomised controlled trial; UN: unclear; Y: yes.

B.2.6 *Clinical effectiveness results of the relevant trials*

Summary of clinical effectiveness results

- Across the BREEZE-AD trial programme, baricitinib treatment was associated with reduced disease burden
- In the monotherapy trials (BREEZE-AD1 and -AD2), a statistically higher proportion of patients in the baricitinib treatment arm achieved the primary outcome with significant improvements in secondary outcomes assessing signs and symptoms of AD, including IGA, EASI and SCORAD outcomes, observed at Week 16 as compared with placebo
- In the combination therapy trials (BREEZE-AD7 and -AD4), a statistically higher proportion of patients in the baricitinib treatment arm achieved the primary outcome with significant improvements in secondary outcomes assessing signs and symptoms of AD, including IGA and EASI outcomes, observed at Week 16 as compared with placebo
- In all pivotal BREEZE-AD RCTs, baricitinib treatment was statistically significantly associated with reduced itch, skin pain and sleep disturbance due to itch, and with significant improvements in HRQoL as assessed by DLQI and ED-5D-5L at Week 16 as compared with

placebo. In the BREEZE-AD4 (JAIN) trial, these significant differences were observed to Week 24.

The anticipated licence dose for baricitinib in AD is 4 mg once daily. For this reason, treatment arms involving administration of other baricitinib doses (1 mg or 2 mg) have been excluded from consideration in this clinical effectiveness summary. However, a dose of 2 mg once daily may be appropriate for some patients, such as those aged 75 years or older, and may be appropriate for patients with a history of chronic or recurrent infections. Data for other baricitinib doses have been presented in figures for completeness, and full results for other baricitinib doses are presented in the clinical study reports that have been provided in the reference pack for this submission.

As discussed in Section B.2.4, for all efficacy endpoints across the BREEZE-AD RCTs, two prespecified censoring rules were applied: the primary censoring rule censored following use of rescue medication or permanent study drug discontinuation; the secondary censoring rule censored following permanent study drug discontinuation only.

Results have been presented for BREEZE-AD4 (Section B.2.6.1), -AD7 (Section B.2.6.2) and -AD1 and -AD2 (Section B.2.6.3) where the conservative primary censoring rule was applied. This primary censoring rule was used to inform the base case economic analysis presented in Section B.3. High rates of rescue may have skewed results for categorical variables where NRI was used to account for censoring, if patients who had received rescue therapy were still benefitting from treatment with baricitinib. Rescue rates and types of rescue medication provided for each of the treatment arms in the originating BREEZE-AD RCTs are presented in Section B.2.10.6.

As such, a scenario analysis was explored where secondary censoring results from BREEZE-AD7 and -AD4 were used (B.3.8.3) so the secondary censoring results for IGA score, EASI outcomes and DLQI outcomes of BREEZE-AD4 (Section B.2.6.1) and -AD7 (Section B.2.6.2) are presented below. Results based on analyses using the secondary censoring rule are available in the relevant CSRs for BREEZE-AD1 and -AD2.

B.2.6.1 Combination therapy trial: BREEZE-AD4 (JAIN)

In this section, results for all efficacy endpoints reported use the conservative primary censoring rule. Additionally, results using the secondary censoring rule have been presented alongside the primary censoring data for IGA, EASI and DLQI outcomes.

Primary efficacy endpoint: EASI75 at Week 16

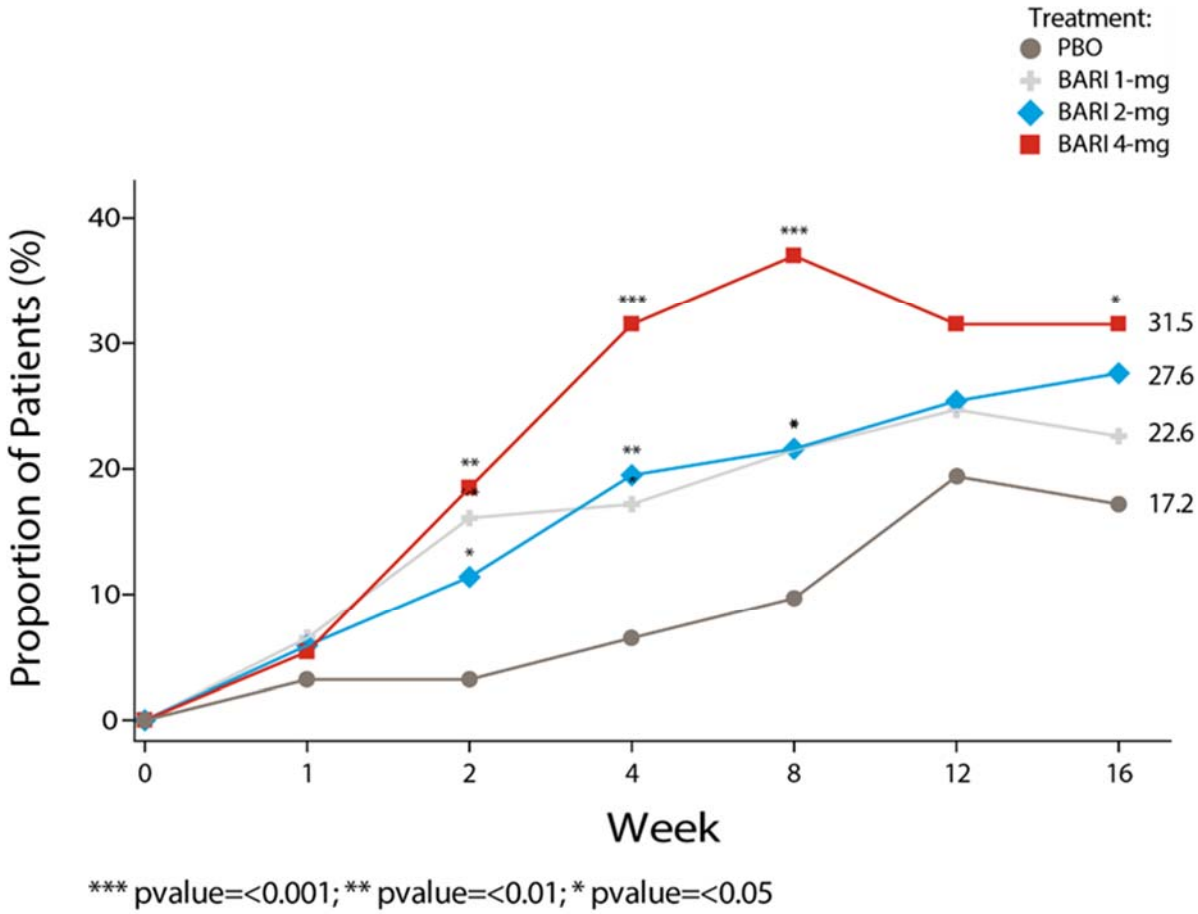
The EASI score measures disease extent at four body regions with higher scores representing higher disease burden, with EASI75 representing an improvement of 75% in EASI score from baseline. The proportion of patients achieving EASI75 at Week 16 is summarised in Table 16. In the 4 mg baricitinib group (primary censoring), a statistically significantly higher proportion of patients achieved EASI75 at Week 16, with 31.5% (95% CI: [redacted], [redacted]) of patients achieving the endpoint as compared with 17.2% (95% CI: [redacted], [redacted]) in the placebo group (odds ratio: [redacted] [95% CI: [redacted], p=[redacted]). This improvement in EASI75 versus placebo was statistically significant at p<0.05 as early as Week 2 and was maintained at p<0.01 from Week 2 through to Week 8 (Figure 8). Results for secondary censoring for EASI75 at Week 16 were consistent with the those of primary censoring.

Table 16: Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI75 at Week 16

EASI75 at Week 16	PBO (N=93)	BARI 4 mg (N=92)
Primary censoring rule		
Response, n (%) [95% CI]	16 (17.2) [redacted]	29 (31.5) [redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]
Secondary censoring rule		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI75: improvement of at least 75% in Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Figure 8: Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI75 over trial period



p-value obtained by Fisher's exact test. Primary censoring data are presented.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Secondary efficacy endpoints

EASI75 at Week 24

The proportion of patients achieving EASI75 at Week 24 is summarised in Table 17. A numerically higher proportion of patients in the 4 mg baricitinib group (primary censoring) achieved EASI75 at Week 24 ([redacted]%, 95% CI: [redacted], [redacted]) as compared with placebo ([redacted]%, 95% CI: [redacted], [redacted]), but this difference failed to reach statistical significance (odds ratio versus placebo [redacted] [95% CI: [redacted], [redacted], p=[redacted]). Results for secondary censoring for EASI75 at Week 24 were consistent with the those of primary censoring and showed that baricitinib 4 mg did not achieve a statistically significant improvement compared to placebo.

With primary censoring, the EASI75 response rate was lower at Week 24 than at Week 16, but with secondary censoring, the EASI75 response rate was similar to that at Week 16.

Table 17: Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI75 at Week 24

EASI75 at Week 24	PBO (N=93)	BARI 4 mg (N=92)
Primary censoring rule		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]
Secondary censoring rule		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI90: improvement of at least 90% in Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

IGA of 0 or 1 at Weeks 16 and 24

The IGA provides a global assessment of AD severity on a 5-point scale. The proportion of patients achieving IGA of 0 or 1 at Weeks 16 and 24 are summarised in Table 18. In the 4 mg baricitinib group (primary censoring), a statistically significantly higher proportion of patients achieved IGA ≤1 (odds ratio versus placebo: [redacted] [95% CI: [redacted]], p=[redacted]) at Week 16, but this difference was not statistically significant at Week 24. The higher proportion of patients with IGA ≤1 versus placebo was statistically significant at p<[redacted] at Week 4 (Figure 9). Results for secondary censoring for IGA ≤1 at Weeks 16 and 24 were consistent with the those of primary censoring and showed that baricitinib 4 mg achieved a statistically significant improvement compared to placebo at Week 16, but not at Week 24. The IGA ≤1 response rate was lower at Week 24 than at Week 16 with both censoring rules.

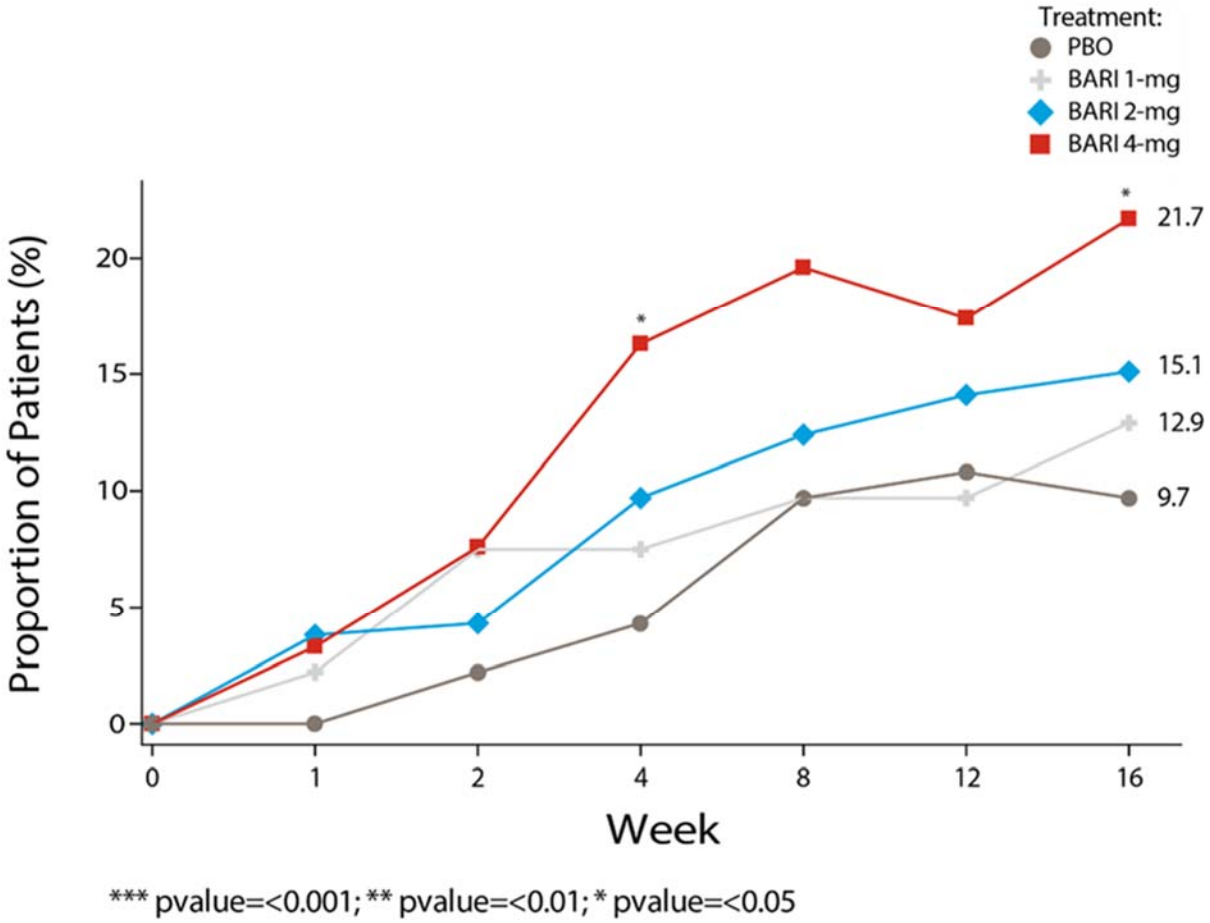
Table 18: Proportion of patients in BREEZE-AD4 (JAIN) achieving IGA ≤1 at Weeks 16 and 24

IGA ≤1	Week 16	Week 24

	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Primary censoring rule				
Response, n (%) [95% CI]	9 (9.7) [redacted]	20 (21.7) [redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
Secondary censoring rule				
Response, n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI75: improvement of at least 75% in Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Figure 9: Proportion of patients in BREEZE-AD4 (JAIN) achieving IGA ≤1 over trial period



p-value obtained by Fisher's exact test. Primary censoring data are presented.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

EASI50 at Week 16 and 24

EASI50 represents an improvement of 50% in EASI score from baseline. The proportion of patients achieving EASI50 at Weeks 16 and 24 using the primary censoring rule is summarised in Table 19. In the 4 mg baricitinib group, a statistically significantly higher proportion of patients achieved EASI50 at Week 16 versus placebo, with 52.2% (95% CI: [redacted]) of patients achieving the endpoint as compared with 35.5% (95% CI: [redacted]) of patients in the placebo group. The odds ratio was [redacted] (95% CI: [redacted]) ([redacted]). At Week 24, a numerically higher proportion of patients in the 4 mg baricitinib group achieved EASI50 as compared with placebo, but this difference did not reach significance (odds ratio versus placebo: [redacted] [95% CI: [redacted]], p=[redacted]). Results for secondary censoring for EASI50 were consistent with the those of primary censoring and showed that baricitinib 4 mg achieved a statistically significant improvement compared with placebo at Week 16 (Table 20).

Table 19: Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI50 at Week 16 and 24 using the primary censoring rule

EASI50	Week 16		Week 24	
	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Response, n (%) [95% CI]	33 (35.5) [redacted]	48 (52.2) [redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI90: improvement of at least 90% in Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Table 20: Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI50 at Week 16 using the secondary censoring rule

EASI50 at Week 16	PBO (N=93)	BARI 4 mg (N=92)
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI90: improvement of at least 90% in Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Source: Secondary censoring data (Data on File).[68]

EASI90 at Week 16 and 24

EASI90 represents an improvement of 90% in EASI score from baseline. The proportion of patients achieving EASI75 at Week 16 is summarised in Table 21. A numerically higher

proportion of patients in the 4 mg baricitinib group (primary censoring) achieved EASI90 at Week 16 (14.1%, 95% CI: [redacted], [redacted]) as compared with placebo (6.5%, 95% CI: [redacted], [redacted]), but this difference failed to reach statistical significance (odds ratio versus placebo [redacted] [95% CI: [redacted], [redacted]], p=[redacted]). Results for secondary censoring for EASI90 were consistent with the those of primary censoring and showed that baricitinib 4 mg did not achieve a statistically significant improvement compared to placebo at Week 16 or Week 24.

Table 21: Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI90 at Week 16 and 24

	Week 16		Week 24	
	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Primary censoring				
Response, n (%) [95% CI]	6 (6.5) [redacted]	13 (14.1) [redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
Secondary censoring				
Response, n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI90: improvement of at least 90% in Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

EASI percent change from baseline at Week 16

The average PCFB in EASI score at Week 16 is summarised in Table 22. Treatment with 4 mg baricitinib was associated with a statistically significant increase in the EASI score PCFB, with a LSM of [redacted] versus [redacted] (95% CI versus placebo: [redacted], p<[redacted]). This improvement versus placebo was statistically significant at [redacted] as early as Week 1 and was maintained through to Week 16 (Figure 10).

Table 22: EASI percent change from baseline at Week 16 in BREEZE-AD4 (JAIN) patients

EASI percent change	PBO (N=54)	BARI 4 mg (N=65)
PCFB, LSM (95% CI vs PBO)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI: Eczema Area and Severity Index; LSM: least squares mean; NA: not applicable; PBO: placebo; PCFB: percent change from baseline.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

Figure 10: Percent change from baseline in EASI score in BREEZE-AD4 (JAIN) patients over trial period

p-values obtained from MMRM models. Primary censoring data are presented.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

SCORAD75 at Weeks 16 and 24

The SCORAD index measures disease severity using six clinical characteristics with higher scores representing higher disease burden. The SCORAD75 outcome represents an improvement of 75% in SCORAD score from baseline. The proportion of patients achieving SCORAD75 at Weeks 16 and 24 is summarised in Table 23. A numerically higher proportion of patients in the 4 mg baricitinib group (primary censoring) achieved SCORAD75 at Week 16 (6.5%, 95% CI: [redacted], [redacted]) as compared with placebo (1.1%, 95% CI: [redacted], [redacted]), but this difference failed to reach statistical significance (odds ratio versus placebo [redacted] [95% CI: [redacted], [redacted]], p=[redacted]). Results at Week 24 were consistent with those at Week 16, with the higher proportion of patients in the 4 mg baricitinib group (primary censoring) achieving SCORAD75 ([redacted]%, 95% CI: [redacted], [redacted]) as compared with placebo ([redacted]%, 95% CI: [redacted], [redacted]) not reaching statistical significance (odds ratio versus placebo [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).The proportion of patients achieving SCORAD75 was significantly higher in the 4 mg baricitinib group than the placebo group at Weeks 8 and 12 (p[redacted]) (Figure 11).

Table 23: Proportion of patients in BREEZE-AD4 (JAIN) achieving SCORAD75 at Weeks 16 and 24

SCORAD75	Week 16		Week 24	
	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Response, n (%) [95% CI]	1 (1.1) [redacted]	6 (6.5) [redacted]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; NA: not applicable; PBO: placebo; SCORAD75: improvement of at least 75% from baseline in SCORing Atopic Dermatitis.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Figure 11: Proportion of patients in BREEZE-AD4 (JAIN) achieving SCORAD75 over trial period

p-value obtained by Fisher's exact test. Primary censoring data are presented.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

Itch NRS ≥4-point improvement at Weeks 1, 2, 4, 16 and 24

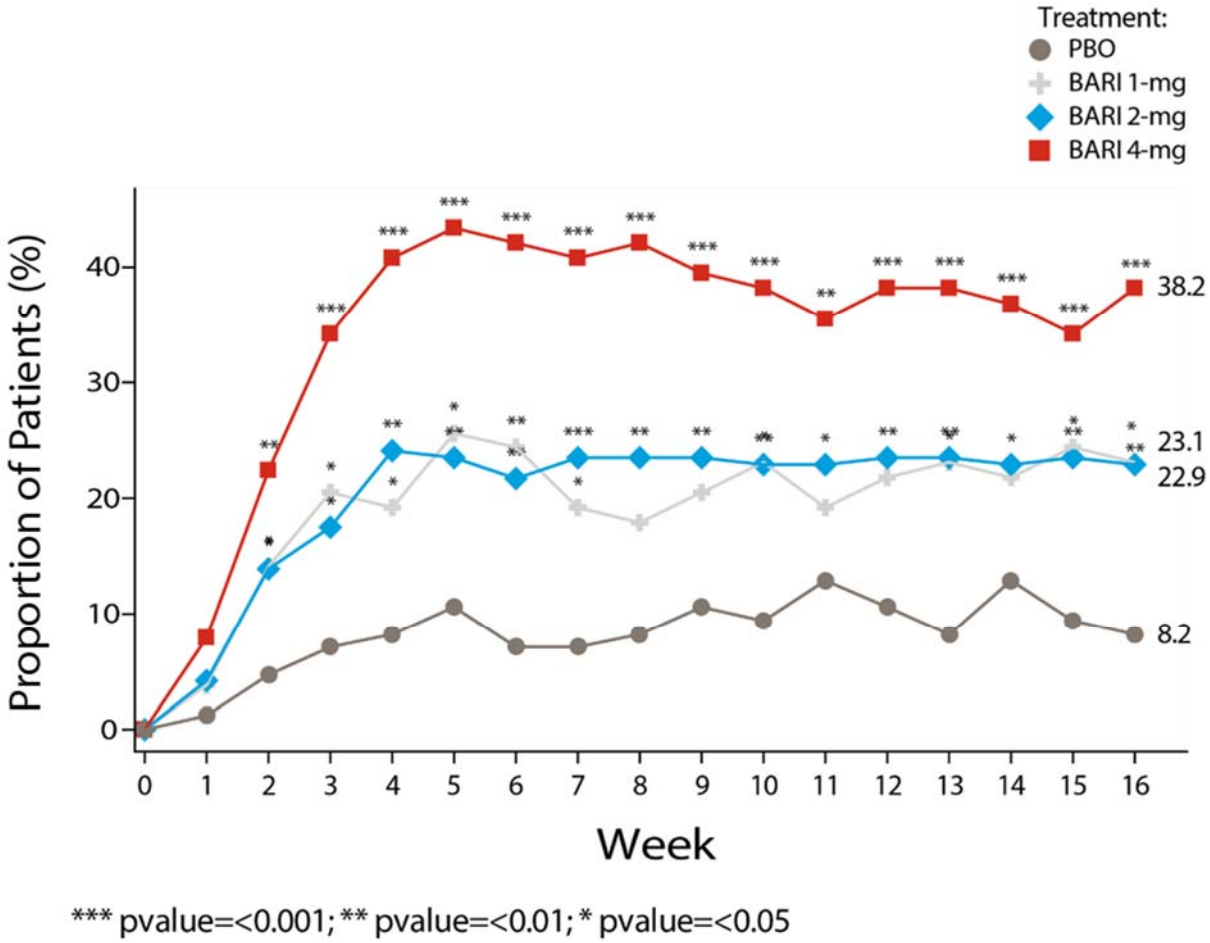
The Itch NRS assesses overall severity of patient itch experienced within the last 24 hours, with higher scores representing worse itch. The proportions of patients achieving a ≥4-point improvement in Itch NRS at Weeks 1, 2, 4, 16 and 24 are summarised in Table 24. In the 4 mg baricitinib group, a statistically significantly higher proportion of patients achieved a ≥4-point improvement in Itch NRS at Week 2 (odds ratio: [redacted] [95% CI: [redacted]], p=[redacted]), Week 4 (odds ratio: [redacted] [95% CI: [redacted]], p<[redacted]), Week 16 (odds ratio: [redacted] [95% CI: [redacted]], p<[redacted]) and Week 24 (odds ratio: [redacted] [95% CI: [redacted]], p=[redacted]). This improvement versus placebo was significant at p<[redacted] as early as Week 2 and was maintained through to Week 16 (Figure 12).

Table 24: Proportion of patients in BREEZE-AD4 (JAIN) with a ≥4 Itch NRS at baseline achieving a ≥4-point Itch NRS improvement at Week 16

≥4-point Itch NRS improvement	Week 1		Week 2		Week 4		Week 16		Week 24	
	PBO (N=85)	4 mg BARI (N=78)	PBO (N=85)	4 mg BARI (N=78)	PBO (N=85)	4 mg BARI (N=78)	PBO (N=85)	4 mg BARI (N=78)	PBO (N=85)	4 mg BARI (N=78)
n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	7 (8.2)	29 (38.2)	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; NA: not applicable; NRS: numeric rating scale; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Figure 12: Proportion of patients in BREEZE-AD4 (JAIN) with a baseline Itch NRS ≥4 achieving a ≥4-point improvement in Itch NRS over trial period



p-value obtained by Fisher's exact test.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Skin pain NRS mean change from baseline at Week 16 and Week 24

The Skin Pain NRS assesses overall severity of patient skin pain experienced within the last 24 hours, with higher scores representing worse pain. The average mean change from baseline (MCFB) in Skin Pain NRS at Week 16 and Week 24 is summarised in Table 25. Treatment with 4 mg baricitinib was associated with a statistically significant increase in the Skin Pain NRS MCFB at Week 16 (LSM: −3.02 versus −1.56; 95% CI versus placebo: [redacted]; p<[redacted]) and Week 24 (LSM: [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] as early as Week 1 and was maintained through to Week 16 (Figure 13).

Table 25: Mean change from baseline in Skin Pain NRS at Week 16 in BREEZE-AD4 (JAIN) patients

Mean change in Skin Pain NRS	Week 16		Week 24	
	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	−1.56 [redacted]	−3.02 [redacted]	[redacted]	[redacted]

p-value[a] vs. PBO				

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; NRS: numeric rating scale; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Figure 13: Mean change from baseline in Skin Pain NRS in BREEZE-AD4 (JAIN) patients over trial period

p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

Item 2 of ADSS mean change from baseline at Weeks 1, 16 and 24

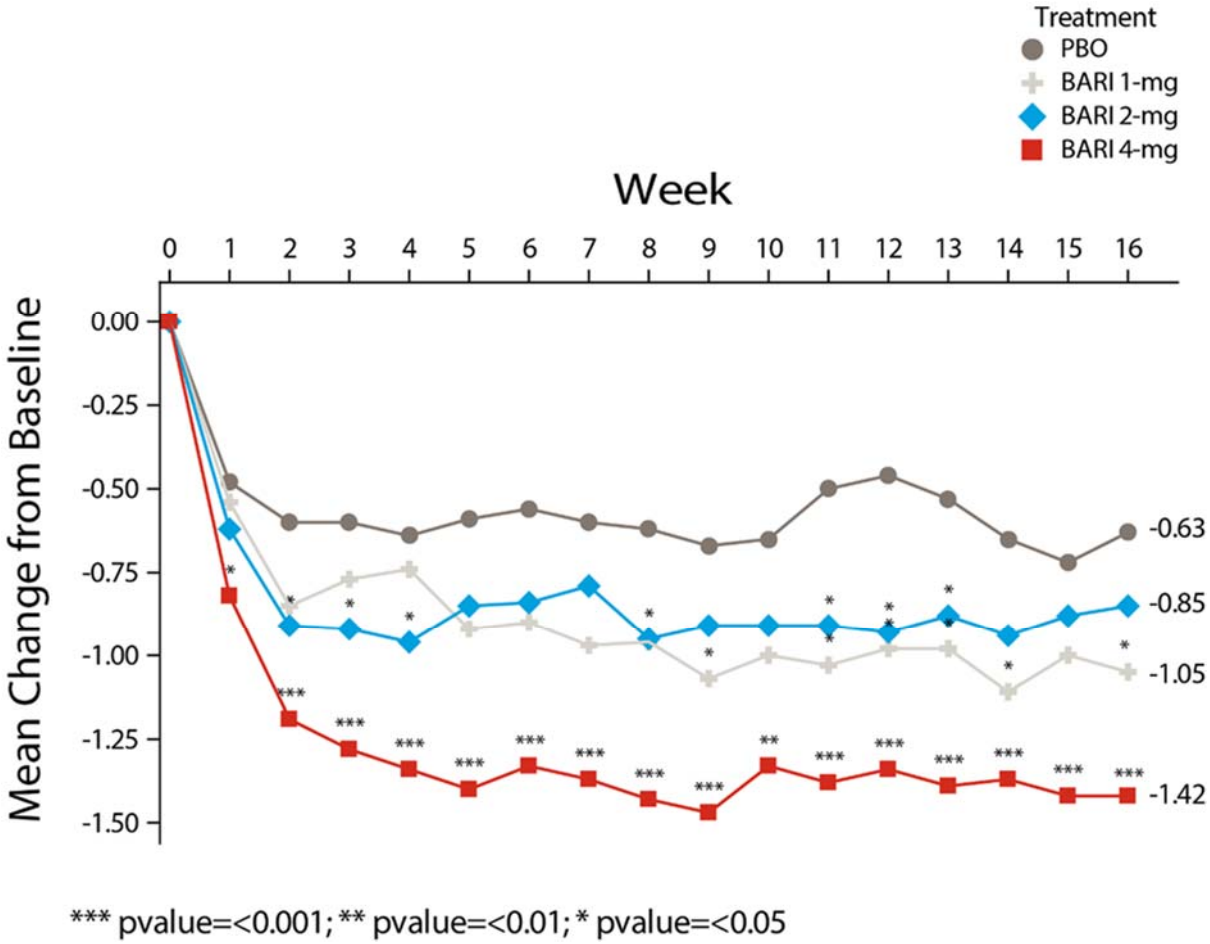
The ADSS assesses the effect of AD-related itch on patient sleep with Item 2 denoting the frequency of waking due to itch the previous night. The average MCFB in ADSS Item 2 at Weeks 1, 16 and 24 is summarised in Table 26. Treatment with 4 mg baricitinib was associated with a statistically significant increase in the MCFB of ADSS Item 2 at Week 1 (LSM versus ; 95% CI versus placebo: ; p=), at Week 16 (LSM −1.42 versus −0.63; 95% CI versus placebo: ; p<) and at Week 24 (LSM versus ; 95% CI versus placebo: ; p=). This improvement versus placebo was statistically significant at p< as early as Week 1 and was maintained at p< from Week 2 through to Week 16 (Figure 14).

Table 26: Mean change from baseline in Item 2 of ADSS at Weeks 1, 16 and 24 in BREEZE-AD4 (JAIN) patients

Mean change in Item 2 of ADSS	Week 1		Week 16		Week 24	
	PBO N=93	BARI 4 mg N=92	PBO N=93	BARI 4 mg N=92	PBO N=93	BARI 4 mg N=92
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]	−0.63 [redacted]	−1.42 [redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: ADSS: atopic dermatitis sleep scale; BARI: baricitinib; CI: confidence interval; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Figure 14: Mean change from baseline in ADSS Item 2 in BREEZE-AD4 (JAIN) patients over trial period



p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Health-related quality of life (HRQoL) endpoints

DLQI at Week 16

The DLQI assess quality of life with higher scores representing greater impairment of life across six domains. The DLQI outcomes at Weeks 16 and 24 are summarised in Table 27.

Treatment with 4 mg baricitinib was associated with a statistically significant increase in the MCFB in DLQI at Week 16 (LSM −7.95 versus −4.95; 95% CI versus placebo: [redacted]; p=[redacted]) and Week 24 (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] from Week 1 to Week 8, and at p<[redacted] from Week 12 to Week 16 (Figure 15).

In the 4 mg baricitinib group, a statistically higher proportion of patients achieved a DLQI score of 0 or 1 at Week 16, with 29.7% (95% CI: [redacted]) of patients achieving the endpoint as compared with 9.7% (95% CI: [redacted]) in the placebo group (odds ratio: [redacted] [95% CI: [redacted]], p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] at Weeks 4, 8 and 16 (Figure 16), but failed to reach significance at Week 24 (p=[redacted]).

Treatment with 4 mg baricitinib was associated with a statistically significantly higher proportion of patients achieving a ≥4-point improvement in DLQI score versus placebo at Week 16 with [redacted]% (95% CI: [redacted]) of patients achieving the endpoint as compared with [redacted]% (95% CI: [redacted]) in the placebo group (odds ratio: [redacted] [95% CI: [redacted]], p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] from Week 1 to Week 4, and at p<0.01 at Weeks 8 and 12 (Figure 17), but failed to reach significance at Week 24 (p=[redacted]). Results for secondary censoring for a ≥4-point improvement in DLQI score were consistent with those of primary censoring and showed that baricitinib 4 mg achieved a statistically significant improvement compared with placebo at Week 16 (Table 28).

Table 27: DLQI outcomes at Weeks 16 and 24 in BREEZE-AD4 (JAIN) patients using the primary censoring rule

DLQI	Week 16		Week 24	
	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB				
MCFB, LSM (95% CI vs. PBO)	−4.95[redacted]	−7.95 [redacted]	[redacted]	[redacted] [redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]
Score of 0 or 1				
Response, n (%) [95% CI]	9 (9.7) [redacted]	27 (29.7) [redacted]	[redacted] [redacted]	[redacted] [redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]
≥4-point improvement[c]	**N=[redacted]**	**N=[redacted]**	**N=[redacted]**	**N=[redacted]**

Response, n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained from MMRM models. [b] p-values obtained by testing odds ratio within logistic regression framework using the primary censoring rule (not presented). [c] Analyses performed on populations with a baseline score ≥4.
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology Life Quality Index; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

Table 28: Proportion of patients in BREEZE-AD4 (JAIN) achieving a ≥4-point improvement in DLQI at Week 16 using the secondary censoring rule

≥4-point improvement in DLQI at Week 16 using the secondary censoring rule[a]	PBO (N=88)	BARI 4 mg (N=83)
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]

[a] Analyses performed on populations with a baseline score ≥4. [b] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology Life Quality Index; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo.
Source: Secondary censoring data (Data on File).[68]

Figure 15: Mean change from baseline in DLQI score in BREEZE-AD4 (JAIN) patients over trial period

p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

Figure 16: Proportion of patients in BREEZE-AD4 (JAIN) achieving a DLQI score of 0 or 1 over trial period

p-value obtained by Fisher's exact test.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

Figure 17: Proportion of patients in BREEZE-AD4 (JAIN) achieving a ≥4-point improvement in DLQI score over trial period

p-value obtained by Fisher's exact test.
Abbreviations: BARI: baricitinib; PBO: placebo.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

EQ-5D-5L at Week 16

The EQ-5D-5L uses a visual analogue scale (VAS) and a descriptive Health Index Score (HIS) to measure self-rated patient health status with lower scores indicating worse disease state. For the clinical effectiveness data presented here, the HIS was based directly on the England-only valuation of EQ-5D-5L by Devlin *et al*, 2018.[69] For the economic evaluation, the EQ-5D-5L scores were cross-walked to EQ-5D-3L and valued using the EQ-5D-3L weights using the algorithm by Dolan *et al*, 1997.[70]

The average MCFB in the two components of the EQ-5D-5L at Week 16 are summarised in Table 29. At Week 16, treatment with 4 mg baricitinib was associated with a statistically significant increase in the Health Index Score (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]), but the difference in the VAS score failed to reach statistical significance (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]).

Table 29: Mean change from baseline in EQ-5D-5L at Week 16 in BREEZE-AD4 (JAIN) patients

EQ-5D-5L	VAS Score		Health Index Score (England algorithm)	
	PBO (N=93)	BARI 4 mg (N=92)	PBO (N=93)	BARI 4 mg (N=92)
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; EQ-5D-5L: 5-level EuroQol 5 Dimensions; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; VAS: visual analogue scale.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.[54]

B.2.6.2 Combination therapy trial: BREEZE-AD7 (JAIY)

In this section, results for all efficacy endpoints reported use the conservative primary censoring rule. Additionally, results using the secondary censoring rule have been presented alongside the primary censoring data for IGA, EASI and DLQI outcomes.

Primary efficacy endpoint: IGA of 0 or 1 at Week 16

The proportion of patients achieving IGA of 0 or 1 at Week 16 are summarised in Table 30. In the 4 mg baricitinib group (primary censoring), a statistically significantly higher proportion of patients achieved IGA ≤1 at Week 16 versus placebo, with [redacted]% (95% CI: [redacted]) of patients achieving the endpoint as compared with [redacted]% (95% CI: [redacted]) of patients in the placebo group. The odds ratio was [redacted] (95% CI: [redacted]) ([redacted]). This improvement versus placebo was statistically significant at [redacted] as early as Week 4 and was maintained through to Week 16 (Figure 18). Results for secondary censoring for IGA ≤1 at Week 16 were consistent with those of primary censoring.

Table 30: Proportion of patients in BREEZE-AD7 (JAIY) achieving IGA ≤1 at Week 16

IGA ≤1 at Week 16	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
Primary censoring		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]
Secondary censoring		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; IGA: Investigator's Global Assessment; NA: not applicable; PBO: placebo; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 18: Proportion of patients in BREEZE-AD7 (JAIY) achieving IGA ≤1 over trial period

p-value obtained by Fisher's exact test. Primary censoring data are presented.
Abbreviations: IGA: Investigator's Global Assessment.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Secondary efficacy endpoints

EASI score at Week 16

The proportion of patients achieving EASI50, EASI75 and EASI90 at Week 16 is summarised in Table 31. In the 4 mg baricitinib group, a statistically significantly higher proportion of patients achieved EASI50 (odds ratio versus placebo: [redacted] [95% CI: [redacted]], p<[redacted]), EASI75 (odds ratio versus placebo: [redacted] [95% CI: [redacted]], p<[redacted]) and EASI90 (odds ratio versus placebo: [redacted] [95% CI: [redacted]], p=[redacted]). The improvement in EASI75 versus placebo was statistically significant at p<[redacted] as early as Week 2 and was maintained to Week 16 (Figure 19). The improvement in EASI90 versus placebo was statistically significant at p[redacted] at Weeks 4, 8 and 12. Results for secondary censoring for EASI50, EASI75 and EASI90 at Week 16 were consistent with the those of primary censoring.

Table 31: Proportion of patients in BREEZE-AD7 (JAIY) achieving EASI50, EASI75 and EASI90 at Week 16

	EASI50		EASI75		EASI90	
	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
Primary censoring						
Response, n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secondary censoring						
Response, n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI(75/90): (improvement of at least 75%/90% in) Eczema Area and Severity Index; NA: not applicable; PBO: placebo; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 19: Proportion of patients in BREEZE-AD7 (JAIY) achieving EASI75 over trial period

p-value obtained by Fisher's exact test. Primary censoring data are presented.
Abbreviations: EASI75: improvement of at least 75% in Eczema Area and Severity Index.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

EASI percent change from baseline at Week 16

The average PCFB in EASI score at Week 16 is summarised in Table 32. Treatment with 4 mg baricitinib was associated with a statistically significant increase in the EASI score PCFB, with a LSM of [redacted] versus [redacted] (95% CI versus placebo: [redacted], p<[redacted]). This improvement versus placebo was statistically significant at [redacted] as early as Week 1 and was maintained through to Week 16 (Figure 20).

Table 32: EASI percent change from baseline at Week 16 in BREEZE-AD7 (JAIY) patients

EASI percent change	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
PCFB, LSM (95% CI vs PBO)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI: Eczema Area and Severity Index; LSM: least-squares mean; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; PCFB: percent change from baseline; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 20: Percent change from baseline in EASI score in BREEZE-AD7 (JAIY) patients over trial period

p-values obtained from MMRM models.
Abbreviations: EASI: Eczema Area and Severity Index.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

SCORAD75 at Week 16

The proportion of patients achieving SCORAD75 at Week 16 is summarised in Table 33. In the 4 mg baricitinib group, a statistically significantly higher proportion of patients achieved SCORAD75 with [redacted]% (95% CI: [redacted]) of patients achieving the endpoint as compared with [redacted]% (95% CI: [redacted]) in the placebo group. The odds ratio was [redacted] (95% CI: [redacted]) (p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] at Weeks 4, 8 and 16 and at p<[redacted] at Week 12.

Table 33: Proportion of patients in BREEZE-AD7 (JAIY) achieving SCORAD75 at Week 16

SCORAD75	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).

Abbreviations: BARI: baricitinib; CI: confidence interval; NA: not applicable; PBO: placebo; SCORAD75: improvement of at least 75% from baseline in SCORing Atopic Dermatitis; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Itch NRS ≥4-point improvement from baseline at Day 2 and Weeks 1, 2, 4 and 16

The proportions of patients achieving a ≥4-point improvement in Itch NRS at Weeks 1, 2, 4 and 16 are summarised in Table 34. In the 4 mg baricitinib group, a statistically significantly higher proportion of patients achieved a ≥4-point improvement in Itch NRS at Week 2 (odds ratio: [redacted] [95% CI: [redacted], p<[redacted]), Week 4 (odds ratio: [redacted] [95% CI: [redacted], p<[redacted]) and Week 16 (odds ratio: [redacted] [95% CI: [redacted], p<[redacted]). The difference in responses was not statistically different between the placebo and 4 mg baricitinib groups at Day 2 (odds ratio: [redacted] [95% CI: [redacted], p=[redacted]) and Week 1 (odds ratio: [redacted] [95% CI: [redacted], p=[redacted]). The higher proportion of patients achieving a ≥4-point improvement versus placebo was statistically significant at p<[redacted] as early as Week 3 and was maintained through to Week 16 (Figure 21).

Table 34: Proportion of patients in BREEZE-AD7 (JAIY) with a ≥4 Itch NRS at baseline achieving a ≥4-point Itch NRS improvement at Day 2 and Weeks 1, 2, 4 and 16

≥4-point Itch NRS improvement	Day 2		Week 1		Week 2		Week 4		Week 16	
	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
n (%) [95% CI]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; NA: not applicable; NRS: numeric rating scale; PBO: placebo; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 21: Proportion of patients in BREEZE-AD7 (JAIY) with a baseline Itch NRS ≥4 achieving a ≥4-point improvement in Itch NRS over trial period

p-value obtained by Fisher's exact test.
Abbreviations: NRS: numeric rating scale.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Skin Pain NRS at Week 16

The average MCFB in Skin Pain NRS at Week 16 is summarised in Table 35. Treatment with 4 mg baricitinib was associated with a statistically significant reduction in the Skin Pain NRS MCFB, with a LSM of [redacted] versus [redacted] (95% CI versus placebo: [redacted], p<[redacted]). This improvement versus placebo was statistically significant at p<[redacted] as early as Week 1 and was maintained through to Week 16 (Figure 22).

Table 35: Mean change from baseline in Skin Pain NRS at Week 16 in BREEZE-AD7 (JAIY) patients

Mean change in Skin Pain NRS	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
Baseline mean	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; NRS: numeric rating scale; PBO: placebo; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 22: Mean change from baseline in Skin Pain NRS in BREEZE-AD7 (JAIY) patients over trial period

p-values obtained from MMRM models.
Abbreviations: NRS: numeric rating scale.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Item 2 of ADSS mean change from baseline at Week 16

The average MCFB in ADSS Item 2 at Week 16 is summarised in Table 36. Treatment with 4 mg baricitinib was associated with a statistically significant increase in the MCFB in ADSS Item 2 at Week 1 (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]) and at Week 16 (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p<[redacted]). This improvement versus placebo was statistically significant at p<[redacted] as early as Week 1 and was maintained at p[redacted] from Week 2 through to Week 16 (Figure 23).

Table 36: Mean change from baseline in Item 2 of ADSS at Week 16 in BREEZE-AD7 (JAIY) patients

Mean change in Item 2 ADSS	Week 1		Week 16	
	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: ADSS: atopic dermatitis sleep scale; BARI: baricitinib; CI: confidence interval; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 23: Mean change from baseline in ADSS Item 2 in BREEZE-AD7 (JAIY) patients over trial period

p-values obtained from MMRM models.
Abbreviations: ADSS: atopic dermatitis sleep scale.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Health-related quality of life (HRQoL) endpoints

DLQI at Week 16

The DLQI outcomes at Week 16 are summarised in Table 37. Treatment with 4 mg baricitinib was associated with a statistically significant increase in the MCFB in DLQI at Week 16, with a LSM of [redacted] versus [redacted] (95% CI versus placebo: [redacted], p<[redacted]). This improvement versus placebo was statistically significant at p<[redacted] as early as Week 1 and was maintained through to Week 16 (Figure 24).

In the 4 mg baricitinib group, a statistically higher proportion of patients achieved a DLQI score of 0 or 1 at Week 16, with [redacted]% (95% CI: [redacted]) of patients achieving the endpoint as compared with [redacted]% (95% CI: [redacted]) in the placebo group (odds ratio: [redacted] [95% CI: [redacted], p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] at Week 2 and at p<[redacted] at Weeks 4, 8 and 16.

Treatment with 4 mg baricitinib was associated with a statistically significantly higher proportion of patients achieving a ≥4-point improvement in DLQI score versus placebo, with [redacted]% (95% CI: [redacted]) of patients achieving the endpoint as compared with [redacted]% (95% CI: [redacted]) in the placebo group (odds ratio: [redacted] [95% CI: [redacted], p=[redacted]). This improvement versus placebo was statistically significant at p<[redacted] as early as Week 2 and was maintained through to Week 16 (Figure 25). Results for secondary censoring for ≥4-point improvement in DLQI score at Week 16 were consistent with the those of primary censoring.

Table 37: DLQI outcomes at Week 16 in BREEZE-AD7 (JAIY) patients

DLQI	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])

Baseline mean	[redacted]	[redacted]
MCFB		
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]
Score of 0 or 1		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]
≥4-point improvement[c]		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]
≥4-point improvement using secondary censoring rule[c]		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]

[a] p-values obtained from MMRM models. [b] p-values obtained by testing odds ratio within logistic regression framework (not presented). [c] Analyses performed on populations with a baseline score ≥4 (PBO: N=102; BARI 4 mg: N=105).
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology Life Quality Index; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; TCS: topical corticosteroids.
Source: BREEZE-AD7 (JAIY) Clinical Study Report,[55] Secondary censoring data (Data on File).[68]

Figure 24: Mean change from baseline in DLQI score in BREEZE-AD7 (JAIY) patients over trial period

p-values obtained from MMRM models.

Abbreviations: DLQI: Dermatology Life Quality Index.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

Figure 25: Proportion of patients in BREEZE-AD7 (JAIY) achieving a ≥4-point improvement in DLQI score over trial period

p-value obtained by Fisher's exact test.
Abbreviations: DLQI: Dermatology Life Quality Index.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

EQ-5D-5L at Week 16

For the clinical effectiveness data presented here, the HIS was based directly on the England-only valuation of EQ-5D-5L by Devlin *et al*, 2018.[69] For the economic evaluation, the EQ-5D-5L scores were cross-walked to EQ-5D-3L and valued using the EQ-5D-3L weights using the algorithm by Dolan *et al*, 1997.[70]

The average MCFB in the two components of the EQ-5D-5L at Week 16 are summarised in Table 38. Treatment with 4 mg baricitinib was associated with a statistically significant increase in VAS (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]) and the Health Index Score (LSM [redacted] versus [redacted]; 95% CI versus placebo: [redacted]; p=[redacted]) at Week 16. This improvement in VAS versus placebo was statistically significant at p<[redacted] as early as Week 2 and was maintained at p<[redacted] through to Week 16, and the improvement in Health Index Score was statistically significant at p<[redacted] as early as Week 1 and was maintained through to Week 16.

Table 38: Mean change from baseline in EQ-5D-5L at Week 16 in BREEZE-AD7 (JAIY) patients

EQ-5D-5L	VAS Score		Health Index Score (England algorithm)	
	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]

MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; EQ-5D-5L: 5-level EuroQol 5 Dimensions; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; TCS: topical corticosteroids; VAS: visual analogue scale.
Source: BREEZE-AD7 (JAIY) Clinical Study Report.[55]

B.2.6.3 Monotherapy trials: BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM)

Primary efficacy endpoint: IGA of 0 or 1 at Week 16

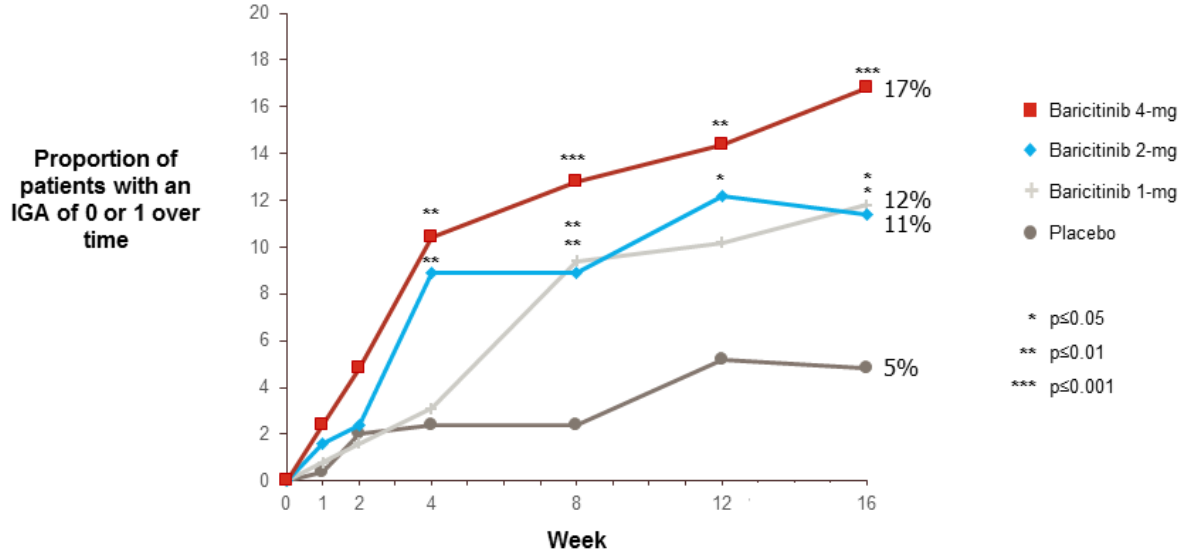
The proportion of patients in BREEZE-AD1 and -AD2 achieving IGA of 0 or 1, representing clear to almost clear disease, at Week 16 are summarised in Table 39. In both trials, a statistically significantly higher proportion of patients in the 4 mg baricitinib group achieved IGA ≤1 at Week 16 versus placebo (p<0.001 in BREEZE-AD1, p=0.001 in BREEZE-AD2); this improvement versus placebo was statistically significant at p<0.01 as early as Week 4 and was maintained through to Week 16 (Figure 26 and Figure 27).

Table 39: Proportion of patients in BREEZE-AD1 and -AD2 achieving IGA ≤1 at Week 16

IGA ≤1 at Week 16	BREEZE-AD1 (JAHL)[56]		BREEZE-AD2 (JAHM)[57]	
	PBO (N=249)	BARI 4 mg (N=125)	PBO (N=244)	BARI 4 mg (N=123)
Response, n (%) [95% CI]	12 (4.8) [2.8, 8.2]	21 (16.8) [11.3, 24.3]	11 (4.5) [2.5, 7.9]	17 (13.8) [8.8, 21.0]
Difference vs PBO, % (95% CI)	NA	12.0 (5.5, 19.8)	NA	9.3 (3.3, 16.8)
Odds ratio vs PBO (95% CI)	NA	4.10 (1.93, 8.70)	NA	3.64 (1.64, 8.05)
p-value[a] vs PBO	NA	<0.001	NA	0.001

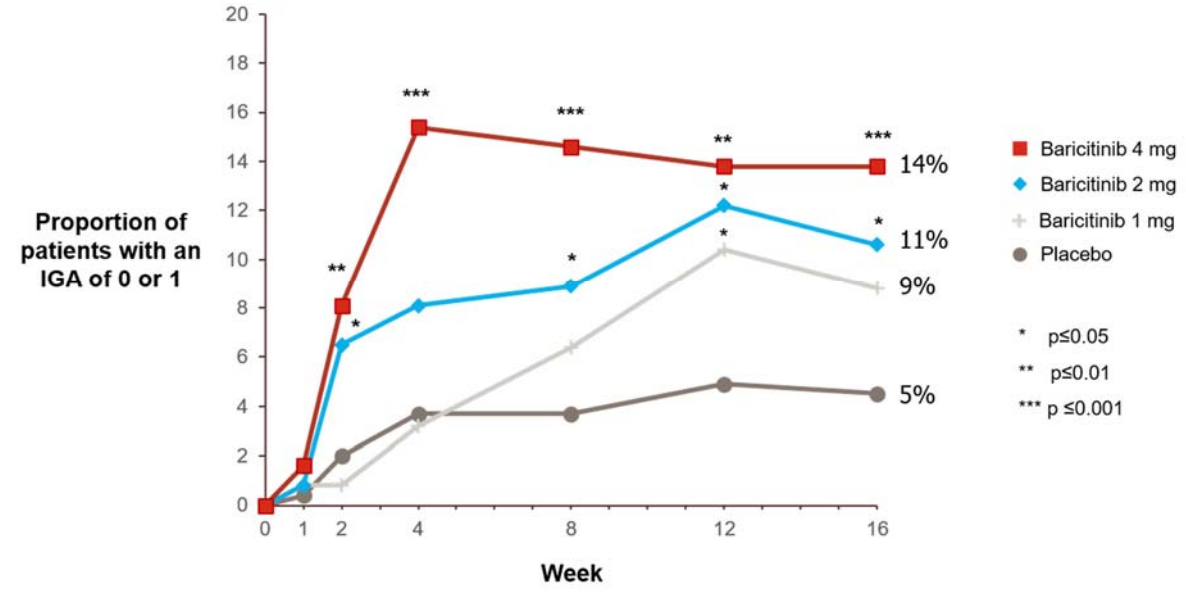
[a] p-value obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; IGA: Investigator's Global Assessment; NA: not applicable; PBO: placebo.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 26: Proportion of patients in BREEZE-AD1 (JAHL) achieving IGA ≤1 over trial period



p-value obtained by Fisher's exact test.
Abbreviations: IGA: Investigator's Global Assessment.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 27: Proportion of patients in BREEZE-AD2 (JAHM) achieving IGA ≤1 over trial period



p-value obtained by Fisher's exact test.
Abbreviations: IGA: Investigator's Global Assessment.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Secondary efficacy endpoints

EASI score at Week 16

The proportion of patients achieving EASI50, EASI75 and EASI90 in BREEZE-AD1 and -AD2 at Week 16 is summarised in Table 40. In both trials, a statistically significantly higher proportion of patients in the 4 mg baricitinib group achieved EASI50, EASI75 and EASI90 at Week 16 as

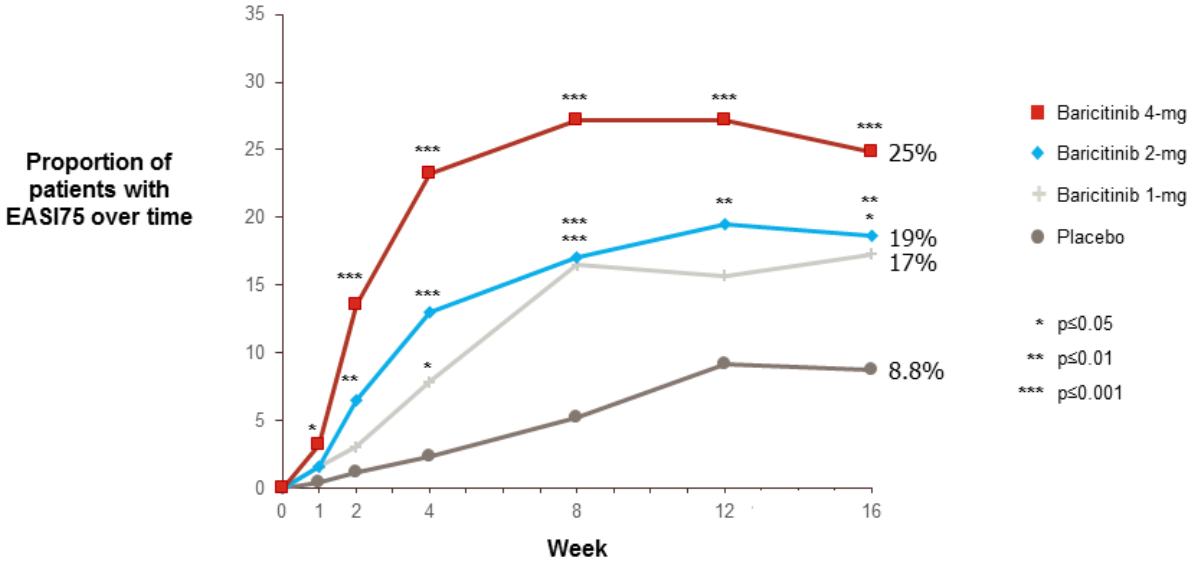
compared with placebo (all p[redacted]). The improvement in EASI75 versus placebo was statistically significant at p<0.05 at Week 1 and at p<0.001 from Week 2 onwards (Figure 28 and Figure 29).

Table 40: Proportion of patients in BREEZE-AD1 and -AD2 achieving EASI50, EASI75 and EASI90 at Week 16

	EASI50		EASI75		EASI90	
	PBO	BARI 4 mg	PBO	BARI 4 mg	PBO	BARI 4-mg
BREEZE-AD1 (JAHL)[56]	**N=[redacted]**	**N=[redacted]**	**N=249**	**N=125**	**N=249**	**N=125**
Response, n (%) [95% CI]	[redacted]	[redacted]	22 (8.8) [5.9, 13.0]	31 (24.8) [18.1, 33.0]	12 (4.8) [2.8, 8.2]	20 (16.0) [10.6, 23.4]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	NA	16.0 (8.0, 24.7)	NA	11.2 (4.8, 18.9)
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	NA	3.72 (2.01, 6.89)	NA	4.13 (1.91, 8.91)
p-value[a] vs PBO	[redacted]	[redacted]	NA	<0.001	NA	<0.001
BREEZE-AD2 (JAHM)[57]	**N=[redacted]**	**N=[redacted]**	**N=244**	**N=123**	**N=244**	**N=123**
Response, n (%) [95% CI]	[redacted]	[redacted]	15 (6.1) [3.8, 9.9]	26 (21.1) [14.9, 29.2)	6 (2.5) [1.1, 5.3]	16 (13.0) [8.2, 20.1]
Difference vs PBO, % (95% CI)	[redacted]	[redacted]	NA	15.0 (7.7, 23.4)	NA	10.5 (5.0, 17.8)
Odds ratio vs PBO (95% CI)	[redacted]	[redacted]	NA	4.41 (2.22, 8.76)	NA	6.20 (2.42, 15.91)
p-value[a] vs PBO	[redacted]	[redacted]	NA	<0.001	NA	<0.001

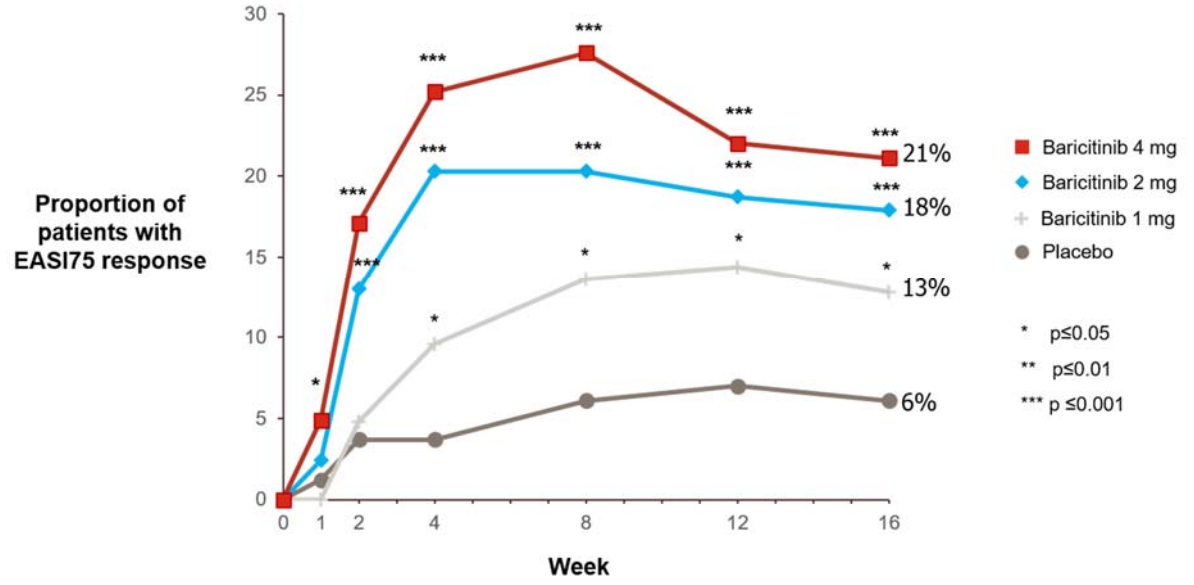
[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI(75/90): (improvement of at least 75%/90% in) Eczema Area and Severity Index; NA: not applicable; PBO: placebo.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 28: Proportion of patients in BREEZE-AD1 (JAHL) achieving EASI75 over trial period



p-value obtained by Fisher's exact test.
Abbreviations: EASI75: improvement of at least 75% in Eczema Area and Severity Index.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 29: Proportion of patients in BREEZE-AD2 (JAHM) achieving EASI75 over trial period



p-value obtained by Fisher's exact test.
Abbreviations: EASI75: improvement of at least 75% in Eczema Area and Severity Index.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

EASI percent change from baseline at Week 16

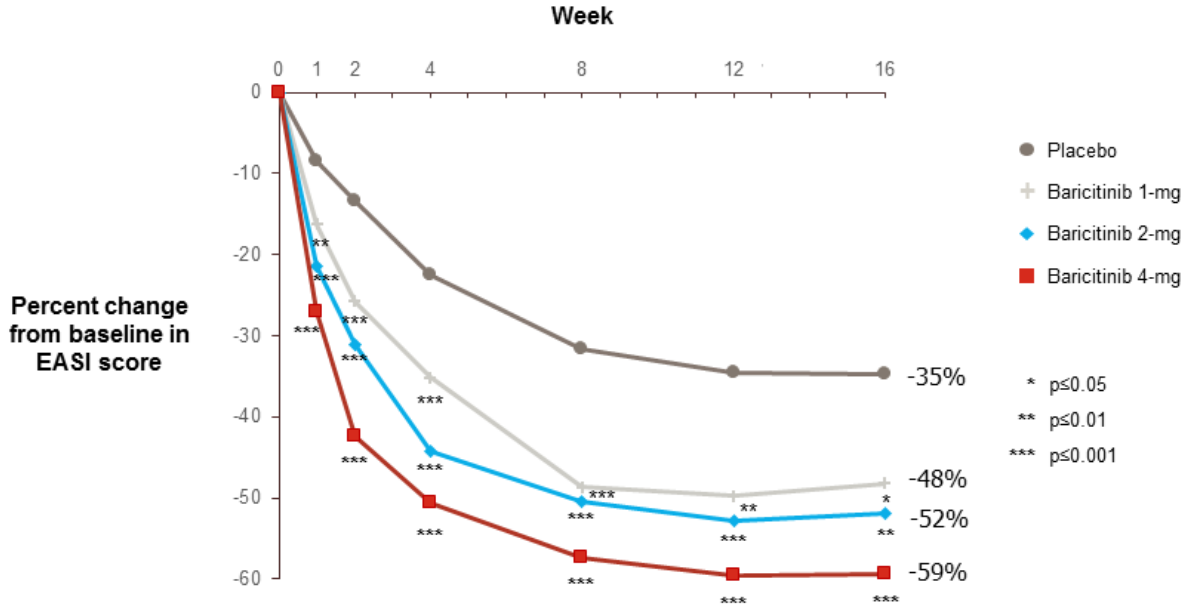
The average percent change from baseline (PCFB) in EASI score at Week 16 for patients in BREEZE-AD1 and -AD2 is summarised in Table 41. In both trials, treatment with 4 mg baricitinib was associated with a statistically significant increase in the EASI score PCFB at Week 16 (both p<0.001), and this improvement versus placebo was statistically significant at p<0.001 as early as Week 1 and was maintained through to Week 16 (Figure 30 and Figure 31).

Table 41: EASI percent change from baseline at Week 16 in BREEZE-AD1 and -AD2 patients

EASI percent change	BREEZE-AD1 (JAHL)[56]		BREEZE-AD2 (JAHM)[57]	
	PBO (N=249)	BARI 4 mg (N=125)	PBO (N=244)	BARI 4 mg (N=123)
PCFB, LSM (95% CI vs PBO)	−34.82 (NA)	−59.36 (−34.84, −14.24)	−28.91 (NA)	−54.88 (−38.29, −13.65)
p-value[a] vs. PBO	NA	<0.001	NA	<0.001

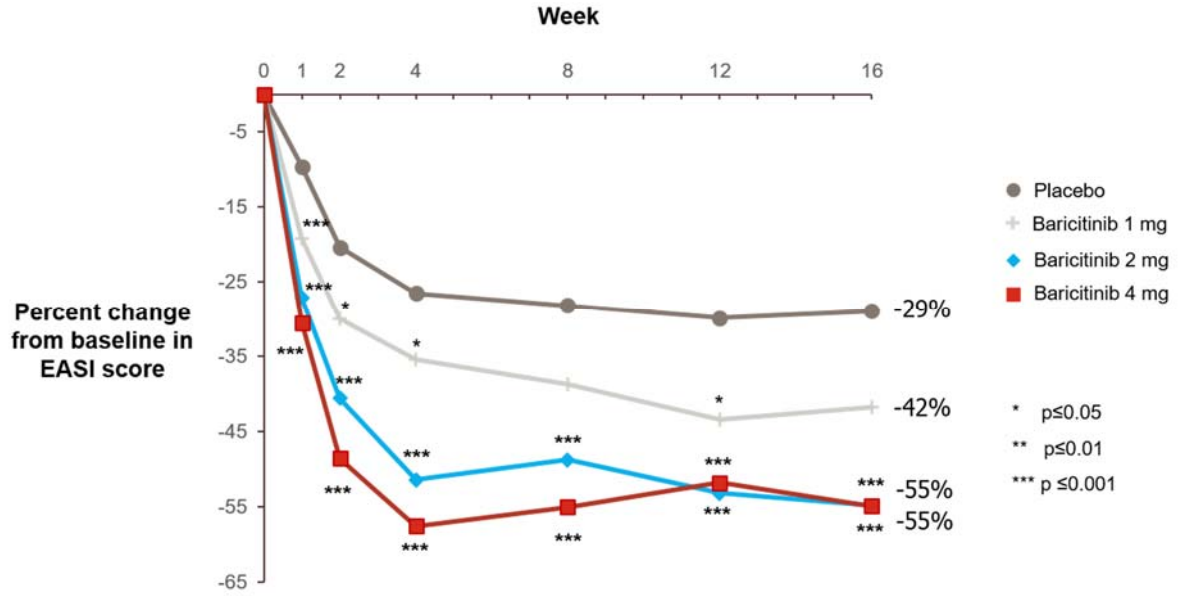
[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; EASI: Eczema Area and Severity Index; LSM: least-squares mean; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; PCFB: percent change from baseline.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 30: Percent change from baseline in EASI score in BREEZE-AD1 (JAHL) patients over trial period



p-values obtained from MMRM models.
Abbreviations: EASI: Eczema Area and Severity Index.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 31: Percent change from baseline in EASI score in BREEZE-AD2 (JAHM) patients over trial period



p-values obtained from MMRM models.
Abbreviations: EASI: Eczema Area and Severity Index.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

SCORAD75 at Week 16

The proportion of patients in BREEZE-AD1 and -AD2 achieving SCORAD75 at Week 16 is summarised in Table 42. A statistically significantly higher proportion of patients in the 4 mg baricitinib group achieved SCORAD75 at Week 16 as compared with placebo, with odds ratios of 8.76 (95% CI: 2.68, 28.58) and 7.40 (95% CI: 2.51, 21.83) (both p<0.001). In BREEZE-AD1, this improvement versus placebo was statistically significant at p<0.05 from Week 4 and was maintained through to Week 16. In BREEZE-AD2, this improvement versus placebo was statistically significant at p<0.01 from Week 4 and was maintained through to Week 16.

Table 42: Proportion of patients in BREEZE-AD1 and -AD2 achieving SCORAD75 at Week 16

SCORAD75	BREEZE-AD1 (JAHL)[56]		BREEZE-AD2 (JAHM)[57]	
	PBO (N=249)	BARI 4 mg (N=125)	PBO (N=244)	BARI 4 mg (N=123)
Response, n (%) [95% CI]	3 (1.2) [0.4, 3.5]	13 (10.4) [6.2, 17.0]	4 (1.6) [0.6, 4.1]	14 (11.4) [6.9, 18.2]
Difference vs. PBO, % (95% CI)	NA	9.2 (4.4, 15.8)	NA	9.7 (4.6, 16.6)
Odds ratio vs. PBO (95% CI)	NA	8.76 (2.68, 28.58)	NA	7.40 (2.51, 21.83)
p-value[a] vs. PBO	NA	<0.001	NA	<0.001

[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; NA: not applicable; PBO: placebo; SCORAD75: improvement of at least 75% from baseline in SCORing Atopic Dermatitis.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Itch NRS ≥4-point improvement at Weeks 1, 2, 4 and 16

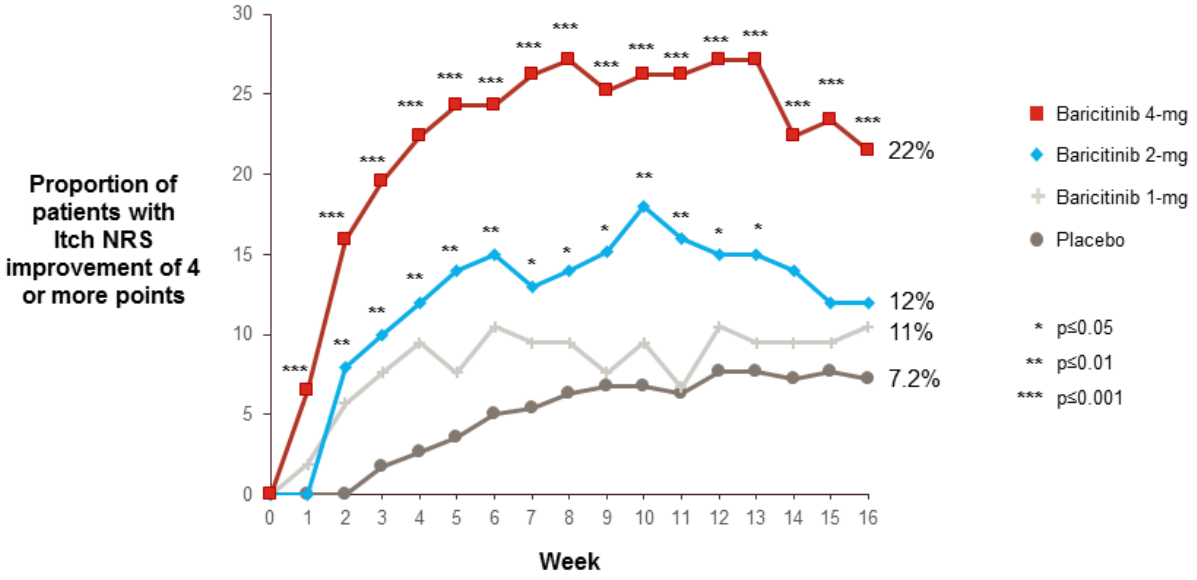
The proportions of patients in BREEZE-AD1 and -AD2 achieving a ≥4-point improvement in Itch NRS at Weeks 1, 2, 4 and 16 are summarised in Table 43. In both trials, a statistically significantly higher proportion of patients in the 4 mg baricitinib group as compared with placebo achieved a ≥4-point improvement in Itch NRS at Week 1 (p=0.010 in BREEZE-AD1, p=0.033 in BREEZE-AD2). From Week 2 onwards, this proportion was significant to ≤0.001 in both trials (Figure 32 and Figure 33).

Table 43: Proportion of patients in BREEZE-AD1 and -AD2 with a ≥4 Itch NRS at baseline achieving a ≥4-point Itch NRS improvement at Weeks 1, 2, 4 and 16

≥4-point Itch NRS improvement	Week 1		Week 2		Week 4		Week 16	
	PBO	4 mg BARI	PBO	4 mg BARI	PBO	4 mg BARI	PBO	4 mg BARI
BREEZE-AD1 (JAHL)[56]	**N=222**	**N=107**	**N=222**	**N=107**	**N=222**	**N=107**	**N=222**	**N=107**
n (%) [95% CI]	0 (0.0) [0.0, 0.0]	7 (6.5) [3.2, 12.9]	0 (0.0) [0.0, 0.0]	17 (15.9) [10.2, 24.0]	6 (2.7) [1.2, 5.8]	24 (22.4) [15.6, 31.2]	16 (7.2) [4.5, 11.4]	23 (21.5) [14.8, 30.2]
Difference vs. PBO, % (95% CI)	NA	6.5 (2.8, 12.9)	NA	15.9 (9.9, 24.0)	NA	19.7 (12.2, 28.6)	NA	14.3 (6.4, 23.4)
Odds ratio vs. PBO (95% CI)	NA	31.93 (2.29, >99.99)	NA	88.26 (5.67, >99.99)	NA	10.00 (4.07, 24.56)	NA	4.80 (2.47, 9.32)
p-value[a] vs. PBO	NA	0.010	NA	0.001	NA	<0.001	NA	<0.001
BREEZE-AD2 (JAHM)[57]	**N=213**	**N=107**	**N=213**	**N=107**	**N=213**	**N=107**	**N=213**	**N=107**
n (%) [95% CI]	1 (0.5) [0.1, 2.6]	4 (3.7) [1.5, 9.2]	2 (0.9) [0.3, 3.4]	11 (10.3) [5.8, 17.5]	5 (2.3) [1.0, 5.4]	20 (18.7) [12.4, 27.1]	10 (4.7) [2.6, 8.4]	20 (18.7) [12.4, 27.1]
Difference vs. PBO, % (95% CI)	NA	3.3 (0.1, 8.8)	NA	9.3 (4.3, 16.6)	NA	16.3 (9.4, 24.9)	NA	14.0 (6.7, 22.7)
Odds ratio vs. PBO (95% CI)	NA	6.65 (1.17, 37.99)	NA	11.03 (2.83, 42.90)	NA	9.93 (3.74, 26.37)	NA	4.91 (2.22, 10.86)
p-value[a] vs. PBO	NA	0.033	NA	<0.001	NA	<0.001	NA	<0.001

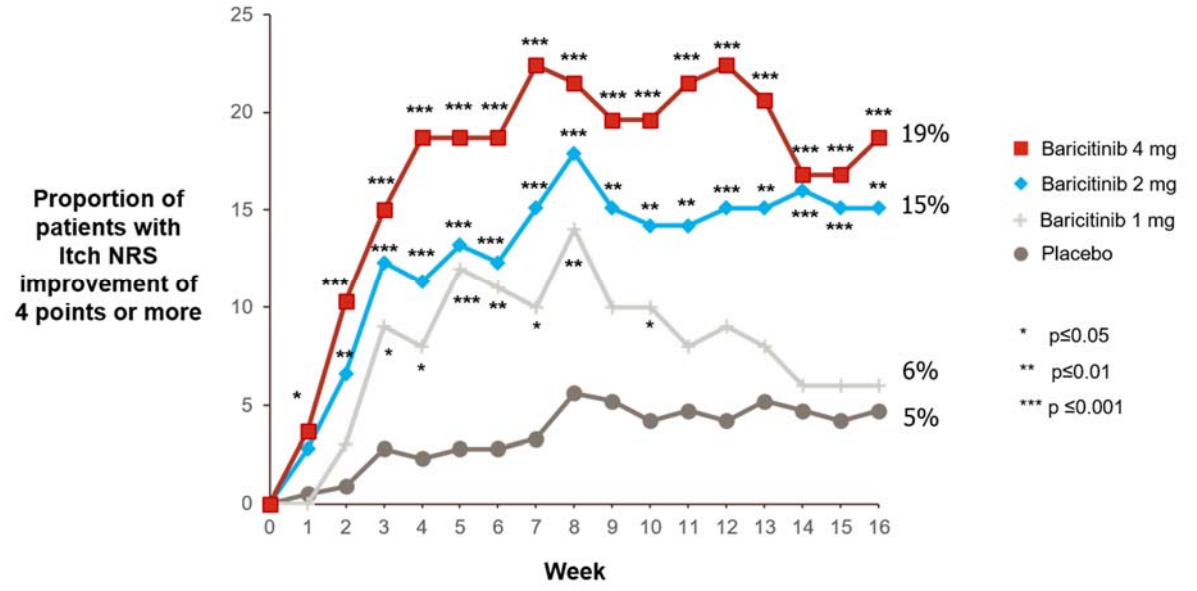
[a] p-values obtained by testing odds ratio within logistic regression framework (not presented).
Abbreviations: BARI: baricitinib; CI: confidence interval; NA: not applicable; NRS: numeric rating scale; PBO: placebo.
Source: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 32: Proportion of patients in BREEZE-AD1 (JAHL) with a baseline Itch NRS ≥4 achieving a ≥4-point improvement in Itch NRS over trial period



p-value obtained by Fisher's exact test.
Abbreviations: NRS: numeric rating scale.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 33: Proportion of patients in BREEZE-AD2 (JAHM) with a baseline Itch NRS ≥4 achieving a ≥4-point improvement in Itch NRS over trial period



p-value obtained by Fisher's exact test.
Abbreviations: NRS: numeric rating scale.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Skin Pain NRS at Week 16

The average MCFB in Skin Pain NRS in patients in the BREEZE-AD1 and -AD2 trials at Week 16 is summarised in Table 44. In both trials, treatment with 4 mg baricitinib was associated with a statistically significant increase in the Skin Pain NRS MCFB as compared with placebo (p=0.002

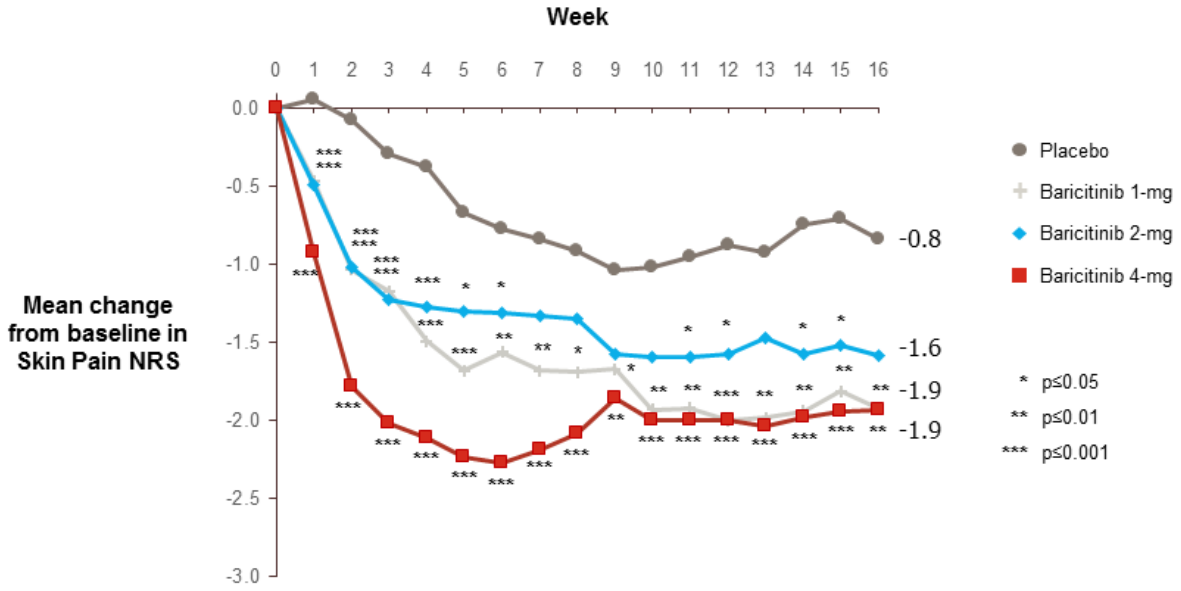
in BREEZE-AD1, p<0.001 in BREEZE-AD2). A significant reduction to <0.01 was maintained until Week 16 (Figure 34 and Figure 35).

Table 44: Mean change from baseline in Skin Pain NRS at Week 16 in BREEZE-AD1 and -AD2 patients

Mean change in Skin Pain NRS	BREEZE-AD1 (JAHL)[56]		BREEZE-AD2 (JAHM)[57]	
	PBO (N=249)	BARI 4 mg (N=125)	PBO (N=244)	BARI 4 mg (N=123)
Baseline mean	6.07	5.74	6.21	5.95
MCFB, LSM (95% CI vs. PBO)	−0.84 (NA)	−1.93 (−1.79, −0.39)	−0.86 (NA)	−2.49 (−2.37, −0.87)
p-value[a] vs. PBO	NA	0.002	NA	<0.001

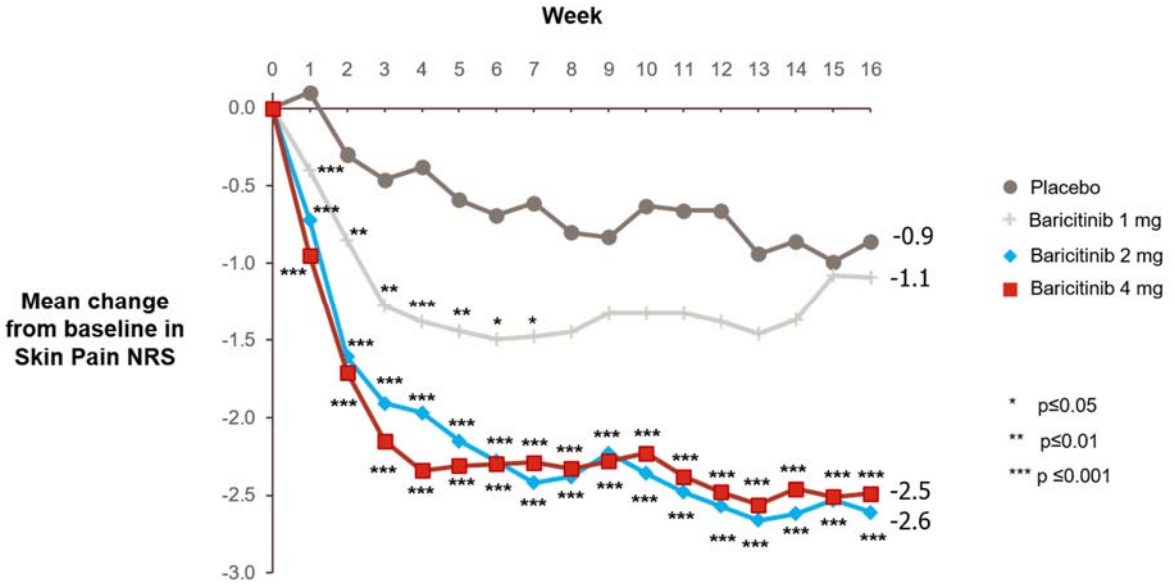
[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; NRS: numeric rating scale; PBO: placebo.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 34: Mean change from baseline in Skin Pain NRS in BREEZE-AD1 (JAHL) patients over trial period



p-values obtained from MMRM models.
Abbreviations: NRS: numeric rating scale.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 35: Mean change from baseline in Skin Pain NRS in BREEZE-AD2 (JAHM) patients over trial period



p-values obtained from MMRM models.
Abbreviations: NRS: numeric rating scale.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Item 2 of ADSS mean change from baseline at Weeks 1 and 16

The average MCFB in ADSS Item 2 at Week 16 in patients in the BREEZE-AD1 and -AD2 trials is summarised in Table 45. In both trials, treatment with 4 mg baricitinib was associated with a statistically significant increase in the MCFB of ADSS Item 2 as compared with placebo at Week 1 (both p<0.001) and at Week 16 (both p<0.01). This improvement versus placebo was statistically significant at p<0.001 as early as Week 1 and significance at this level was maintained through at p<0.01 to Week 15 in both trials (Figure 36 and Figure 37).

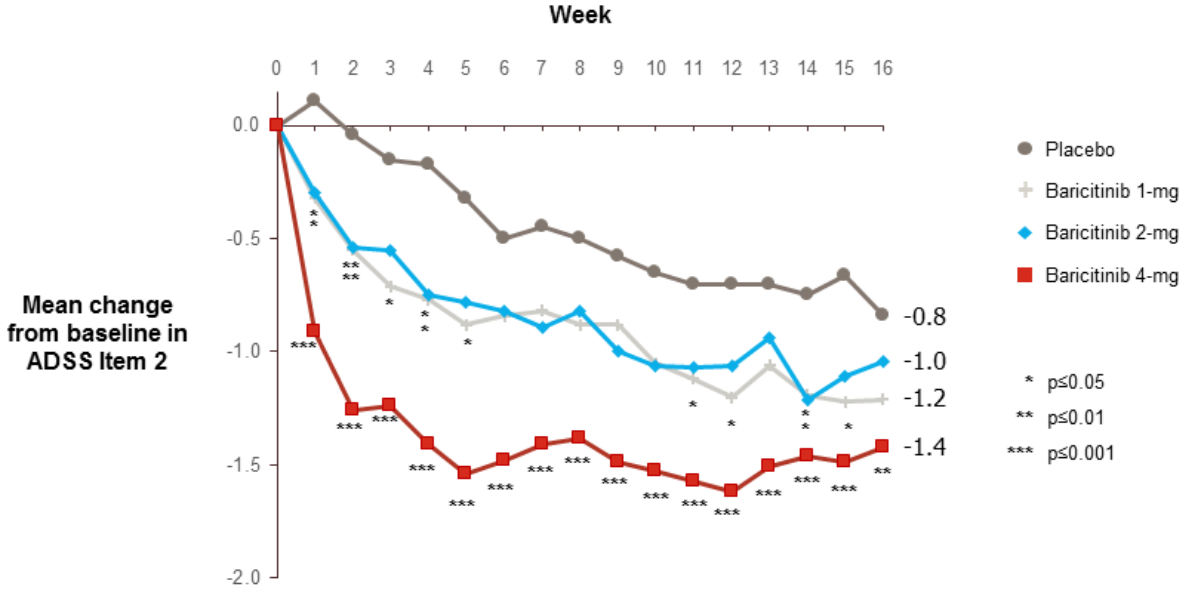
Table 45: Mean change from baseline in Item 2 of ADSS at Week 16 in BREEZE-AD1 and -AD2 patients

Mean change in Item 2 ADSS	Week 1		Week 16	
	PBO	BARI 4 mg	PBO	BARI 4 mg
BREEZE-AD1 (JAHL)[56]	**N=249**	**N=125**	**N=249**	**N=125**
Baseline mean	3.41	3.26	3.41	3.26
MCFB, LSM (95% CI vs. PBO)	0.11 (NA)	−0.91 (−1.38, −0.66)	−0.84 (NA)	−1.42 (−1.00, −0.17)
p-value[a] vs. PBO	NA	<0.001	NA	0.006
BREEZE-AD2 (JAHM)[57]	**N=244**	**N=123**	**N=244**	**N=123**
Baseline mean	1.83	1.91	1.83	1.91
MCFB, LSM (95% CI vs. PBO)	−0.02 (NA)	−0.58 (−0.79, −0.33)	−0.50 (NA)	−1.13 (−0.96, −0.29)
p-value[a] vs. PBO	NA	<0.001	NA	<0.001

[a] p-values obtained from MMRM models.
Abbreviations: ADSS: atopic dermatitis sleep scale; BARI: baricitinib; CI: confidence interval; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo.

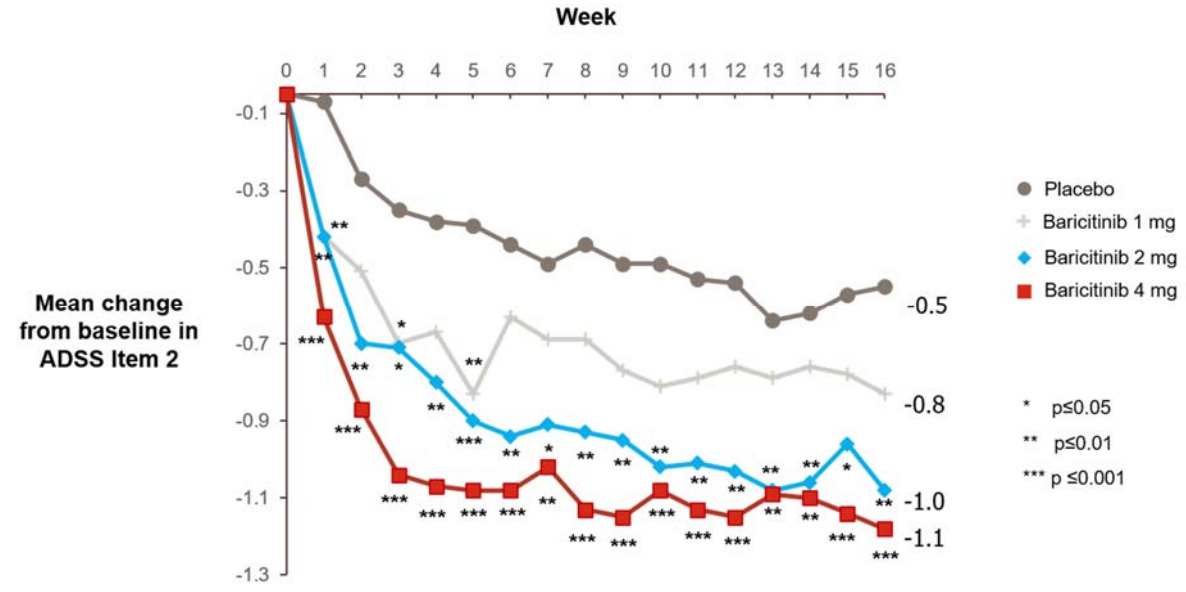
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 36: Mean change from baseline in ADSS Item 2 in BREEZE-AD1 (JAHL) patients over trial period



p-values obtained from MMRM models.
Abbreviations: ADSS: atopic dermatitis sleep scale.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 37: Mean change from baseline in ADSS Item 2 in BREEZE-AD2 (JAHM) patients over trial period



p-values obtained from MMRM models.
Abbreviations: ADSS: atopic dermatitis sleep scale.
Sources: Simpson *et al,* 2020,[49] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Health-related quality of life (HRQoL) endpoints

DLQI score at Week 16

The DLQI outcomes at Week 16 for patients in the BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) studies are summarised in Table 46 and Table 47, respectively.

In both trials, treatment with 4 mg baricitinib was associated with a statistically significant increase in the MCFB in DLQI, a statistically higher proportion of patients achieving a DLQI score of 0 or 1 and a statistically significantly higher proportion of patients achieving a ≥4-point improvement in DLQI score at Week 16 as compared with placebo (all p<[redacted]).

The proportion of patients achieving a ≥4-point improvement in DLQI score across the BREEZE-AD1 and -AD2 trial periods was significantly higher at p<[redacted] in the baricitinib-treated group at Week 1, and was maintained at p<[redacted] until Week 16 (Figure 38 and Figure 39).

Table 46: DLQI outcomes at Week 16 in BREEZE-AD1 (JAHL) patients

DLQI	PBO (N=[redacted])	BARI 4 mg (N=[redacted])
Baseline mean	[redacted]	[redacted]
MCFB		
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]
Score of 0 or 1		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]
≥4-point improvement[c]		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]

[a] p-values obtained from MMRM models. [b] p-values obtained by testing odds ratio within logistic regression framework (not presented). [c] Analyses performed on populations with a baseline score ≥4 (PBO: N=233; BARI 4 mg: N=116).
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology Life Quality Index; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo.
Source: BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Table 47: DLQI outcomes at Week 16 in BREEZE-AD2 (JAHM) patients

DLQI	PBO (N=[redacted])	BARI 4 mg (N=[redacted])
Baseline mean	[redacted]	[redacted]
MCFB		
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]
Score of 0 or 1		
Response, n (%) [95% CI]	[redacted]	[redacted]

Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]
≥4-point improvement[c]		
Response, n (%) [95% CI]	[redacted]	[redacted]
Difference vs. PBO, % (95% CI)	[redacted]	[redacted]
Odds ratio vs. PBO (95% CI)	[redacted]	[redacted]
p-value[b] vs. PBO	[redacted]	[redacted]

[a] p-values obtained from MMRM models. [b] p-values obtained by testing odds ratio within logistic regression framework (not presented). [c] Analyses performed on populations with a baseline score ≥4 (PBO: N=224; BARI 4 mg: N=112).
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology Life Quality Index; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo.
Source: BREEZE-AD2 (JAHM) Clinical Study Report.[57]

Figure 38: Proportion of patients in BREEZE-AD1 (JAHL) achieving a ≥4-point improvement in DLQI score over trial period

p-value obtained by Fisher's exact test.
Abbreviations: DLQI: Dermatology Life Quality Index.
Source: BREEZE-AD1 (JAHL) Clinical Study Report.[56]

Figure 39: Proportion of patients in BREEZE-AD2 (JAHM) achieving a ≥4-point improvement in DLQI score over trial period

p-value obtained by Fisher's exact test.
Abbreviations: DLQI: Dermatology Life Quality Index.
Source: BREEZE-AD2 (JAHM) Clinical Study Report.[57]

EQ-5D-5L at Week 16

For the clinical effectiveness data presented here, the HIS was based directly on the England-only valuation of EQ-5D-5L by Devlin *et al*, 2018.[69] For the economic evaluation, the EQ-5D-5L scores were cross-walked to EQ-5D-3L and valued using the EQ-5D-3L weights using the algorithm by Dolan *et al*, 1997.[70]

The average MCFB in the two components of the EQ-5D-5L for patients in the BREEZE-AD1 and -AD2 trials at Week 16 are summarised in Table 48. In both trials at Week 16, treatment with 4 mg baricitinib was associated with a statistically significant increase in VAS (both p<[redacted]) and the Health Index Score (both p<[redacted]) as compared with placebo. The improvements in VAS and the Health Index Score were maintained in both trials to Week 16 to p<[redacted] and p<[redacted], respectively.

Table 48: Mean change from baseline in EQ-5D-5L at Week 16 in BREEZE-AD1 (JAHL) patients

EQ-5D-5L	VAS Score		Health Index Score (England algorithm)	
	PBO	4 mg BARI	PBO	4 mg BARI
BREEZE-AD1 (JAHL)[56]	N=[redacted]	N=[redacted]	N=[redacted]	N=[redacted]
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]
BREEZE-AD2 (JAHM)[57]	N=[redacted]	N=[redacted]	N=[redacted]	N=[redacted]

Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB, LSM (95% CI vs. PBO)	[redacted]	[redacted]	[redacted]	[redacted]
p-value[a] vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-values obtained from MMRM models.
Abbreviations: BARI: baricitinib; CI: confidence interval; EQ-5D-5L: 5-level EuroQol 5 Dimensions; LSM: least-squares mean; MCFB: mean change from baseline; MMRM: mixed-effect model for repeated measures; NA: not applicable; PBO: placebo; VAS: visual analogue scale.
Source: BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

B.2.7 *Subgroup analysis*

Prespecified subgroup analyses were conducted on the pooled Phase III monotherapy population (BREEZE-AD1 and BREEZE-AD2) and on the combination therapy population (BREEZE-AD7 [JAIY]). Subgroup data were not available from BREEZE-AD4 (JAIN) at the time of submission.

The full list of subgroups tested for interaction in each trial are provided in Section B.2.3.1. Tests of subgroup interactions were executed with regression models, and subgroup by therapy interaction terms were tested at the $\alpha<0.01$ level of significance.

The proportion of patients achieving IGA ≤1, EASI75 or a ≥4-point improvement in Itch NRS at Week 16 for subgroups with significant interactions ($p<0.1$) is presented in Table 49. In the combination therapy patients from BREEZE-AD7 (JAIY), significant interactions ($p<0.05$) were observed at Week 16 for gender in IGA ≤1, gender, region, specific region (Japan versus all others and East Asia versus all others) and ciclosporin failure for EASI75 and specific region (East Asia versus all others) for ≥4-point improvement in Itch NRS. In the pooled monotherapy patients from BREEZE-AD1 and -AD2, a significant interaction ($p<0.05$) was observed at Week 16 for baseline IGA score in EASI75.

Across the combination therapy population and the monotherapy population, many of the statistically significant treatment by subgroup interactions were likely driven by differential responses across subgroups in the placebo and 1 mg baricitinib treatment groups. There was no evidence of a reversal of treatment effect as compared with the ITT population, with 4 mg baricitinib consistently favourable versus placebo across subgroups, suggesting a quantitative rather than qualitative interaction in these subgroups.

Table 49: Proportion of combination therapy patients (BREEZE-AD7 [JAIY]) and monotherapy patients (Pooled BREEZE-AD1 and -AD2) achieving IGA ≤1, EASI75 or a ≥4-point improvement in Itch NRS at Week 16 for subgroups with significant interactions (p<0.1)

Outcome	Subgroup	Category	Response at Week 16 (%)				RR vs PBO	p-value[a]
			PBO	1 mg BARI	2 mg BARI	4 mg BARI		
Combination therapy: BREEZE-AD7 (JAIY) (N=[redacted])								
IGA ≤1	Gender	Male (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
EASI75	Gender	Male (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Baseline IGA score	IGA 3 (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		IGA 4 (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Region	Europe (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Japan (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		ROW (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk (Europe vs Japan)	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk (Europe vs ROW)	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	Europe (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		All other (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	Japan (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		Not Japan (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	East Asia (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		All other (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		Yes (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]

	Ciclosporin failure	No (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
Itch NRS improvement of 4 or more points	Prior systemic therapy	Yes (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		No (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	Europe (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		No (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	East Asia (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		All other (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
	TCI failure or inadvisable	Yes (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
		No (N=[redacted])	[redacted]	NA	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	NA	[redacted]	[redacted]	[redacted]	[redacted]
Pooled monotherapy: BREEZE-AD1 and -AD2 (JAHL and JAHM) (N=[redacted])								
IGA ≤1	TCI failure or inadvisable	Yes (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		No (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EASI75	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Baseline IGA score	IGA 3 (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		IGA 4 (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	Europe (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		All other (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] p-value shows treatment by subgroup interaction value and includes all doses of baricitinib.
Abbreviations: BARI: baricitinib; PBO: placebo; ROW: Rest of World; RR: risk ratio.

B.2.8 *Meta-analysis*

A network meta-analysis (NMA) is a common method used to compare two or more interventions. Dupilumab was the only comparator for which double-blind, parallel, placebo-controlled studies which reported results in a manner comparable to the baricitinib evidence base were identified. Therefore, an indirect treatment comparison (ITC) was performed to synthesise the evidence concerning baricitinib and dupilumab (Section B.2.9).

B.2.9 *Indirect and mixed treatment comparisons*

Summary of indirect treatment comparison

- An ITC was performed to assess the clinical effectiveness of 4 mg baricitinib versus 300 mg dupilumab in adult patients with moderate-to-severe AD who have experienced failure with, are intolerant to, or are contraindicated to ciclosporin, in line with the eligibility criteria for the JAIN trial.
  - In the pooled analysis where JAIN-like JAIY patients and CAFÉ-like CHRONOS patients were included (primary censoring), the results indicated similar efficacy between baricitinib and dupilumab in achieving EASI75 response at Week 16 (RR: [redacted], 95% CI: [redacted]). However, these results indicated that baricitinib was associated with [redacted] (in terms of the RR) compared to dupilumab in achieving EASI50 response with (RR: [redacted]; 95% CI: [redacted]) and without (RR: [redacted]; 95% CI: [redacted]) ≥4-point improvement in DLQI at Week 16.
  - In the pooled analysis of JAIN-like JAIY patients and CAFÉ-like CHRONOS patients using the secondary censoring rule, results were consistent with those of primary censoring. Baricitinib was associated with a [redacted] (in terms of the RR) compared to dupilumab in achieving EASI50 response with a ≥4-point improvement in DLQI at Week 16 (RR: [redacted]; 95% CI: [redacted]). The results indicated similar efficacy between baricitinib and dupilumab in achieving EASI50 (RR: [redacted]; 95% CI: [redacted]) and EASI75 (RR: [redacted]; 95% CI: [redacted]) at Week 16.
  - In the analysis comparing JAIN versus CAFÉ, no statistical difference (in terms of the RR) was observed between baricitinib (in combination with TCS) and dupilumab in achieving EASI50 (RR: [redacted]; 95% CI: [redacted]), EASI75 (RR: [redacted]; 95% CI: [redacted]) and EASI90 (RR: [redacted]; 95% CI: [redacted]) at Week 16, but results [redacted]. Baricitinib also showed [redacted] in achieving itch reduction (RR: [redacted]; 95% CI: [redacted]), but this was not statistically significant.
  - Similar results were observed in the analysis where only European patients from the JAIN trial were included. Whilst there were no statistical differences (in terms of the RR) observed between baricitinib and dupilumab in achieving EASI50 (RR: [redacted]; 95% CI: [redacted]), EASI75 (RR: [redacted]; 95% CI: [redacted]), EASI90 (RR: [redacted]; 95% CI: [redacted]) and itch reduction (RR: [redacted]; 95% CI: [redacted]) at Week 16, all comparisons [redacted] in this analysis.
  - Results of the scenario analysis considering JAIN-like JAHL/JAHM patients and CAFÉ-like SOLO1/SOLO2 patients indicated that baricitinib monotherapy showed similar efficacy to dupilumab in achieving EASI50 response (RR: [redacted]; 95% CI: [redacted]), EASI75 response (RR: [redacted]; 95% CI: [redacted]) and EASI50 with a ≥4-point DLQI improvement (RR: [redacted]; 95% CI: [redacted]) at Week 16, with no statistically significant differences (in terms of the RR) observed.
- In summary, differences between baricitinib (4 mg QD) and dupilumab (300 mg Q2W) were often not statistically significant, with the confidence intervals of the calculated ORs and RRs spanning a value of 1. In the analysis of adult patients with moderate-to-severe AD who have experienced failure with, are intolerant to, or are contraindicated to ciclosporin (JAIN + JAIN-like JAIY), results [redacted] in terms of skin clearance as assessed by EASI score, but in the analysis of JAIN-only patients, results [redacted] in terms of itch improvement.

B.2.9.1 Study identification

As discussed in Section B.2.1, an SLR was conducted to identify all relevant clinical evidence on the efficacy and safety of baricitinib and potential comparators for the treatment of adults with moderate-to-severe AD. 62 publications were ultimately included in the SLR, reporting on 40 unique studies. Full details of the methodology and results of the SLR are presented in Appendix D.

The SLR was designed to capture evidence for a broader patient population than the population of relevance for this submission and included a broader range of potential comparators than are relevant in UK clinical practice. The population of relevance for this submission is patients with moderate-to-severe AD who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. As discussed in Section B.1.3.3, the relevant comparators in this setting are limited to dupilumab and BSC. This is in line with TA534, where it was considered appropriate to include a comparison to BSC only in the 5th-line setting. Of the 36 studies included in the SLR, 3 published studies investigated the use of baricitinib and 12 investigated the use of dupilumab in patients with moderate-to-severe AD (see Table 50).

Table 50: Clinical effectiveness evidence for baricitinib and dupilumab for patients with moderate-to-severe AD

Study ID	Trial No.	Full reference
Baricitinib		
Guttman-Yassky 2019b[48]	NCT02576938	Guttman-Yassky E, Silverberg JI, Nemoto O, et al. Baricitinib in adult patients with moderate-to-severe atopic dermatitis: A phase 2 parallel, double-blinded, randomized placebo-controlled multiple-dose study. J Am Acad Dermatol 2019;80:913-921.e9.
BREEZE-AD1 (JAHL)[49]	NCT03334396	Simpson EL, Lacour JP, Spelman L, et al. Baricitinib in patients with moderate-to-severe atopic dermatitis and inadequate response to topical corticosteroids: results from two randomized monotherapy phase III trials. Br J Dermatol 2020.
BREEZE-AD2 (JAHM)[49]	NCT03334422	
Dupilumab		
C4[50]	NCT01639040	Beck LA, Thaçi D, Hamilton JD, et al. Dupilumab treatment in adults with moderate-to-severe atopic dermatitis. New England Journal of Medicine 2014;371:130-139.
Guttman-Yassky 2019a[51]	NCT01979016	Guttman-Yassky E, Bissonnette R, Ungar B, et al. Dupilumab progressively improves systemic and cutaneous abnormalities in patients with atopic dermatitis. Journal of Allergy and Clinical Immunology 2019;143:155-172.
LIBERTY AD CAFÉ[71]	NCT02755649	de Bruin-Weller M, Thaci D, Smith CH, et al. Dupilumab with concomitant topical corticosteroid treatment in adults with atopic dermatitis with an inadequate response or intolerance to ciclosporin A or when this treatment is medically inadvisable: a placebo-controlled, randomized phase III clinical trial (LIBERTY AD CAFE). Br J Dermatol 2018;178:1083-1101.
LIBERTY AD CHRONOS[72]	NCT02260986	Blauvelt A, de Bruin-Weller M, Gooderham M, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corticosteroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. The Lancet 2017;389:2287-2303.

LIBERTY AD EVALUATE[52]	NCT02210780	Blauvelt A, Simpson EL, Tyring SK, et al. Dupilumab does not affect correlates of vaccine-induced immunity: a randomized, placebo-controlled trial in adults with moderate-to-severe atopic dermatitis. Journal of the American Academy of Dermatology 2019;80:158-167. e1.
LIBERTY AD SOLO-CONTINUE[53]	NCT02395133	Worm M, Simpson EL, Thaçi D, et al. Efficacy and safety of multiple dupilumab dose regimens after initial successful treatment in patients with atopic dermatitis: a randomized clinical trial. JAMA dermatology 2020;156:131-143.
M12[50]	NCT01548404	Beck LA, Thaçi D, Hamilton JD, et al. Dupilumab treatment in adults with moderate-to-severe atopic dermatitis. New England Journal of Medicine 2014;371:130-139.
M4A[50]	NCT01259323	
M4B[50]	NCT01385657	
SOLO 1[73]	NCT02277743	Simpson EL, Bieber T, Guttman-Yassky E, et al. Two Phase 3 Trials of Dupilumab versus Placebo in Atopic Dermatitis. New England Journal of Medicine 2016;375:2335-2348.
SOLO 2[73]	NCT02277769	
Thaci 2016[74]	NCT01859988	Thaçi D, Simpson EL, Beck LA, et al. Efficacy and safety of dupilumab in adults with moderate-to-severe atopic dermatitis inadequately controlled by topical treatments: a randomised, placebo-controlled, dose-ranging phase 2b trial. The Lancet 2016;387:40-52.

Abbreviations: AD: atopic dermatitis.

Direct evidence for the relative efficacy of baricitinib versus BSC is provided by the placebo-controlled BREEZE-AD trials (placebo can be considered a proxy for BSC). However, no head-to-head clinical trials comparing baricitinib versus dupilumab were identified. Therefore, in order to estimate the comparative effectiveness of baricitinib versus dupilumab, the evidence identified in the SLR for dupilumab was reviewed for the purposes of conducting an ITC. The ITC included analysis of broader patient populations than the population of relevance for this submission. The characteristics of studies included in the analysis of relevance for the submission are presented in Table 51.

Table 51: Characteristics of studies included in the ITC analysis of relevance for the submission

	Inclusion criteria	Exclusion criteria
Patient population	• As per the eligibility criteria for the SLR (Appendix D) • Includes available data for patients who have experienced failure with, or are intolerant to or have a contraindication to, ciclosporin	• As per the eligibility criteria for the SLR (Appendix D)
Interventions/ Comparators	• Includes at least one trial arm reporting on the licensed dose of baricitinib (4 mg QD) or dupilumab (300 mg Q2W)	• Does not include at least one trial arm reporting on the licensed dose of baricitinib (4 mg QD) or dupilumab (300 mg Q2W)
Outcome	• As per the eligibility criteria for the SLR (Appendix D)	• As per the eligibility criteria for the SLR (Appendix D)
Study design	• All randomised, controlled trials for moderate-to-severe AD identified in the SLR	• Pilot study/ Phase I studies/Phase IIa studies • Non-comparative studies with no active comparator arm or no

		placebo arm • Any clinical trial program that may have been interrupted

Abbreviations: AD: atopic dermatitis; QD: daily; Q2W: every two weeks; SLR: systematic literature review.

Table 52 summarises the studies included in and excluded from the ITC analysis of interest. A total of 8 studies were included in the ITC: BREEZE-AD1 (JAHL), -AD2 (JAHM), -AD4 (JAIN) and -AD7 (JAIY) investigating baricitinib and LIBERTY AD CAFÉ, LIBERTY AD CHRONOS and SOLO1 and SOLO2 investigating dupilumab. Individual patient data (IPD) were available for the BREEZE-AD trials, and thus despite the broader eligibility criteria for the BREEZE-AD1, -AD2 and -AD7 trials, data could be extracted for the subgroups of patients who met the eligibility criteria for the ITC. IPD were not available for the dupilumab trials. However, data were available for the relevant population from post-hoc pooled analyses presented in TA534.[1]

Table 52: Summary of studies included and excluded from the ITC

Trial name	Patient population	Subgroup data available for the relevant patient population	Interventions	Comparator	Phase	Included (Yes/No)	Reason for exclusion
Guttman-Yassky 2019b	Moderate to severe AD (background TCS)	No	Baricitinib: • 4 mg QD (n=38) • 2 mg QD (n=37)	Placebo (n=49)	II	No	Not a relevant population
BREEZE-AD1 (JAHL)	Moderate to severe AD	Yes	Baricitinib: • 4 mg QD (n=125) • 2 mg QD (n=123) • 1 mg QD (n=127)	Placebo (n=249)	III	Yes	-
BREEZE-AD2 (JAHM)	Moderate to severe AD	Yes	Baricitinib: • 4 mg QD (n= 123) • 2 mg QD (n= 123) • 1 mg QD (n= 125)	Placebo (n=244)	III	Yes	-
BREEZE-AD4 (JAIN)[a]	Moderate to severe adult patients with AD who have experienced failure with, are intolerant to, or have contraindication to, ciclosporin	NA	Baricitinib: • 4 mg QD+ TCS (n=92) • 2 mg QD+ TCS (n=185) 1 mg QD+ TCS (n=93)	Placebo + TCS (n=93)	III	Yes	-
BREEZE-AD7 (JAIY)[a]	Moderate to severe AD	Yes	Baricitinib: • 4 mg QD + TCS (n=111) • 2 mg QD + TCS (n= 109)	Placebo + TCS (n=109)	III	Yes	-
LIBERTY AD CAFÉ	Moderate to severe AD, ciclosporin inadvisable	NA	Dupilumab: • 300 mg Q2W + TCS (n=107) • 300 mg QW + TCS (n=110)	Placebo + TCS (n=108)	III	Yes	-

LIBERTY AD CHRONOS	Moderate to severe AD	Yes – TA534: Post-hoc subgroup of patients who cannot take ciclosporin or who did not adequately respond to ciclosporin (n=137), pooled with CAFÉ (n=325)	Dupilumab: • 300 mg Q2W + TCS (n=106) • 300 mg QW + TCS (n=319)	Placebo + TCS (n=315)	III	Yes	-
SOLO 1	Moderate to severe AD	Yes – TA534: Post-hoc subgroup of patients who previously used systemics (commonly ciclosporin) (n=288)	Dupilumab: • 300 mg Q2W (n=224) • 300 mg QW (n=223)	Placebo (n=224)	III	Yes	-
SOLO 2	Moderate to severe AD		Dupilumab: • 300 mg Q2W (n=233) • 300 mg QW (n=239)	Placebo (n=236)	III	Yes	-
Thaci 2016 (AD-1021)	Moderate to severe AD	No	Dupilumab: • 300 mg Q2W (n=64) • 300 mg QW (n=63) • 200 mg Q2W (n=62) • 300 mg Q4W (n=65) • 100 mg Q4W (n=65)	Placebo [QW] (n=61)	IIb	No	• Not a relevant population
C4	Moderate to severe AD	No	Dupilumab: • 300 mg QW +TCS (n=21)	Placebo +TCS (n=10)	IIa	No	• Not a relevant study design or population • Did not include approved dupilumab dose
M12	Moderate to severe AD	No	Dupilumab: • 300 mg QW (n=55)	Placebo QW (n=54)	IIa	No	

M4A	Moderate to severe AD	No	Dupilumab: • 75 mg QW (n=8) • 150mg QW (n=8) • 300 mg QW (n=8)	Placebo QW (n=6)	I	No	
M4B	Moderate to severe AD	No	Dupilumab: • 150 mg QW (n=14) • 300 mg QW (n=13)	Placebo QW (n=10)	I	No	
LIBERTY AD EVALUATE	Moderate to severe AD	No	Dupilumab: • 300 mg QW (n= 97)	Placebo (n=97)	II	No	• Not a relevant study design or population • Did not include approved dupilumab dose
Guttman-Yassky 2019a	Moderate to severe AD	No	Dupilumab: • 200 mg QW (n=27)	Placebo (n=27)	II	No	• Not a relevant study design or population • Did not include approved dupilumab dose
LIBERTY AD SOLO-CONTINUE	Moderate to severe AD	No	Dupilumab: • 300 mg Q8W (n=84) • 300 mg Q4W (n=86) • 300 mg QW or Q2W (n=169)	Placebo (n=83)	III	No	Results were not available

Grey text indicates that that population, dose or trial design does not match the eligibility criteria for the ITC.
[a]Whilst not identified in the SLR, BREEZE-AD4 (JAIN) and BREEZE-AD7 (JAIY) were included in the ITC.
Abbreviations: AD: atopic dermatitis; QD: once daily; QW: once weekly. Q2W: once every 2 weeks. Q4W: once every 4 weeks; TCS: Topical corticosteroids.

B.2.9.2 Feasibility assessment

Study design and outcomes

A comparison of the study designs of the trials considered in the ITC is presented in Table 54. All trials considered in the ITC were Phase III, double-blind, randomised, placebo-controlled trials, and all trials collected efficacy endpoints at Week 16. All trials were international, however CAFÉ only included European patients.

Washout periods for topical treatments prior to randomisation were longer in the BREEZE-AD trials than CHRONOS and the SOLO1/2 trials, and CAFÉ utilised a wash-in period (during the initial 2 weeks of the screening period, patients could use TCS at investigator discretion).

Table 53: Comparison of study design for studies considered in the ITC

Study ID	Study design	Region	Intervention/ comparator	Washout period[a]	Duration (weeks)	Timepoints for efficacy assessment (weeks)
JAHL	Double-blind, randomised, placebo-controlled, Phase III trial	International	Baricitinib monotherapy vs PBO	2 weeks	16	16
JAHM		International			16	16
JAIN		International	Baricitinib +TCS vs PBO +TCS		52	16, 24, 52
JAIY		International			16	16
CAFÉ		International (Europe only)	Dupilumab +TCS vs PBO +TCS	2 weeks (wash-in)	16	16
CHRONOS		International		1 week	52	16
SOLO 1		International	Dupilumab monotherapy vs PBO	1 week	16	16
SOLO 2		International			16	16

[a] Washout period for topical AD treatments.
Abbreviations: ITC: indirect treatment comparison; PBO: placebo; TCS: topical corticosteroids.

The patient populations included in the BREEZE-AD1, -AD2, -AD7, CHRONOS, SOLO1 and SOLO2 trials were broader than the population of relevance for the ITC (patients with moderate-to-severe AD versus patients with moderate-to-severe AD who have experienced failure with, are intolerant to or have a contraindication to, ciclosporin). In order to investigate the feasibility of indirect comparisons in the relevant population, data were extracted from the BREEZE-AD1, -AD2, and -AD7 trials for the "JAIN-like" subgroups of patients who had a history of intolerance or inadequate response to ciclosporin. To maximise sample sizes, data were pooled for baricitinib monotherapy (JAHL + JAHM JAIN-like) and baricitinib +TCS (JAIN + JAIY JAIN-like). Data were available for dupilumab for the relevant population from post-hoc pooled analyses presented in TA534 (CAFÉ + CHRONOS CAFÉ-like and SOLO 1/2 CAFÉ-like).[1] These populations have been considered further in the feasibility assessment.

A comparison of the outcomes included in the relevant populations for the ITC is presented in Table 54. Evidence for the composite response endpoint of EASI50 and ≥4-point improvement in DLQI was available from TA534 for the post-hoc subgroups of CAFÉ plus CAFÉ-like patients from CHRONOS and CAFÉ-like patients from SOLO1 and SOLO2, but not for the CAFÉ trial. Evidence for EASI75 and EASI50 was available from all relevant trial populations. EASI90 data were also available from the CAFÉ trial.

Table 54: Comparison of efficacy outcomes available for the ITC

Outcomes	Baricitinib			Dupilumab		
	BREEZE-AD4 (JAIN)	JAIN + JAIY JAIN-like pooled	JAHL/JAHM JAIN-like pooled	CAFÉ	CAFÉ + CHRONOS CAFÉ-like pooled	SOLO 1/2 CAFÉ-like pooled
EASI50 and ≥4-point DLQI improvement	✓	✓	✓	×	✓	✓
EASI50	✓	✓	✓	✓	✓	✓
EASI75	✓	✓	✓	✓	✓	✓
EASI90	✓	✓	✓	✓	×	×
Itch NRS ≥4-point Improvement	✓	✓	✓	✓	×	×
SCORAD Total Score PCFB	✓	✓	✓	×	×	×
SCORAD Sleep Score PCFB	✓	✓	✓	×	×	×
DLQI Score MCFB (ACFB)	✓	✓	✓	×	×	×
EQ5D (absolute CFB)	✓	✓	✓	×	×	×
POEM (absolute CFB)	✓	✓	✓	×	×	×
HADS anxiety score (ACFB)	✓	✓	✓	×	×	×
HADS depression score (ACFB) value	✓	✓	✓	×	×	×
EASI score (PCFB)	✓	✓	✓	×	×	×
Itch NRS score (ACFB)	✓	✓	✓	×	×	×
Itch NRS score (PCFB)	✓	✓	✓	×	×	×
BSA at Week 16 (ACFB)	✓	✓	✓	×	×	×

IGA data for dupilumab have not been extracted and a comparison based on IGA has not been conducted since the IGA outcome for the dupilumab clinical trials programme does not match the scale used in the baricitinib trials. A tick (✓) denotes that the outcome was reported at Week 16, and a cross (×) denotes the outcome was not reported.
Abbreviations: ACFB: absolute change from baseline; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; IGA: Investigator's Global Assessment; ITC: indirect treatment comparison; NRS: numeric rating scale; PCFB: percentage change from baseline; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.

Based on the populations and outcomes included in the relevant clinical trials for baricitinib and dupilumab, a number of comparisons were considered for inclusion in the ITC (Table 55). For combination therapies, in addition to the direct comparison of JAIN versus CAFÉ, a comparison was considered using pooled data for patients from JAIN and the post-hoc subgroup of JAIN-like patients from JAIY versus the pooled CAFÉ + CAFÉ-like CHRONOS data from TA534. For monotherapies, a comparison was considered using pooled data for the post-hoc subgroups of JAIN-like patients from JAHL and JAHM versus the pooled CAFÉ-like SOLO1 and SOLO2 data from TA534.

Table 55: Summary of analyses considered for the ITC

Comparison	Populations		Outcomes (Week 16)
	Baricitinib	Dupilumab	
Baricitinib + TCS versus dupilumab + TCS	**JAIN** All trial data	**CAFÉ** All trial data	• EASI50 • EASI75 • EASI90 • Itch NRS ≥4-point Improvement
	JAIN + JAIN-like JAIY JAIN trial data combined with *post hoc* data from the subgroup of patients with ciclosporin failure, intolerance or contradiction from the JAIY study	**CAFÉ + CAFÉ-like CHRONOS** CAFÉ trial data combined with *post hoc* data from the subgroup of patients with ciclosporin failure, intolerance or contradiction from the CHRONOS study	• EASI50 + DLQI ≥4-point Improvement • EASI50 • EASI75
Baricitinib monotherapy versus dupilumab monotherapy	**JAIN-like JAHL + JAHM** Pooled *post hoc* data from the subgroup of patients with ciclosporin failure, intolerance or contradiction from the JAHL and JAHM studies	**CAFÉ-like SOLO1 and SOLO2** Pooled data for the sub-population of patients with ciclosporin failure, intolerance or contradiction from the SOLO1 and SOLO2 studies	• EASI50 + DLQI ≥4-point Improvement • EASI50 • EASI75

Abbreviations: DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; ITC: indirect treatment comparison.

Patient population

A comparison of the baseline characteristics of the relevant populations considered for the ITC is presented in Table 56, to identify any heterogeneity that could influence relative treatment or baseline treatment effects. Populations were similar in terms of sex, age, and baseline scores, with the exception of HADS. BREEZE-AD4 included a considerably higher proportion of Asian patients compared with CAFÉ (17% and 20% in the baricitinib and placebo arms of BREEZE-AD4, respectively, versus 2% in both the dupilumab and placebo arms of CAFÉ). As discussed in Section B.2.7, in the BREEZE-AD7 trial, significant interactions ($p<0.05$) were observed at Week 16 for specific region (Japan versus all others and East Asia versus all others) for EASI75 and specific region (East Asia versus all others) for ≥4-point improvement in Itch NRS, indicating that geographic region might be a treatment effect modifier. As such, a sensitivity analysis was conducted for the comparison of JAIN versus CAFÉ where only European patients from JAIN were included.

Table 56: Comparison of baseline characteristics for the populations considered in the ITC

	Baricitinib						Dupilumab					
	BREEZE-AD4 (JAIN)		JAIN + JAIY JAIN-like pooled		JAHL/JAHM JAIN-like pooled		CAFÉ		CAFÉ + CHRONOS CAFÉ-like pooled		SOLO 1/2 CAFÉ-like pooled	
Intervention	PBO +TCS	BARI 4 mg QD +TCS	PBO +TCS	BARI 4 mg QD +TCS	PBO	BARI 4 mg QD	PBO +TCS	DUPI 300 mg Q2W +TCS	PBO +TCS	DUPI 300 mg Q2W +TCS	PBO	DUPI 300 mg Q2W
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	108	107	169	130	88	104
Males, %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	63	61	60	59	63	72
Race, n (%)												
White	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	104 (96.3)	104 (97.2)	152 (89.9)	121 (93.1)	52 (59.1)	75 (72.1)
Black	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0 (0)	0 (0)	3 (1.8)	1 (0.8)	0 (0)	1 (1)
Asian	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	2 (1.9)	2 (1.9)	12 (7.1)	7 (5.4)	30 (34.1)	23 (22.1)
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	38.9 (13.3)	37.5 (12.9)	38.1 (13)	37.8 (12.9)	38.8 (12.9)	38 (13.5)
Baseline scores												
EASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	32.9 (10.8)	33.3 (9.3)	34.8 (12)	33.6 (10.5)	35.6 (14.3)	36.9 (14.6)
SCORAD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	67 (12.2)	68.6 (11.9)	68.7 (12.8)	69.3 (12.9)	72.8 (13.4)	72.2 (13.9)
IGA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3.5 (0.5)	3.5 (0.5)	3.5 (0.5)	3.5 (0.5)	3.6 (0.5)	3.7 (0.5)
DLQI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	13.2 (7.6)	14.5 (7.6)	14.8 (7.7)	14.6 (7.5)	16.6 (7.9)	15.7 (6.8)
Itch NRS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-	-	-	-
Pruritis NRS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	6.4 (2.23)	6.6 (2.3)	6.9 (2.1)	6.9 (2.1)	7.8 (1.5)	7.6 (1.6)
BSA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55 (20.51)	56.1 (17.83)	58.9 (21.7)	57.3 (18.5)	59.9 (23.7)	58.8 (21.9)
POEM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19.1 (5.9)	19.3 (6.2)	19.9 (6)	19.8 (6.1)	21.9 (5.6)	22 (5.4)
HADS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	13 (7.85)	12.8 (8.01)	13.2 (8.1)	12.8 (7.9)	14.8 (8.8	13.3(7.7)
EQ-5D VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-	-	-	-

Baseline characteristics have only been reported for the licensed doses of baricitinib (4 mg QD) or dupilumab (300 mg Q2W) and placebo.
Abbreviations: EASI: Eczema Area and Severity Index; SCORAD: Scoring Atopic Dermatitis; IGA: Investigator's global assessment; DLQI: Dermatology Life Quality Index; NRS: Itch Numeric Rating Scale; BSA: Body Surface Area; POEM: Patient Oriented Eczema Measure; HADS: Hospital Anxiety and Depression Scale; EQ-5D: European Quality of Life-5 Dimensions; TCS: Topical corticosteroids; qw: once Weekly. q2w: once every 2 Weeks. q4w: once every 4 Weeks; DB: double blind

B.2.9.3 Methodology

An ITC was performed using the Cheetah-tool (Indirect Comparison on results from 2 Meta-Analyses version 1.1), a program developed by Eli Lilly based on R package 'meta'.[75] Fixed effects (FE) models were used to obtain the pooled estimator of the treatment effect for all analyses, given that no between-study heterogeneity (p>0.2) was identified. The full methodology of the ITC is presented in Appendix D.

The ITC analysis of interest for the submission was aligned with the eligibility criteria for the JAIN trial: adult patients with moderate-to-severe AD who have experienced failure with, are intolerant to, or are contraindicated to ciclosporin. This is broadly in line with the population of interest for this submission. The common comparator in all analyses was placebo. Given the lack of data for the composite endpoint of EASI50 + ≥4-point improvement in DLQI for the comparison of JAIN versus CAFÉ, the comparison of pooled JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients was used in the base case analysis of the model, and has thus been presented first here.

The results of the ITC are presented in the following sections:

- Section B.2.9.4: combination therapy results: pooled JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients
- Section B.2.9.5: combination therapy results: JAIN versus CAFÉ
    - Sensitivity analysis: combination therapy results: JAIN (European patients only) versus CAFÉ
- Section B.2.9.6: monotherapy: post-hoc pooled JAIN-like JAHL/JAHM versus CAFÉ-like SOLO1/SOLO2 patients

The results presented here are based on data where the primary censoring rule was applied, in line with the clinical data presented in Section B.2.6 and the data that informs the economic model. All analyses were carried out for the 4 mg dose of baricitinib. Binary endpoints were assessed in the ITC, including the proportion of patients achieving EASI50, EASI75, EASI90, and patients achieving a ≥4-point improvement in itch NRS and a ≥4-point improvement in DLQI at Week 16, as shown in Table 55. Odds ratios (ORs), risk ratios (RRs) and risk difference (RDs) were estimated.

B.2.9.4 Combination therapy results: pooled JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients

EASI50 and DLQI ≥4-point improvement

The pairwise results for proportion of patients achieving EASI50 and ≥4-point improvement in DLQI at Week 16 for JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients are presented in Table 57. Significant differences were observed in favour of dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted], p[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted], p[redacted]) for the comparison between 4 mg baricitinib and 300 mg dupilumab (primary censoring). Results for secondary censoring were consistent with those of primary censoring.

Table 57: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI50 and a ≥4-point improvement in DLQI at Week 16 in JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
Primary censoring				
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]
Secondary censoring				
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

* indicates statistical difference favouring dupilumab. All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology life Quality Index; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI50

The pairwise results for EASI50 at Week 16 for JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients are presented in Table 58. [redacted] in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) for the comparison between 4 mg baricitinib and 300 mg dupilumab (primary censoring). Analyses using the secondary censoring data aligned with the primary censoring data in finding statistically significant differences in favour of dupilumab in terms of OR versus 4 mg baricitinib (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]), but no significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 58: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI50 at Week 16 in JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
Primary censoring				
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]
Secondary censoring				
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

* indicates statistical difference favouring dupilumab. All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI75

The pairwise results for EASI75 at Week 16 for JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients are presented in Table 59. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) in the primary censoring analysis, and the secondary censoring analysis was consistent with this.

Table 59: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI75 at Week 16 in JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
Primary censoring				
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]

CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]
Secondary censoring				
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

B.2.9.5 Combination therapy results: JAIN versus CAFÉ

EASI50

The pairwise results for EASI50 at Week 16 in the full trial populations of JAIN and CAFÉ are presented in Table 60. A [redacted] was observed in terms of the OR between 4 mg baricitinib and 300 mg dupilumab (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]). No significant difference in terms of the RR was observed (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 60: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI50 at Week 16 in all trial patients of the JAIN and CAFÉ trials, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (All trial data)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

* indicates statistical difference favouring dupilumab. All p-values were derived from fixed-effects model.

Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI75

The pairwise results for EASI75 at Week 16 in the full trial populations of JAIN and CAFÉ are presented in Table 61. 4 mg baricitinib showed similar odds of patients achieving EASI75 at Week 16 as 300 mg dupilumab (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]), and no significant difference in terms of the RR was observed (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 61: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI75 at Week 16 in all trial patients of the JAIN and CAFÉ trials, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (All trial data)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI90

The pairwise results for EASI90 at Week 16 in the full trial populations of JAIN and CAFÉ are presented in Table 62. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 62: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI90 at Week 16 in all trial patients of the JAIN and CAFÉ trials, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (All trial data)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]

Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

Itch NRS ≥4-point improvement

The pairwise results for the proportion of patients achieving ≥4-point improvement in itch NRS at Week 16 in the full trial populations of JAIN and CAFÉ are presented in Table 63. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 63: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in ≥4-point improvement in Itch NRS at Week 16 in all trial patients of the JAIN and CAFÉ trials, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (All trial data)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; NRS: numerical rating scale; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

Sensitivity analysis: JAIN (European patients only) versus CAFÉ

EASI50

The pairwise results for EASI50 at Week 16 for European patients from the JAIN trial and all patients from the CAFÉ trial are presented in Table 64. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 64: Relative treatment effect of pairwise comparisons expressed as OR and RR k (with 95% CI) in EASI50 at Week 16 in European patients of the JAIN trial and all patients in the CAFÉ trial, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN	4 mg BARI qd + TCS vs	[redacted]	[redacted]	[redacted]

(European patients only)	PBO + TCS	[redacted]	[redacted]	
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI75

The pairwise results for EASI75 at Week 16 for European patients from the JAIN trial and all patients from the CAFÉ trial are presented in Table 65. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 65: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI75 at Week 16 in European patients of the JAIN trial and all patients in the CAFÉ trial, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (European patients only)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI90

The pairwise results for EASI90 at Week 16 for European patients from the JAIN trial and all patients from the CAFÉ trial are presented in Table 66. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 66: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI90 at Week 16 in European patients of the JAIN trial and all patients in the CAFÉ trial, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (European patients only)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

Itch NRS ≥4-point improvement

The pairwise results for the proportion of patients achieving ≥4-point improvement in itch NRS at Week 16 at Week 16 for European patients from the JAIN trial and all patients from the CAFÉ trial are presented in Table 67. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 67: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in itch NRS ≥4-point improvement at Week 16 in European patients of the JAIN trial and all patients in the CAFÉ trial, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN (European patients only)	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ (All trial data)	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; NRS: numerical rating scale; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

B.2.9.6 Monotherapy scenario results: post-hoc pooled JAIN-like JAHL/JAHM versus CAFÉ-like SOLO1/SOLO2 patients

EASI50 with a ≥4-point improvement in DLQI

The pairwise results for proportion of patients achieving EASI50 and ≥4-point improvement in DLQI at Week 16 for JAIN-like JAHL and JAIN-like JAHM patients and CAFÉ-like SOLO1 and CAFÉ-like SOLO2 patients are presented in Table 68. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted], p=[redacted]).

Table 68: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI50 and a ≥4-point improvement in DLQI at Week 16 in JAIN-like JAHL and JAIN-like JAHM patients and CAFÉ-like SOLO1 and CAFÉ-like SOLO2 patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN-like JAHL and JAIN-like JAHM	4 mg BARI qd vs PBO	[redacted]	[redacted]	[redacted]
CAFÉ-like SOLO1 and SOLO2	300 mg Dupi q2w vs PBO	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI vs 300 mg Dupi	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; DLQI: Dermatology life Quality Index; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI50

The pairwise results for EASI50 at Week 16 for JAIN-like JAHL and JAIN-like JAHM patients and CAFÉ-like SOLO1 and CAFÉ-like SOLO2 patients are presented in Table 69. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted], p=[redacted]).

Table 69: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI50 at Week 16 in JAIN-like JAHL and JAIN-like JAHM patients and CAFÉ-like SOLO1 and CAFÉ-like SOLO2 patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN-like JAHL and JAIN-like JAHM	4 mg BARI qd vs PBO	[redacted]	[redacted]	[redacted]

CAFÉ-like SOLO1 and SOLO2	300 mg Dupi q2w vs PBO	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI vs 300 mg Dupi	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

EASI75

The pairwise results for EASI75 at Week 16 for JAIN-like JAHL and JAIN-like JAHM patients and CAFÉ-like SOLO1 and CAFÉ-like SOLO2 patients are presented in Table 69. No significant difference was observed between 4 mg baricitinib and 300 mg dupilumab in terms of the OR (OR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]) and RR (RR: [redacted] [95% CI: [redacted], [redacted]], p=[redacted]).

Table 70: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI75 at Week 16 in JAIN-like JAHL and JAIN-like JAHM patients and CAFÉ-like SOLO1 and CAFÉ-like SOLO2 patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN-like JAHL and JAIN-like JAHM	4 mg BARI qd vs PBO	[redacted]	[redacted]	[redacted]
CAFÉ-like SOLO1 and SOLO2	300 mg Dupi q2w vs PBO	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI vs 300 mg Dupi	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

B.2.9.7 Conclusions of the indirect treatment comparison

The results of the ITC indicate that baricitinib (4 mg QD) has similar efficacy to dupilumab (300 mg Q2W) in patients who have previously failed ciclosporin due to intolerance, contraindication or inadequate disease control. In the majority of analyses, the [redacted], but differences were often not statistically significant, with the confidence intervals of the calculated ORs and RRs spanning a value of 1.

In the pooled analysis where JAIN-like JAIY patients and CAFÉ-like CHRONOS patients were included, the results indicated similar efficacy between baricitinib and dupilumab in achieving EASI75 response at Week 16 (RR: [redacted], 95% CI: [redacted]). However, these results indicated that baricitinib was associated with [redacted] (in terms of the

RR) compared to dupilumab in achieving EASI50 response with (RR: [redacted]; 95% CI: [redacted]) and without (RR: [redacted]; 95% CI: [redacted]) ≥4-point improvement in DLQI at Week 16. In the analysis of this patient population using the secondary censoring rule, results were consistent with those of primary censoring. Baricitinib was associated with a [redacted] (in terms of the RR) compared to dupilumab in achieving EASI50 response with a ≥4-point improvement in DLQI at Week 16 (RR: [redacted]; 95% CI: [redacted]). The results indicated similar efficacy between baricitinib and dupilumab in achieving EASI50 (RR: [redacted]; 95% CI: [redacted]) and EASI75 (RR: [redacted]; 95% CI: [redacted]) at Week 16.

In the analysis comparing JAIN versus CAFÉ, no statistical difference (in terms of the RR) was observed between baricitinib (in combination with TCS) and dupilumab in achieving EASI50, (RR: [redacted]; 95% CI: [redacted]), EASI75 (RR: [redacted]; 95% CI: [redacted]) and EASI90 (RR: [redacted]; 95% CI: [redacted]), but results [redacted]. Baricitinib also showed [redacted] compared to dupilumab in achieving itch reduction (RR: [redacted]; 95% CI: [redacted]), but this was not statistically significant. Similar results were observed in the analysis where only European patients from the JAIN trial were included: whilst there were no statistical differences (in terms of the RR) observed between baricitinib and dupilumab in achieving EASI50 (RR: [redacted]; 95% CI: [redacted]), EASI75 (RR: [redacted]; 95% CI: [redacted]), EASI90 (RR: [redacted]; 95% CI: [redacted]) and itch reduction at Week 16 (RR: [redacted]; 95% CI: [redacted]), all comparisons [redacted] in this analysis.

Results of the scenario analysis considering JAIN-like JAHL/JAHM patients and CAFÉ-like SOLO1/SOLO2 patients indicated that baricitinib monotherapy showed similar efficacy to dupilumab in achieving EASI50 response (RR: [redacted]; 95% CI: [redacted]), EASI75 response (RR: [redacted]; 95% CI: [redacted]) and 4-point DLQI improvement at Week 16 (RR: [redacted]; 95% CI: [redacted]), with no statistically significant differences (in terms of the RR) observed.

B.2.9.8 Uncertainties in the indirect and mixed treatment comparisons

Given the available data, indirect comparisons versus dupilumab could only be explored in the short term (16 weeks), so the efficacy of baricitinib compared to dupilumab in the long term is uncertain. Also, safety endpoints could not be evaluated as the studies were not comparable enough, i.e. in terms of rescue medication. Data for the composite endpoint of EASI50 + ≥4-point improvement in DLQI were not available from the CAFÉ trial. Thus, the composite endpoint of EASI50 and a ≥4-point improvement in DLQI could only be explored for the pooled comparisons including patients from the CHRONOS and SOLO1/SOLO2 trials.

This ITC does not capture a number of additional benefits that may be associated with baricitinib treatment, including the convenience of administration for patients of an oral drug such as baricitinib versus the injectable form of administration of dupilumab. Improvements in sleep disturbance due to itch and skin pain were seen with BREEZE-AD trials. These represent additional benefits of baricitinib treatment compared to dupilumab, particularly given the rapid onset of action of baricitinib. However, these endpoints are not reflected in the indirect comparison, since skin pain or sleep disturbance were not assessed in a similar manner in the dupilumab trials. Also, a comparison based on IGA could not be conducted because the dupilumab and baricitinib clinical programmes used different IGA scales.

B.2.10 *Adverse events*

Summary of safety results

- Across all BREEZE-AD trials, no clinically meaningful difference in the overall frequencies of

AEs was observed between the placebo and 4 mg baricitinib groups
- Across most BREEZE-AD trials, a higher proportion of patients in the placebo arm reported SAEs than in the 4 mg baricitinib arm
- As compared with placebo, baricitinib treatment was associated with a slightly higher proportion of TEAEs and adverse events leading to permanent discontinuation from study treatment across most trials
- No deaths occurred in the placebo or 4 mg baricitinib treatment groups across any of the trials

B.2.10.1 Summary of adverse events

The safety of baricitinib with or without concurrent TCS use versus placebo was evaluated in the BREEZE-AD trials. An integrated safety analysis of the placebo-controlled treatment period (Weeks 0–16) of the BREEZE-AD1, -AD2 and -AD7 trials is presented for all adverse event (AE) data in order to maximise the size of the safety database and improve the likelihood of observing less frequently reported AEs in patients with AD. This integrated analysis also includes data from the Phase II JAHG trial, which has not otherwise been considered within this submission due to the availability of the more relevant Phase III trials. An overview of the patients contributing to the integrated safety analysis is presented in Table 71.

Table 71: Overview of the patients contributing to the integrated safety analysis

Study	PBO (± TCS)	4 mg BARI (± TCS)
JAHG	[redacted]	[redacted]
BREEZE-AD1 (JAHL)	249	125
BREEZE-AD2 (JAHM)	244	123
BREEZE-AD7 (JAIY)	[redacted]	[redacted]
Total	[redacted]	[redacted]

[a] One patient in BREEZE-AD7 (JAIY) failed screening as was randomised to PBO in error but did not receive study treatment.
Abbreviations: BARI: baricitinib; PBO: placebo: TCS: topical corticosteroids.
Source: Simpson *et al,* 2020.[49]

Individual safety data from Weeks 0–16 of the combination therapy RCT BREEZE-AD4 (JAIN) are also presented. The safety data from Weeks 0–24 of the BREEZE-AD4 (JAIN) trial can be found in the CSR provided in the reference pack for this submission, and were found to be consistent with the Week 16 data, with no new safety signals observed.[54] Safety data from the individual BREEZE- AD1, -AD2 and -AD7 trials are presented in Appendix L, and long-term safety data from Weeks 0–52 (originating study Weeks 16–68) of the extension study BREEZE-AD3 (JAHN) are presented in Appendix M.

Across all BREEZE-AD RCTs, numerically more patients had treatment-emergent adverse events (TEAEs) in the 4 mg baricitinib group than the placebo group although no clinically meaningful differences were observed (Table 72), with nasopharyngitis representing the most common TEAE (Table 73). Despite more patients in the 4 mg baricitinib group reporting adverse events (AEs) which necessitated permanent discontinuation of study drug administration (Table 75) and adverse events of special interest (AESIs) (Table 76), more patients in the placebo group reported serious AEs (SAEs) compared to the 4 mg baricitinib groups across all BREEZE-AD trials (Table 74). No patients receiving baricitinib 4 mg or placebo died during any of the BREEZE-AD studies.

Table 72: Summary of adverse events in the BREEZE-AD RCTs

	BREEZE-AD4 (JAIN)[54]		BREEZE-AD1, -AD2, -AD7 and Study JAHG (integrated analysis)	
	PBO + TCS (N=93)	4 mg BARI + TCS (N=93)	PBO (N=[redacted])	PBO + TCS (N=[redacted])
Patients with ≥1 TEAE, n (%)	50 (53.8)	69 (75.0)	[redacted]	[redacted]
SAEs, n (%)	2 (2.2)	6 (6.5)	[redacted]	[redacted]
AEs leading to permanent discontinuation from study treatment, n (%)	1 (1.1)	1 (1.1)	[redacted]	[redacted]
AESIs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; AESI: adverse event of special interest; BARI: baricitinib; PBO: placebo; SAE: serious adverse event; TEAE: treatment emergent adverse event.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020.[58]

B.2.10.2 Treatment-emergent adverse events

TEAEs were defined as untoward medical occurrences which emerged or worsened during the treatment period and were not necessarily causally related to the treatment. All TEAEs affecting >2% of either treatment arm across all BREEZE-AD RCTs are presented in Table 73. A slightly higher proportion of patients from the 4 mg baricitinib arm of all BREEZE-AD RCTs reported TEAEs as compared with placebo, with nasopharyngitis consistently representing the most common AE.

Table 73: Summary of TEAEs affecting >3% of patients in the placebo and 4 mg baricitinib treatment groups in the BREEZE-AD trials

TEAEs affecting >3% of patients, n (%)	PBO (± TCS)	4 mg BARI (± TCS)
BREEZE-AD4 (JAIN)[54]	**N=93**	**N=92**
≥1 TEAE	50 (53.8)	69 (75.0)
Nasopharyngitis	12 (12.9)	24 (26.1)
Headache	6 (6.5)	7 (7.6)
Influenza	2 (2.2)	6 (6.5)
Abdominal pain, upper	2 (2.2)	5 (5.4)
Diarrhoea	3 (3.2)	5 (5.4)
Oral herpes	3 (3.2)	5 (5.4)
Oedema, peripheral	0 (0.0)	4 (4.3)
Abdominal pain	3 (3.2)	3 (3.3)
Back pain	3 (3.2)	3 (3.3)
Asthma	[redacted]	[redacted]
Dry eye	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
BREEZE-AD1, -AD2, -AD7 and Study JAHG (integrated analysis)	**N=[redacted]**	**N=[redacted]**
Patients with ≥1 TEAE	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Headache	[redacted]	[redacted]

Blood creatine phosphokinase increased	[redacted]	[redacted]
URTI	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; PBO: placebo; TCS: topical corticosteroid; TEAE: treatment-emergent adverse event; URTI: upper respiratory tract infection.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

B.2.10.3 Serious adverse events

SAEs were defined as any AE which resulted in death, a life-threatening experience, persistent or significant disability or incapacity, a congenital abnormality or birth defect or any important medical event which jeopardises the patient or requires intervention to prevent any of the other outcomes previously listed. All SAEs recorded in either treatment arm across all BREEZE-AD RCTs is presented in Table 74. In BREEZE-AD4, a slightly higher proportion of patients in the baricitinib treatment arm experienced SAEs than in the placebo arm, while in all other BREEZE-AD trials, a smaller proportion of patients in the 4 mg baricitinib arm experienced SAEs than in the placebo arm. In all BREEZE-AD RCTs, atopic dermatitis represented the most commonly reported SAE.

Table 74: Serious adverse events in the placebo and 4 mg baricitinib treatment groups in the BREEZE-AD trials

SAEs, n (%)	PBO (± TCS)	4 mg BARI (± TCS)
BREEZE-AD4 (JAIN)[54]	**N=93**	**N=92**
≥1 SAE	2 (2.2)	6 (6.5)
Dermatitis, atopic	1 (1.1)	2 (2.2)
Bowen's disease	1 (1.1)	0 (0.0)
Conjunctivitis, allergic	0 (0.0)	1 (1.1)
Erysipelas	1 (1.1)	0 (0.0)
Ligament rupture	0 (0.0)	1 (1.1)
Pyelitis	0 (0.0)	1 (1.1)
Soft tissue inflammation	0 (0.0)	1 (1.1)
Staphylococcal infection	0 (0.0)	1 (1.1)
BREEZE-AD1, -AD2, -AD7 and Study JAHG (integrated analysis)	**N=[redacted]**	**N=[redacted]**
≥1 SAE	[redacted]	[redacted]
Dermatitis, atopic	[redacted]	[redacted]
Eczema herpeticum	[redacted]	[redacted]
Intervertebral disc protrusion	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Alcohol poisoning	[redacted]	[redacted]
Asthma	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Breast cancer	[redacted]	[redacted]
Cataract	[redacted]	[redacted]
Clavicle fracture	[redacted]	[redacted]
Dermatitis exfoliative, generalised	[redacted]	[redacted]
Eye infection, toxoplasmal	[redacted]	[redacted]

Hypertension	[redacted]	[redacted]
Large intestine polyp	[redacted]	[redacted]
Papillary thyroid cancer	[redacted]	[redacted]
Postoperative abscess	[redacted]	[redacted]
Pulmonary embolism	[redacted]	[redacted]
Retinal detachment	[redacted]	[redacted]
Rib fracture	[redacted]	[redacted]
Suicide attempt	[redacted]	[redacted]
Tonsillitis	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; PBO: placebo; SAE: serious adverse event; TCS: topical corticosteroids.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020.[58]

B.2.10.4 Adverse events leading to permanent discontinuation from study treatment

The criteria for permanent discontinuation from the study treatment are presented in Appendix L. All AEs which resulted in permanent discontinuation from study treatment in either treatment arm across all BREEZE-AD studies is presented in Table 75. The occurrence of AEs necessitating permanent discontinuation from study treatment was well balanced between the placebo and 4 mg baricitinib arms of BREEZE-AD4. Across other BREEZE-AD trials where discontinuation due to AEs was proportionally slightly higher in the baricitinib groups, the specific AE was varied within treatment groups with no particular AE of concern emerging.

Table 75: Adverse events leading to permanent discontinuation from study treatment in the placebo and 4 mg baricitinib treatment groups in the BREEZE-AD trials

AE leading to permanent discontinuation from study treatment, n (%)	PBO (± TCS)	BARI 4 mg (± TCS)
BREEZE-AD4 (JAIN)	**N=93**	**N=92**
AE leading to permanent discontinuation from study treatment	1 (1.1)	1 (1.1)
Blood alkaline phosphatase increased	1 (1.1)	0 (0.0)
Skin infection	0 (0.0)	1 (1.1)
BREEZE-AD1, -AD2, -AD7 and Study JAHG (integrated analysis)	**N=[redacted]**	**N=[redacted]**
AE leading to permanent discontinuation from study treatment	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]
Dizziness	[redacted]	[redacted]
Toxic skin eruption	[redacted]	[redacted]
White blood cell count decreased	[redacted]	[redacted]
Dermatitis atopic	[redacted]	[redacted]
Eczema	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Asthma	[redacted]	[redacted]
Breast cancer	[redacted]	[redacted]
Dermatitis exfoliative, generalised	[redacted]	[redacted]
Haematuria	[redacted]	[redacted]

Headache	[redacted]	[redacted]
Lymphocyte count abnormal	[redacted]	[redacted]
Papillary thyroid cancer	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Postoperative abscess	[redacted]	[redacted]
Pulmonary embolism	[redacted]	[redacted]
Skin ulcer	[redacted]	[redacted]
URTI	[redacted]	[redacted]

Abbreviation: AE: adverse event; BARI: baricitinib; PBO: placebo; URTI: upper respiratory tract infection.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020[58]

B.2.10.5 Adverse events of special interest (AESI)

AESIs were defined as infections, malignancies, hepatic events as defined by abnormal clinical liver tests, major adverse cardiovascular events including myocardial infarction or stroke, and thrombotic events, including deep vein thrombosis. All AESIs recorded in either treatment arm across all BREEZE-AD studies is presented in Table 76. Across all trials, infections were the most common AESIs. While a slightly higher proportion of patients in the 4 mg baricitinib arm reported treatment-emergent infections than those in the placebo arm across all BREEZE-AD trials, the proportion of serious infections was well balanced across treatment arms in all BREEZE-AD studies.

Table 76: Adverse events of special interest in the placebo and 4 mg baricitinib treatment groups in the BREEZE-AD trials

AESI, n (%)	PBO (± TCS)	BARI 4 mg (± TCS)
BREEZE-AD4 (JAIN)	**N=93**	**N=92**
Any TE infection	[redacted]	[redacted]
Serious infection	[redacted]	[redacted]
Herpes zoster	[redacted]	[redacted]
Herpes simplex[a]	[redacted]	[redacted]
Infections led to study drug treatment interruption	[redacted]	[redacted]
Infections led to study drug treatment discontinuation	0 (0.0)	1 (1.1)
Other	[redacted][b]	[redacted]
BREEZE-AD1, -AD2, -AD7 and Study JAHG (integrated analysis)	**N=[redacted]**	**N=[redacted]**
Any TE infection	[redacted]	[redacted]
Serious infection	[redacted]	[redacted]
Herpes simplex	[redacted]	[redacted]
Herpes zoster	[redacted]	[redacted]
Infections led to study drug treatment interruption	[redacted]	[redacted]
Infections led to study drug treatment discontinuation	[redacted]	[redacted]
Patients with ≥1 skin infection requiring antibiotic treatment	[redacted]	[redacted]

Other[c]	[redacted]	[redacted]

[a] Herpes simplex includes the terms oral herpes, herpes simplex, eczema herpeticum, genital herpes simplex, genital herpes and Kaposi's varicelliform eruption. It does not include herpes zoster, of which no cases were observed in either treatment arm of BREEZE-AD4. [b] Other AESI in the BREEZE-AD4 placebo group was non-melanoma skin cancer (N=1). [c] Does not include data from Study JAHG. [d] Other AESI in the integrated analysis PBO group were breast cancer (N=1), papillary thyroid cancer (N=1) and opportunistic infection (N=1). [e] Other AESI in the integrated analysis 4 mg baricitinib group was pulmonary artery embolus (N=1).
Abbreviations: AESI: adverse event of special interest; BARI: baricitinib; PBO: placebo; TE: treatment-emergent.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] Bieber *et al*, 2020.[58]

B.2.10.6 Patients rescued and rescue therapies used

The number of patients who received rescue during the trial period of the originating BREEZE-AD RCTs are summarised in Table 77. Overall at Week 16, fewer patients were rescued in the 4 mg baricitinib group than the placebo group in BREEZE-AD7, -AD1 and -AD2, with similar rescue rates observed between treatment arms in BREEZE-AD4. A summary of the types of rescue medication used in these trials is provided in Table 78.

Table 77: Cumulative number of patients rescued during the BREEZE-AD RCTs

Time point (week)	Cumulative number of patients rescued, n (%)							
	BREEZE-AD4 (JAIN)[54]		BREEZE-AD7 (JAIY)[55]		BREEZE-AD1 (JAHL)[56]		BREEZE-AD2 (JAHM)[57]	
	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO + TCS (N=[redacted])	BARI 4 mg + TCS (N=[redacted])	PBO (N=[redacted])	4 mg BARI (N=[redacted])	PBO (N=[redacted])	4 mg BARI (N=[redacted])
1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]	100 (40.2)	15 (12.0)	132 (54.1)	35 (28.5)
4	[redacted]	[redacted]	[redacted]	[redacted]	132 (53.0)	25 (20.0)	159 (65.2)	45 (36.6)
8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	[redacted]	[redacted]	[redacted]	[redacted]	161 (64.7)	51 (40.8)	186 (76.2)	70 (56.9)
16	[redacted]	[redacted]	[redacted]	[redacted]	166 (66.7)	51 (40.8)	187 (76.6)	72 (58.5)
20	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; PBO: placebo; TCS: topical corticosteroids.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] BREEZE-AD7 (JAIY) Clinical Study Report,[55] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report,[57] Simpson *et al,* 2020.[49]

Table 78: Summary of the types of rescue medications used in the BREEZE-AD RCTs

Rescue medications, n (%)	PBO (± TCS)	BARI 4 mg (± TCS)
BREEZE-AD4 (JAIN)[54]	N=[redacted]	N=[redacted]
Any rescue	[redacted]	[redacted]
Rescue TCS[a,b]	[redacted]	[redacted]

Phototherapy[a]	[redacted]	[redacted]
Systemic medications[a]	[redacted]	[redacted]
Corticosteroids[a]	[redacted]	[redacted]
Biologics	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]
BREEZE-AD7 (JAIY)[55]	**N=[redacted]**	**N=[redacted]**
Any rescue	[redacted]	[redacted]
Rescue topical corticosteroids	[redacted]	[redacted]
Systemic medications	[redacted]	[redacted]
Corticosteroids	[redacted]	[redacted]
Ciclosporin	[redacted]	[redacted]
BREEZE-AD1 (JAHL)[56]	**N=[redacted]**	**N=[redacted]**
Any rescue	[redacted]	[redacted]
TCS	[redacted]	[redacted]
TCI	[redacted]	[redacted]
Systemic medications	[redacted]	[redacted]
Corticosteroids	[redacted]	[redacted]
Ciclosporin	[redacted]	[redacted]
BREEZE-AD2 (JAHM)[57]	**N=[redacted]**	**N=[redacted]**
Any rescue	[redacted]	[redacted]
TCS	[redacted]	[redacted]
TCI	[redacted]	[redacted]
Systemic medications	[redacted]	[redacted]
Corticosteroids	[redacted]	[redacted]
Ciclosporin	[redacted]	[redacted]

[a] Percentages have been adjusted from those presented in the Clinical Study Report to present the proportion of patients in the treatment arm who received the rescue therapy. [b] Includes high potency, ultra-high potency and unclassified.
Abbreviations: BARI: baricitinib; PBO: placebo; TCI: topical calcineurin inhibitor; TCS: topical corticosteroids.
Sources: BREEZE-AD4 (JAIN) Clinical Study Report,[54] BREEZE-AD7 (JAIY) Clinical Study Report,[55] BREEZE-AD1 (JAHL) Clinical Study Report,[56] BREEZE-AD2 (JAHM) Clinical Study Report.[57]

B.2.11 *Ongoing studies*

The BREEZE-AD3 and -AD4 trials are ongoing. Additional data from BREEZE-AD4 may become available in October 2020, and additional data from BREEZE-AD3 in November 2020. However, given the current COVID-19 pandemic, these dates are may be subject to considerable change.

B.2.12 *Innovation*

Baricitinib (Olumiant®) has a novel, targeted mode of action, selectively and reversibly inhibiting the JAK family of protein tyrosine kinases, specifically JAK1 and JAK2, which mediate pathways involved in the inflammatory processes underlying AD. Baricitinib is administered orally as a monotherapy or in combination with TCS.

As discussed in Section B.1.3.3, it is expected that clinicians will use baricitinib as an alternative to dupilumab following consideration of a systemic immunosuppressant agent. This is in line with

the clinical positioning of baricitinib in current UK practice and the eligibility criteria for the BREEZE-AD4 (JAIN) trial. Dupilumab has been recommended by NICE for adults with severe-to-moderate AD who experience failure with, are intolerant to or have contraindication to at least one systemic therapy.[1] Whilst dupilumab may be effective in controlling the disease, there are considerable limitations to its use. Unlike baricitinib, dupilumab is administered via subcutaneous injection every other week. Many patients experience injection site reactions, with over 1 in 10 patients experiencing swelling at the site of injection, and more than 1 in 100 reporting redness, pain or itch at the injection site.[46] Eye disorders such as conjunctivitis are also common adverse events of dupilumab treatment. In the CAFÉ trial, 28% patients receiving dupilumab (every other week in combination with TCS) experienced conjunctivitis, which was severe in 0.9% and moderate in 12.1% patients.[47] These adverse events result in additional health care resource use through the need for consultant ophthalmologist visits. As such, there is a clear unmet clinical need for an effective, tolerable, easily-administered treatment option for patients whose only alternative is dupilumab.

As an orally administrated therapy, baricitinib is not associated with such limitations, and thus has the potential to dramatically simplify the treatment paradigm for patients in this setting, and potentially facilitate a reduction in health care resource utilisation. The efficacy and safety of baricitinib (as a monotherapy and in combination with TCS) has been demonstrated in four randomised, placebo-controlled, Phase III studies (BREEZE-AD1, -AD2, -AD4 and -AD7), leading to statistically significant and clinically meaningful improvements in the signs and symptoms of AD compared with placebo (IGA, EASI, SCORAD, Itch NRS, Skin Pain NRS, ADSS) (See Section B.2.6). The results of the ITC also indicate that baricitinib has similar efficacy to dupilumab in terms of EASI response and Itch NRS. Skin itch, skin pain and sleep disturbance have been shown to have substantial impact on HRQoL in patients with moderate-to-severe AD,[26] and are alleviated through treatment with baricitinib as shown by statistically significant improvements in novel PROs: Itch NRS, Skin Pain NRS and ADSS. These improvements are reflected by statistically significant improvements in HRQoL outcomes (DLQI an EQ-5D-5L).

Baricitinib provides an effective, tolerable, easily-administered treatment option for patients whose only alternative is dupilumab, and thus has the potential to be first in class for patients with moderate-to-severe AD who experienced failure with, are intolerant to or have contraindication to at least one systemic therapy.

B.2.13 *Interpretation of clinical effectiveness and safety evidence*

Principal findings from the clinical evidence base

The efficacy and safety of baricitinib (as a monotherapy or in combination with TCS) has been demonstrated in four randomised, placebo-controlled, Phase III studies (BREEZE-AD1, -AD2, -AD4 and -AD7). At baseline across all the trials, patients had moderate-to-severe disease, with at least a third of participants with IGA 4 at screening. Despite this, results from these trials show baricitinib to be an effective treatment option for moderate-to-severe AD associated with robust and rapid improvement of symptoms, including a significant reduction in itch by Week 2 and improved sleep by Week 1 as determined by ADSS Item 2. These trials also found baricitinib to have a tolerable safety profile with nasopharyngitis consistently representing the most common AE and no safety signals of concern observed.

In the context of current clinical practice within the NHS in England, this submission positions baricitinib as an alternative treatment to dupilumab following inadequate response to a systemic therapy, or if these are contraindicated or not tolerated. This is narrower than the full marketing authorisation for baricitinib and the population specified in the NICE scope. Therefore, the clinical effectiveness evidence in this submission focusses on the BREEZE-AD4 (JAIN) trial, the inclusion criteria of which reflects this population of interest, and an ITC was performed to provide relative effectiveness data versus dupilumab. The results of the ITC indicate that both baricitinib (4 mg QD) monotherapy and in combination with TCS have similar efficacy to dupilumab (300 mg Q2W).

Strengths and limitations of the clinical evidence base

The clinical evidence presented within this submission has been derived from an SLR of clinical trials investigating the efficacy and safety of treatment options, including baricitinib, for moderate-to-severe AD.

The BREEZE-AD trials represent the primary sources of evidence for baricitinib as a treatment for adult patients with moderate-to-severe AD. The BREEZE-AD1, -AD2, -AD4 and -AD7 trials are large, placebo-controlled RCTs, and thus provide robust evidence for the safety and efficacy of baricitinib for the treatment of adult patients with AD. Additionally, as discussed in Section B.2.5, the BREEZE-AD trials can be considered of good quality. To synthesise relative effectiveness and safety data of baricitinib versus dupilumab in the population of relevance for this submission, an ITC was conducted. According to UK clinical experts consulted as part of the submission, the patient baseline characteristics of patients included in the ITC are considered to be generally consistent with what may be expected of patients in clinical practice in England.

A key limitation of the evidence base is the lack of direct evidence identified for baricitinib versus dupilumab to inform relative efficacy estimates, since the BREEZE-AD trials are placebo-controlled. However, the SLR identified 11 RCTs investigating dupilumab, all of which were placebo-controlled (see Section B.2.9.1 and Appendix D). As such, it was possible to conduct an ITC with placebo as a common comparator. Despite being associated with some uncertainty due to heterogeneity in study design and a small number of included studies, the ITCs allowed derivation of relative efficacy estimates for baricitinib versus dupilumab, its most relevant clinical comparator. The ITC demonstrates that baricitinib has similar efficacy compared with dupilumab for the population of relevance for this submission, and these results appeared to be generally consistent for both baricitinib as a monotherapy and in combination with TCS across a number of different endpoints and scenario analyses.

B.3 Cost effectiveness

B.3.1 *Published cost-effectiveness studies*

An SLR was conducted to identify any relevant economic evaluations for the treatment of adult patients with moderate-to-severe AD. The original SLR was performed in March 2018 and was updated in February 2020. In total, 17 studies featuring relevant health state utility or cost and resource use data associated with the treatment of adult patients with moderate-to-severe AD were identified: 15 from the original SLR, and a further 2 in the update SLR. Of these, 2 publications, both HTA reports, were subsequently used to inform inputs within the economic analysis presented in this submission; these are presented in Table 79. Full details of the SLR search strategy, study selection process and results are reported in Appendix G.

Table 79: Summary of relevant cost-effectiveness studies

Study	Institute for Clinical and Economic Review (2017)	NICE (2018)
Summary of model	Markov model	Combined one-year decision tree and three-state Markov model
Intervention	Dupilumab	Dupilumab
Comparator(s)	Usual care (emollients, TCS, TCI, phototherapy or ciclosporin)	BSC (emollients, low-to-mid potency TCS and rescue therapy, such as higher potency TCS, oral corticosteroids or TCIs)
Patient population	Adults with moderate-to-severe AD who have experienced inadequate response or contraindication to topical therapies	Adults with moderate-to-severe AD who have experience inadequate response, intolerance or contraindication to ciclosporin
QALYs (intervention, comparator)	1.91	Confidential
Costs (intervention, comparator)	*Using net price:* USD $389,415 *Using list price:* USD $476,264	Confidential
ICER (cost per QALY gained)	*Using net price:* USD $101,800 *Using list price:* USD $124,541	*Pooled CAFÉ + CAFÉ-like CHRONOS population:* £28,874 *CAFÉ population:* £24,703

Abbreviations: AD: atopic dermatitis; BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life years; TCI: topical calcineurin inhibitors; TCS: topical corticosteroids.

B.3.2 *Economic analysis*

The objective of this economic analysis was to assess the cost effectiveness of baricitinib compared with dupilumab and BSC for the treatment of moderate to severe AD patients who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. The base case population is considered to be relevant to UK clinical practice, reflecting the anticipated positioning for baricitinib in the treatment pathway and the highest unmet clinical need, The SLR did not identify any studies evaluating the cost-effectiveness of baricitinib in moderate-to-severe AD. A *de novo* cost-effectiveness analysis of

baricitinib versus comparators relevant to the decision problem for this submission was therefore performed. A Markov structure was deemed appropriate to adequately capture the key features of AD. In line with the NICE reference case, the analysis was conducted from the perspective of the NHS and Personal Social Services (PSS) and included direct medical costs only over a lifetime time horizon.[76] Sections B.3.2.1, B.3.2.2 and B.3.2.3 present the patient population considered in the model, the model structure and the included interventions and comparators, respectively.

B.3.2.1 Patient population

This economic evaluation considers the cost-effectiveness of baricitinib in adult patients with moderate-to-severe AD who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control, in line with the indication of relevance for this submission.

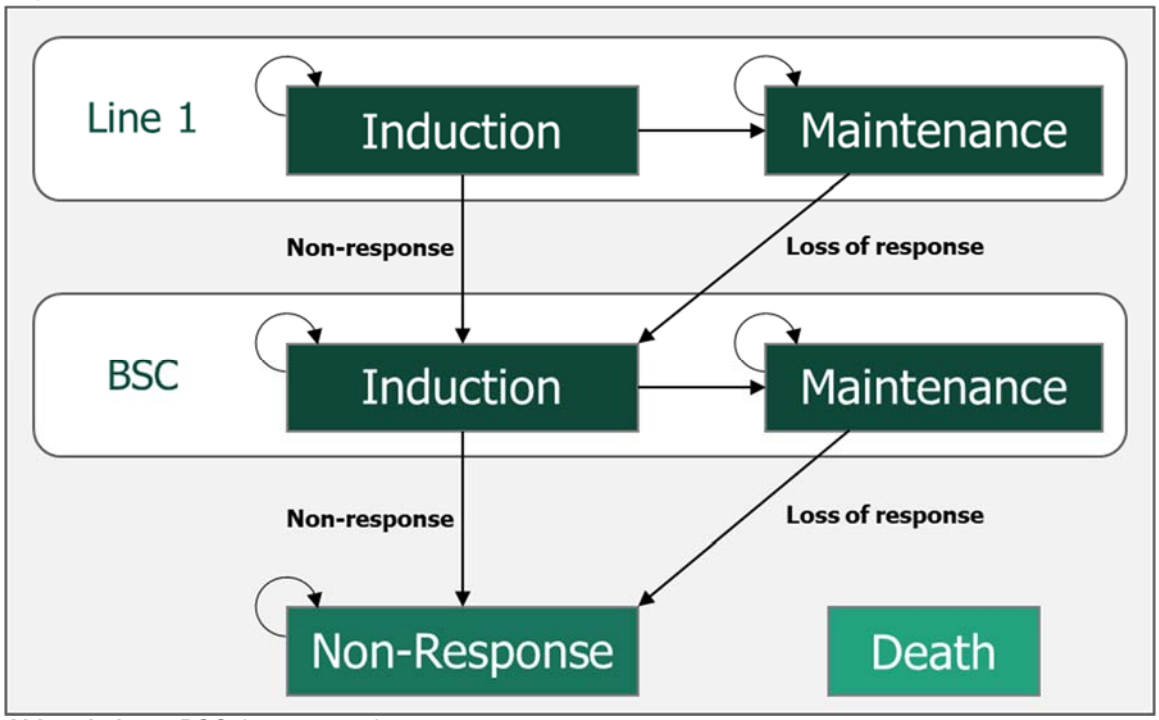
As shown in Figure 4 in Section B.2.1, the patient populations included in the economic evaluation (JAIN, JAIN + JAIY JAIN-like and JAHL + JAHM JAIN-like) are in line with the eligibility criteria for the JAIN trial: patients who had a history of intolerance, inadequate response or contraindication to ciclosporin. All of these patients meet the population definition considered in the economic evaluation. Whilst not interchangeable, it can be assumed that ciclosporin is broadly comparable to azathioprine, methotrexate, and mycophenolate mofetil, as these systemic therapies are considered at the same stage of the treatment pathway in UK clinical practice. The pooled population of JAIN + JAIY JAIN-like patients, which is used in the base case analysis, includes patients who received baricitinib and concomitant TCS, reflecting the anticipated treatment pathway in the UK. As such, the efficacy data for these populations from the Phase III trials was considered generalisable to the target population in UK clinical practice. The patient population considered in the economic evaluation is also narrower than the population specified in the NICE final scope and the full anticipated marketing authorisation for baricitinib in AD: adult patients with moderate-to-severe AD who are candidates for systemic therapy. However, this narrower population is consistent with the anticipated eligible patient population for baricitinib in UK clinical practice where it is expected to provide clinicians with an alternative to dupilumab following consideration of a systemic immunosuppressant agent.

B.3.2.2 Model structure

A cohort Markov state transition model was chosen to evaluate the cost-effectiveness of baricitinib versus dupilumab and best supportive care (BSC) in the target population and was constructed in Microsoft Excel. The model structure aimed to adequately capture the key features of AD be reflective of clinical practice in the UK.

The model structure is presented in Figure 40. The model facilitates pairwise comparisons and fully incremental analysis of treatments, in line with previous assessments. The model includes four health states: “Induction”, which is represented by a set of tunnel states, followed by “Maintenance”, “Non-Response” and “Death”. The model as built allows treatment sequencing to be evaluated, but this feature is not presented in this submission as it is not relevant to the UK decision problem given the positioning of baricitinib and dupilumab as fifth line therapies.

Figure 40: Model structure



Abbreviations: BSC: best supportive care.
Arrows to the Death health state have been removed for clarity; Death can be reached from any other health state at any time.

Upon entering the model, patients are allocated to baricitinib or a comparator treatment (dupilumab or BSC) and enter the Induction health state for that treatment. The length of the Induction period is 16 weeks for baricitinib, dupilumab and BSC aligning with the double-blinded treatment period of the Phase III RCTs for baricitinib and dupilumab and the time-point for clinical assessment of response in the UK.[54, 55, 71, 72] This 16 week Induction period is achieved through using four tunnel states, each with a cycle length of four weeks. A tunnel state is a type of temporary health state which can only be visited once in a fixed sequence.[77] In the model, patients cannot discontinue during the Induction period.

At the end of the Induction period, patient response to treatment is assessed; in the base case, response is defined as EASI50 with a DLQI improvement of four or more points (ΔDLQI ≥4), which was the preferred option of the Appraisal Committee during the dupilumab NICE appraisal (TA534).[1] Patients who respond to treatment transition to the Maintenance health state, whilst patients who do not respond transition to BSC treatment.

Patients who enter the Maintenance health state are modelled to receive continuous treatment, during which they are at risk of discontinuation as a consequence of loss of response or due to other factors such as severe AEs (captured by all-cause discontinuation). The probability of annual all-cause discontinuation is assumed to be constant (5.2% for baricitinib and dupilumab) and reflects the withdrawal probability observed in the dupilumab CHRONOS trial because Week 52 data are not yet available from the BREEZE-AD4 (JAIN) trial (see Section B.3.3.3).[72]

Patients remain in the Maintenance health state until they discontinue treatment, after which they transition to BSC treatment. No further lines of treatment are available for patients who do not respond to BSC or who discontinue BSC following loss of response, and these patients transition to the Non-Response health state. Upon entering the Non-Response health state, patients

remain there until death or the end of the simulation. In the model, Death represents the absorbing state, accumulating patient flows from all health states. There is no assumption for treatment effect on mortality and thus it is assumed that the probability of transition from any of the other health states to Death is equal within each cycle. The model includes normal UK population mortality (see Section B.3.3.5).

Features of economic analysis

The key features of the economic analysis and their justifications are presented in Table 80. Health state utility values are derived by cross-walking EQ-5D-5L scores collected in the BREEZE-AD trials to EQ-5D-3L scores using the algorithm presented in van Hout *et al*. 2012, in line with the NICE reference case.[52, 78] These scores are subsequently used to generate utility index values using the UK value set by Dolan *et al*. 1997.[70] Costs considered within the model include treatment acquisition costs, associated administration costs and adverse event costs. Effectiveness measures include life years (LYs) and quality-adjusted life years (QALYs). The incremental cost-effectiveness ratio (ICER) of baricitinib versus each comparator is evaluated in terms of the incremental cost per QALY gained. An annual discount of 3.5% is applied for both costs and QALYs.

The analysis is conducted from the perspective of the UK NHS and Personal Social Services (PSS) over a lifetime horizon which is considered appropriate given the chronic nature of AD. Maximal lifetime for patients is set to 100 years, reflecting that the ONS life tables for mortality end at 100.[35]

The cycle length employed in the Markov model is four weeks and half-cycle correction was not included in the model due to the short cycle lengths. Given the different time reference of model inputs, including annually or per three-month period, calculations are performed in the model to rescale all variables to four-week duration. Two methods for rescaling are used, depending on the nature of the input. For probabilities, the probability is converted to a constant instantaneous rate, which is in turn converted to the desired length probability of four weeks. For the inputs related to absolute levels, such as annual frequency or annual number of flares, linear conversion is applied by dividing the number of days in the desired length of four weeks by the number of days per year and multiplying this by the annual frequency of the event.

Table 80: Features of the economic analysis

Factor	Previous appraisal: TA534[1]	Current appraisal	
		Chosen values	Justification
Model structure	One-year decision tree followed by a three-state Markov model	Markov state transition model with 4-week cycles	A Markov state transition model approach was chosen as this model structure is in line with previous models in AD and is reflective of clinical practice in the UK[1]
Time horizon	Lifetime	Lifetime	In line with the NICE reference case[52] and considered to reflect that AD is a chronic disease expected to affect a patient over a

			lifetime and will ensure the model captures all costs and benefits of intervention and comparators
Source of utilities	Utility values were estimated based on a mixed model regression. The utility values were adjusted multiplicatively for the impact of ageing on HRQoL.	Health state utility values are derived by cross-walking EQ-5D-5L scores collected in the BREEZE-AD trials to EQ-5D-3L scores, using the algorithm presented in van Hout *et al*. 2012, in line with the NICE reference case, and are subsequently used to generate utility index values using the UK value set by Dolan *et al*, 1997.[52, 70, 78] The utility values were adjusted for the impact of ageing on HRQoL.	In line with previous models in AD and with the NICE reference case[1, 52]
Source of costs	BNF (2017), the PSSRU and National Reference Costs (2015), National Schedule of Reference Costs (2015–16) and NHS reference costs (2014–2015)	National schedule of NHS Costs (2018–19) PSSRU and National Reference Costs (2019) and the BNF (2019)	Established sources of costs within the NHS. In line with the NICE reference case and TA534[1, 52]
Resource use	The adverse events considered in the model were based on those reported in the dupilumab clinical trials	Resource use was derived from TA534[1]	Resource use was not captured within the BREEZE-AD trials but the TA534 was considered a relevant resource use data source. Other sources are established sources of costs within the NHS
Health effects measure	QALYs	QALYs	NICE reference case[52]
Half cycle correction applied?	Yes – yearly cycles with half-cycle correction	No	Half-cycle correction was not included in the model due to the short 4-week cycle length

Abbreviations: MIMS: Monthly Index of Medical Specialities; PSSRU: Personal Social Services Research Unit
Source: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534).[1]

B.3.2.3 Intervention technology and comparators

The intervention of interest is 4 mg baricitinib administered orally once a day. This is in line with the regimen used in the Phase III BREEZE-AD trials supporting the submission and the SmPC

for baricitinib.[79] In the base case, baricitinib is modelled to be used in combination with TCS as this is considered to represent typical AD management in the UK.[79]

As discussed in Section B.1.3.3, for patients who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control, dupilumab is the recommended treatment option. If dupilumab fails to control the disease, or in patients for whom use of dupilumab is not recommended or contraindicated, no further safe and effective treatment options are available, so patients are treated with best supportive care (BSC). Baricitinib is positioned as a 5th line therapy for the treatment of adult patients with moderate-to-severe AD as an alternative to dupilumab following consideration of a systemic immunosuppressant agent. The comparators included in the model (dupilumab and BSC) therefore reflect the standard of care for patients in this setting in UK clinical practice (as reflected in TA534), and the available evidence identified in the clinical SLR.[80]

The dose for dupilumab included in the model was a loading dose of 600 mg s.c. followed by 300 mg s.c. injection every other week and is aligned with the licensed indication for its use in AD and the Phase III RCTs CAFÉ and CHRONOS.[34, 71, 72]

If dupilumab fails to control the disease, or in patients for whom use of dupilumab is not recommended or contraindicated, no further safe and effective treatment options are available, so patients are treated with best supportive care (BSC) which remains poorly defined in UK clinical practice. In line with the NICE draft scope for this appraisal and based on placebo regimens included in the BREEZE-AD trials, BSC in the model is defined as emollients, low-to-mid potency topical corticosteroids, phototherapy, psychological support, and rescue therapy including higher potency topical or oral corticosteroids or topical calcineurin inhibitors.

Within the economic model, a stopping rule is applied for patients who do not respond to treatment at 16 weeks for baricitinib, dupilumab and BSC.

B.3.3 *Clinical parameters and variables*

As described in Section B.3.2.2, four distinct AD health states are defined; in the base case, these are defined based on achievement of EASI50 with ΔDLQI ≥4. Patients transition between the Induction and Maintenance (i.e. responder) or Non-Responder health states depending on changes in AD severity experienced following treatment with the intervention or comparators. Patients in the Maintenance health state may over time transition to the Non-Responder or Death health states. Once patients enter the Non-Responder health state, they remain in that state until the end of the model simulation or death, with Death representing the absorbing state.

Key efficacy data and utility inputs for baricitinib are derived from the pivotal BREEZE-AD trials. As discussed in Section B.2.9, in the absence of head-to-head evidence between baricitinib and the comparator dupilumab, an ITC versus dupilumab was performed to inform the base case economic analysis. Given the lack of data for the composite endpoint of EASI50 + ≥4-point improvement in DLQI for the comparison of JAIN versus CAFÉ, the comparison of pooled JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients was used in the base case analysis of the model, and has thus been presented first here. The placebo + TCS data from the pooled analysis of JAIN + JAIN-like JAIY patients are employed to represent the effectiveness of BSC alone in the model.

The sources for the clinical parameters used in the economic model are summarised in Table 81 and discussed below in turn.

Table 81: Summary of sources of data used in the economic model

Parameter	Baricitinib	Dupilumab	Reference in submission
Baseline characteristics	JAIN + JAIN-like JAIY patients		Section B.3.3.1
Treatment response	ITC: JAIN + JAIN-like JAIY	ITC: CAFÉ + CAFÉ-like CHRONOS	Section B.3.3.2
Sustained effectiveness up to 52 weeks	Dupilumab submission (TA534)		Section B.3.3.3
Long-term treatment discontinuation	Dupilumab submission (TA534)		Section B.3.3.3
Adverse events	JAIN + JAIN-like JAIY patients	Dupilumab submission (TA534)	Section B.3.3.4
Mortality	General UK population		Section B.3.3.5

Source: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534).[1]

B.3.3.1 Baseline characteristics

The baseline characteristics of the modelled cohort and their source are presented in Table 82. The start age and the proportion of male patients in the modelled cohort is derived from the pooled population of JAIN + JAIN-like JAIY patients.[54, 55] No differences in population characteristics are assumed between interventions.

Table 82: Baseline characteristics for base case JAIN + JAIY JAIN-like population

Component	Base case value	Source
Start age (years)	[redacted]	JAIN + JAIN-like JAIY patients (n=[redacted])[54, 55]
Male, %	[redacted]	
EASI score, mean	[redacted]	
EQ-5D HIS, mean (SD)	[redacted]	

Abbreviations: EASI: Eczema Area and Severity Index; EQ-5D: EuroQol 5 Dimensions; HIS: Health index Score; SD: standard deviation.

B.3.3.2 Treatment response

Treatment response rates for baricitinib and dupilumab are based on the results of the ITC analysis presented in Section B.2.9.

Treatment response rates for BSC are based on data for patients receiving placebo in the pooled population of JAIN + JAIN-like JAIY patients. The level of response to placebo, which comprised emollients and TCS, observed in the clinical trial is unlikely to be observed in clinical practice, since patients eligible for baricitinib are candidates for systemic treatment and therefore topical treatments have previously not been sufficient to control the disease. The base case response rates used in the model are presented in Table 83.

Table 83: Response rates for EASI50 + ΔDLQI≥4 employed in the base case analysis

	Response probability, % (SE%)
Baricitinib	48.99 (4.09)

Dupilumab	79.25 (3.00)
BSC	31.25 (3.86)

Abbreviations: BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; SE: standard error.

Sensitivity analyses were conducted using alternative response definitions to determine whether patients in the JAIN + JAIN-like JAIY population continued on the same treatment during the Maintenance period to which they were assigned in the Induction period, or whether they switched to the next treatment in the sequence. A summary of the response rates for the available alternative response definitions included in the model is presented in Table 84.

Table 84: Summary of response probabilities for alternative response definitions included in the model

	Response probabilities, % (SE%)	
	EASI50	EASI75
Baricitinib	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]
BSC	[redacted]	[redacted]

The outcomes of EASI90, ΔItch NRS ≥4 at Week 4 and ΔItch NRS ≥4 at Week 16 are not available for the base case population of JAIN/JAIN-like versus CAFÉ/CAFÉ-like CHRONOS.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; SE: standard error.

B.3.3.3 Sustained response and long-term treatment discontinuation

Sustained response up to 52 weeks

After the end of the trial period, responders enter the Maintenance treatment phase and receive continuous treatment. During this phase, patients are at risk of discontinuation as a consequence of loss of response or due to other factors such as severe AEs. The sustained effectiveness of baricitinib, dupilumab and BSC up to Week 52 is modelled by applying a discontinuation rate reflecting loss of response (and any other factors such as severe AEs). Discontinuation rates for dupilumab and BSC are informed by the conditional probability of response at Week 52 given response at Week 16 in the dupilumab submission (TA534).[1] A reliable and valid estimate for discontinuation rates based on JAIN data could not be generated, as the extrapolation would have been dependent on one single time point (i.e. 16–24 weeks), which is the only available data for the sustained effectiveness of baricitinib to date. This would likely overestimate the discontinuation rate for baricitinib. From an economic standpoint, overestimating the discontinuation rate from Week 16 to 52 in the model risks underestimating the total cost of baricitinib treatment and biasing the model in favour of baricitinib. For this reason, an assumption of equivalence to dupilumab is made for the model base case.

The conditional probabilities of response at Week 52 are used to estimate the probability of treatment discontinuation due to loss of response between the end of the Induction period (Week 16 in the base case analysis) and Week 52. This is performed by deducting the conditional probability of response at Week 52 from the probability of response at the time of response assessment. Thereafter, the four-week probability of treatment discontinuation is derived as described in Section B.3.2.2. This approach assumes that patients who have lost response between the response assessment and Week 52 have done so at a continuous and constant rate and that they discontinue treatment once they have lost response.

The base case response probabilities at Week 52 used in the model are presented in Table 85.

Table 85: Response probabilities for EASI50 + ΔDLQI≥4 at Week 52 conditional upon response at Week 16 for baricitinib and comparators employed in the base case analysis

	Response probability, % (SE%)[a]
Baricitinib[b]	93.9 (2.8)
Dupilumab	93.9 (2.8)
BSC	76.7 (4.8)

[a] Assumed to be the same as observed at Week 52 in CHRONOS. [b] Assumed to be the same as dupilumab.
Abbreviations: BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; SE: standard error.

The response probabilities at Week 52 using alternative response definitions were assessed in scenario analyses. These response probabilities are presented in Table 86.

Table 86: Response probabilities at Week 52 conditional upon response at Week 16 for baricitinib and comparators on categorial endpoints

	Response probabilities, % (SE%)[a]	
	EASI50[c]	EASI75
Baricitinib[b]	82.1 (5.3)	82.1 (5.3)
Dupilumab	82.1 (5.3)	82.1 (5.3)
BSC	70.6 (6.4)	70.6 (6.4)

[a] Assumed to be the same as observed at Week 52 CHRONOS. [b] Assumed to be the same as dupilumab.
[c] Assumed to be the same as EASI75.
The outcomes of EASI90, ΔItch NRS ≥4 at Week 4 and ΔItch NRS ≥4 at Week 16 are not available for the base case population of JAIN/JAIN-like versus CAFÉ/CAFÉ-like CHRONOS.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; NRS: numerical rating scale; SE: standard error.

Long-term discontinuation after 52 weeks

After 52 weeks, the model includes an annual probability of discontinuation that represents the annual rate at which patients discontinue baricitinib or dupilumab each year due to lack of long-term efficacy, adverse event, patient preference, or physician preference.

After 52 weeks, there are no placebo-controlled trial data and it is therefore assumed that dupilumab and baricitinib have the same annual probability of treatment discontinuation. The annual probability of discontinuation is applied to patients in the Maintenance health state starting at the second year of the model. First year data are based on sustained response data definitions. Patients who discontinue dupilumab or baricitinib during this time transition to BSC.

The annual probability of discontinuation for the second and subsequent years for dupilumab and for baricitinib has been set to 3.7% for the composite response criterion and 5.1% when EASI75 is chosen as response criterion, reflecting the withdrawal probabilities observed in the dupilumab CHRONOS trial.[72] A reliable and valid estimation for long-term discontinuation rates (beyond 52 weeks) based on JAIN data is not yet available.

Table 87: Annual probabilities of discontinuation after Week 52

Trial response at week 16	Annual probability of discontinuation, %
EASI50 AND DLQI ≥4	3.7

EASI75	5.1

Abbreviations: DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index.

There is limited data on the loss of response to BSC outside the trial setting in the relevant patient segment. However, in the base case we implemented the annual probability of study withdrawal or use of rescue medication from the BSC arm in the CHRONOS trial: 57.0%.

B.3.3.4 Adverse events

The AEs considered in the model are based on the most frequent and serious reported adverse events reported in the baricitinib clinical trials for AD and the dupilumab NICE submission (TA534) for dupilumab and BSC.[1] As adverse event data for baricitinib are only available for the 16 week trial period, annual probabilities were calculated by transforming the 16-week probabilities into 16-week rates. Annual rates were then calculated, which were then transformed back into annual probabilities to be inputted into the model.

The probabilities of AEs are assumed to remain constant over the treatment duration, meaning that patients have the same risk of AEs in each cycle. Given that baricitinib, dupilumab and BSC have different methods of administration and modes of action, the modelled AEs differ by treatment. In the model, AEs are not modelled as separate health states; instead the rates and the consequences of AEs impact the costs and utility accumulated in each cycle. The adverse event probabilities used in the base case of the model are presented in Table 88.

Table 88: Annual adverse event probabilities used in the model base case

AE probability, %	4 mg QD baricitinib	Dupilumab Q2W	BSC
Injection site reaction	[redacted]	0.091	0.000
Allergic conjunctivitis	[redacted]	0.401	0.188
Infectious conjunctivitis	[redacted]	0.255	0.033
Oral herpes	[redacted]	0.055	0.110
URTI	[redacted]	0.000	0.000

Abbreviations: AE: adverse event; BSC: best supportive care; Q2W: twice a week; QD: once a day; URTI: upper respiratory tract infection.

B.3.3.5 Mortality

All-cause mortality was considered in the cost effectiveness analysis based on the Office for National Statistics National life tables with no adjustment for AD-specific mortality.[81] Age- and gender-specific rates are combined to a blended rate, based on the proportion of men and women in the model and the starting age, as reported in Section B.3.3.1.

B.3.4 *Measurement and valuation of health effects*

B.3.4.1 Health-related quality-of-life studies

An SLR was conducted to identify any relevant HRQoL data for adult patients with moderate-to-severe AD. The original SLR was performed in March 2018 and was updated in February 2020. In total, 23 studies featuring relevant health state utility data associated with the treatment of adult patients with moderate-to-severe AD were identified: 16 from the original SLR, and a further

7 in the update SLR. Full details of the SLR search strategy, study selection process and results are reported in Appendix H.

The SLR yielded no results related to utility data associated with baricitinib treatment of adults with AD. Therefore, the utility values applied in the base case were derived from the EQ-5D-5L data collected in the BREEZE-AD trials.

B.3.4.2 Health-related quality-of-life data from clinical trials and mapping

As described in Section B.2.6, the BREEZE-AD trials assessed HRQoL via the EQ-5D-5L health utilities instrument up to Week 16. For use in the model, health state utility values were derived in line with the NICE reference case: EQ-5D-5L scores collected in the BREEZE-AD trials were cross-walked to EQ-5D-3L scores using the algorithm presented in van Hout *et al*. 2012, and are subsequently used to generate utility index values using the UK value set by Dolan *et al*. 1997.[52, 70, 78] Therefore, these utility index values are reflective of the preferences of a sample representative of the population of interest in UK clinical practice.

B.3.4.3 Adverse reactions

Disutilities associated with adverse events are not included in the model since the AEs observed in the BREEZE-AD trials were mild. Therefore, it is not expected that a significant detriment in QoL would be associated with these events. It is further assumed that QoL decrements due to AEs would be captured in the utility data obtained from the BREEZE-AD trials and thus the exclusion of AE-related disutility from the base case avoids double-counting of this disutility. This approach is in line with the dupilumab submission (TA534).[1]

B.3.4.4 Health-related quality-of-life data used in the cost-effectiveness analysis

Within the cost-effectiveness analyses, health state utility values are derived by cross-walking EQ-5D-5L scores collected in the BREEZE-AD4 (JAIN) and BREEZE-AD7 (JAIY) trials to EQ-5D-3L scores using the algorithm presented in van Hout *et al*. 2012, in line with the NICE reference case.[52, 78] These scores are subsequently used to generate utility index values using the UK value set by Dolan *et al*. 1997.[70]

Patient-level utility index values for the pooled population of JAIN + JAIN-like JAIY patients were used to derive the health state utility values for the base case analysis.[54, 55] All observed values across patients receiving all baricitinib dose groups and placebo were included in the analysis. In order to capitalise on the longitudinal nature of the data collection from the BREEZE-AD programme, a mixed model repeated measurement (MMRM) approach was used to generate the health state utility values from patient-level utility index values, accounting for the parameters presented in Table 89. Statistical significance was set to 0.05; all analyses were run using the SAS software, version 9.4.

Table 89: MMRM model parameters used to generate health state utility values

Parameter	
Dependent variable	EQ-5D score change from baseline to week 16
Factors	Response variable[a]
	Gender

	Visit (reflecting time)
Covariates	Age
	EQ-5D baseline score
Interaction term	Visit-by EQ-5D baseline score

[a]The "response variable" was directly linked to the health state of the cost-effectiveness model, and was dependent on the chosen response definition.
Abbreviations: EQ-5D: EuroQol 5 Dimensions; MMRM: mixed-effect model for repeated measurement.

As described in Section B.3.2.2, in the base case, health states were defined based on achievement of EASI50 with ΔDLQI ≥4. Parameter estimates for the mixed model based on EASI50 with ΔDLQI ≥4 response categories are presented in Table 90. The resulting EQ-5D-3L utility scores are presented in Table 91, including the number of observations included in the analysis.

Table 90: Parameter estimates for the mixed model based on EASI50 with ΔDLQI ≥4 response categories

Fixed effects	Utility model Coefficients	p value
EASI50 with ΔDLQI ≥4: Yes	[redacted]	[redacted]
EASI50 with ΔDLQI ≥4: No	[redacted]	[redacted]
Week 1	[redacted]	[redacted]
Week 2	[redacted]	[redacted]
Week 4	[redacted]	[redacted]
Week 8	[redacted]	[redacted]
Week 12	[redacted]	[redacted]
Sex (Male)	[redacted]	[redacted]
Age	[redacted]	[redacted]
Base	[redacted]	[redacted]
Fit Criteria		
AIC	[redacted]	[redacted]
BIC	[redacted]	[redacted]

Abbreviations: AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index.

Table 91: EQ-5D-3L utility score at baseline and Week 16 by EASI50 with ΔDLQI ≥4 response category at Week 16

EASI50 with ΔDLQI ≥4 (Week 16)	Baseline EQ-5D-3L[a]			Change in EQ-5D-3L (baseline to Week 16)[b]	
	Number of patients	Mean	Standard deviation	LS Mean	95% CIs
Overall	[redacted]	0.5979	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	0.1821	[redacted]
No	[redacted]	[redacted]	[redacted]	0.2042	[redacted]

[a]Observed values. [b]from mixed model. Number of observations used = 2378.
Abbreviations: CI: confidence interval; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; LS: least squares; SD: standard deviation.

In the base case analysis, the baseline utility value was applied in the induction state and the utility value for those with an EASI50 with ΔDLQI ≥4 response at Week 16 is applied in the maintenance state. Non-responders were assigned baseline utility based on advice from clinical experts and in line with the assumptions in the US ICER model. The utility values used in the base case cost-effectiveness analysis are presented in Table 92.

Table 92: Summary of utility values used in the base case cost-effectiveness analysis

State	Utility value (mean)	Reference in submission	Justification
Base case (response: EASI50 with ΔDLQI ≥4)			
Induction (baseline)	0.5979	Section B.3.4.1, Page 144	BREEZE-AD trials
Maintenance	0.7800		
Non-response	0.5979		

Abbreviations: DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index.

Scenario analyses were explored for alternative study populations and where health states were defined using alternative response definitions. Parameter estimates and results of the MMRM models used to generate utility values used in scenario analyses are provided in the reference pack for this submission. The utility values used in scenario analyses are presented in Table 93. In all scenarios, the baseline utility value was applied in the induction and non-response states. Utility values for those achieving response at Week 16 were applied in the maintenance state.

Where possible, the model accounted for conditional response within the maintenance health state, in order to allow for differentiation in efficacy between treatments within the group of responders. For example, if EASI50 was chosen as the response definition for the JAIN + JAIY JAIN-like population, the utility values for the EASI50 to <75 and EASI75 response categories were applied in the maintenance state for the relative proportions of patients achieving EASI50 and EASI75, respectively.

Age adjustment

With increasing age, health utility is expected to decline. Given the base case time horizon of the model, which spans a patient's lifetime, the model base case includes an annual adjustment factor for age derived from UK data from Ara *et al.* (2011), in line with the assumptions made in TA534.[1, 82] Utility values are multiplied by the adjustment factor "Y", which is derived using the following formula based on the age of cohort in a given model cycle:

Y = (1.0708-0.0044*[mean age])/0.901

Table 93: Utility data applied in scenario analyses

Response category (Week 16)		Population				
		JAIN+JAIY JAIN-like	JAHL/JAHM JAIN-like	JAIN (Europe)	JAIN	JAIN+JAIY JAIN-like (secondary censoring)
EASI50 + ΔDLQI ≥4						
Baseline, mean (SD)		0.5979 (0.2776)	[redacted]	[redacted]	[redacted]	[redacted]
CfB at Week 16, LS mean (95%CI)		0.1821 (0.1376, 0.2266)	[redacted]	[redacted]	[redacted]	[redacted]
EASI response status						
Baseline, mean (SD)		0.6203 (0.2789)	[redacted]	[redacted]	[redacted]	[redacted]
CfB at Week 16, LS mean (95%CI)	EASI 50 (to <75)	0.1799 (0.1489, 0.2109)	[redacted]	[redacted]	[redacted]	[redacted]
	EASI 75 to <90	0.2316 (0.2051, 0.2581)[a]	[redacted]	[redacted]	[redacted]	[redacted]
	EASI ≥90		[redacted]	[redacted]	[redacted]	
ΔItch NRS ≥4						
Baseline, mean (SD)		NA	[redacted]	[redacted]	[redacted]	[redacted]
CfB at Week 16, LS mean (95%CI)		NA	[redacted]	[redacted]	[redacted]	[redacted]

[a] EASI ≥75 (EASI75 to <90 combined with EASI ≥90).
Abbreviations: CfB: change from baseline; CI: confidence interval; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; LS: least squares; NA: Not available; NRS: numerical rating scale; SD: standard deviation.

B.3.5 *Cost and healthcare resource use identification, measurement and valuation*

An SLR was conducted to identify any relevant cost or resource use data for adult patients with moderate-to-severe AD. The original SLR was performed in March 2018 and was updated in February 2020. In total, 4 studies featuring relevant cost and resource use data associated with the treatment of adult patients with moderate-to-severe AD were identified: 3 from the original SLR, and a further 1 in the update SLR. Full details of the SLR search strategy, study selection process and results are reported in Appendix I.

The following cost categories are included in the model:

- Drug acquisition costs (Section B.3.5.1)
- Administration costs (Section B.3.5.1)
- Treatment initiation and monitoring resource use (Section B.3.5.1)
- AEs (Section B.3.5.3)

The economic analysis was conducted from an NHS and PSS perspective and therefore included only costs that would be incurred by the NHS and PSS. Cost inputs are based on British National Formulary,[5] Monthly Index of Medical Specialities (MIMS),[83] National Health Service Reference costs (2018–2019),[84] and Personal Social Services Research Unit (PSSRU).[85]

B.3.5.1 Intervention and comparator costs and resource use

Drug acquisition and administration costs

For drug acquisition costs for treatments, presented in Table 94, the dose and frequency are based on approved doses obtained from Summary of Product Characteristics (SmPC). Drug acquisition costs have been calculated based on the cost per unit of each treatment and the required number of units per cycle. The administration cost associated with dupilumab has been derived from the dupilumab submission (TA534) since it was validated by clinicians and accepted by NICE.[1] These costs are used for analyses in the model.

BSC consists of a mixture of treatment modalities. The composition and acquisition costs for BSC are based on expert advice and are presented in Table 95. For each treatment option, the cost per administration was calculated, based on the dose per application, frequency and the pack cost of the respective treatment. Using the proportion of use of each treatment option, a weekly cost of £14.73 for BSC was calculated and used for analyses in the model.

Table 94: Drug acquisition and administration costs for treatments

Treatment	Pack cost, £	Number of units per pack	Dose per unit, mg	Cost per unit, £	Number of doses administered		Subcutaneous injection training, unit (cost)	
					Induction period (16 weeks)	Maintenance (annual)	Induction period	Maintenance
Baricitinib	[redacted]	28	4 or 2	[redacted]	112	365	0	0
Dupilumab	1,264.89	2	300	632.45	10	26	1 (£56.50)[a]	0

[a] One injection training session at the start of treatment as presented in NICE (TA534).[1]

Table 95: Drug acquisition costs for BSC

Treatment	Pack cost, £	Number of units per pack	Dose per unit	Dose per application	Administration frequency (number of doses administered)		Proportion of use, %
					Induction period (16 weeks)	Maintenance (annual)	
BSC							100
Mometasone (class II TCS)	9.50	100g	0.10%	32 g	Daily, 112	Daily, 365	66.70
TCI (Tacrolimus)	47.28	60g	0.10%	1.75 g	Twice per week, 28	Twice per week, 104	22.20
Oral corticosteroids (Prednisolone)	1.48	28	30 mg	10 mg	Daily, 4.31	One course, 14	5.00
Weighted cost of BSC per week: £14.73							

Abbreviations: BSC: best supportive care; N/A: not applicable; TCI: topical calcineurin inhibitor TCS: topical corticosteroids.
Sources: MIMS,[83] UK Medical advisory board (April 2019).

Costs of concomitant medication

The model includes the weekly cost of concomitant medication, consisting of bathing and emollient products, mid-potency background TCS (mometasone 0.1% ointment) and TCI (Protopic 0.1% ointment, tacrolimus). These costs are presented in Table 96.

The weekly cost of bathing products and emollients used in the model was derived by averaging the weekly cost of each of the treatments within each of these categories. The costs for responders are applied to responders receiving baricitinib, dupilumab or BSC. The non-responder costs correspond to the non-responder health states as presented in Section B.3.5.2.

Health care resource use was obtained from the dupilumab submission (TA534) where it was assumed that responders had a 50% reduction of resource use in bathing products and emollients compared to non-responders.[1] For TCS, the resource use has been derived from the dupilumab submission (TA534) as no long term data is available for baricitinib. This assumption has been confirmed by clinical experts. For TCI use, the resource use has been derived from the dupilumab submission (TA534), where clinical experts concluded that for facial involvement, TCIs are more appropriate then steroid treatments, and Protopic 0.1% ointment (tacrolimus) is the preferred option. The clinical experts also concluded that responders to treatment would not require TCI treatment.[1] Given this clinical validation and approval by NICE, these were identified as the most relevant sources of cost and resource use associated with concomitant medication for the treatment of AD.

Table 96: Costs of bathing products and emollients used in the model

Medication	Pack costs	Pack size	Proportion of product prescribed	Amount per week (non-responders)	Weekly costs (non-responders)	Weekly costs (responders)[a]	Resource use (induction, 16 weeks)	Resource use (maintenance, annual)
Bathing products					**£4.26**	**£2.13**	16.0	52.0
Aqueous cream	£4.70	500mg	33%	Assume 1 pack per week	£4.70	£2.35		
Dermol 200 Shower Emollient	£3.55	200ml	25%	Use as a soap substitute, assumed 1 pack per week	£3.55	£1.78		
Aveeno Bath Oil	£7.29	300ml	17%	30ml per bath, assumed daily	£7.29	£3.65		
Dermol 600 Bath Emollient	£7.55	600ml	15%	30ml per bath, assumed daily (210ml/week, 45% of 600ml)	£3.40	£1.70		
Oilatum Bath Formulation	£5.02	300ml	10%	140ml (20ml per bath)	£2.36	£1.18		
Emollients					**£5.24**	**£2.22**	16.0	52.0
Aveeno cream	£6.47	500ml	-	1	£6.47	£3.24		
Cetraben ointment	£5.39	450g	-	1	£5.39	£2.70		
Dermol cream	£6.63	500g	-	1	£6.63	£3.32		
Diprobase ointment	£5.99	500g	-	1	£5.99	£3.00		
Epaderm ointment	£12.25	1000g	-	0.5	£6.13	£1.53		
Hydromol ointment	£8.20	1000g	-	0.5	£4.10	£1.03		
White soft paraffin 50%/ liquid paraffin 50% ointment	£4.57	500g	-	1	£4.57	£2.29		
Oilatum cream	£5.28	500ml	-	0.5	£2.64	£0.66		
TCS								
Mometasone 0.1% ointment	£9.50	100g	-	112.04	£10.64	£5.39[b]	16	52
TCI								
Protopic 0.1% ointment, tacrolimus	£47.28	60g	-	1.75	£1.38	£0.00[c]	16	52

[a] Assuming 50% reduction from non-responder. [b] Assuming usage of 56.70g per week in responders. [c] Assuming no usage of TCIs in responders
Sources: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534);[1] MIMS.[83]

Costs of medications to treat flares

Due to the nature of AD, rescue medication following a flare is often required. Flare was not an endpoint in the baricitinib studies, but the receipt of rescue medication can be considered a proxy for flare. Weekly costs, updated to the current cost year, the proportions of treatments of flares and annual flare rate were all assumed to be the same as those presented in the dupilumab NICE submission since these are the most plausible estimates and since long term data for baricitinib is pending.[1] The proportion of treatment of flares and annual flare rate are presented in Table 97, and the estimates for flare cost treatment are presented in Table 98.

Table 97: Proportion of treatment of flare and annual flare rate

	Baricitinib	Dupilumab	BSC
Proportion TCS (potent) at 52 weeks	0.42	0.42	0.54
Proportion TCS (very potent) at 52 weeks	0.23	0.23	0.27
Proportion systemic steroids at 52 weeks	0.29	0.29	0.13
Proportion TCI at 52 weeks	0.00	0.00	0.06
Annual flare rate	0.18	0.18	0.78

Abbreviations: BSC: best supportive care; TCI: topical calcineurin inhibitor; TCS: topical corticosteroids.
Source: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534).[1]

Table 98: Flare treatment cost

Treatment class	Product	Pack cost	Pack size	Resource use assumptions	Cost
TCS (potent)	Betamethasone valerate cream	£7.83	100	1 pack	£21.51
	Cutivate 0.05% cream	£4.24	30	3 and 1/3 packages	
TCS (very potent)	Eumovate 0.05% ointment	£5.44	100	1 pack	£13.34
	Dermovate 0.05%cream	£7.90	100	1 pack	
Systemic steroid	Predisolone 5mg	£1.48	28	1 pack	£1.48
TCI	Protopic 0.1% ointment, tacrolimus	£47.28	60	5.7g/dose every 3 days over 4 weeks	£19.02

Abbreviations: TCI: topical calcineurin inhibitor; TCS: topical corticosteroids.
Sources: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534),[1] MIMS.[83]

Monitoring costs and resource use

In the model, patients incur monitoring costs and resource use that differ depending on treatment and health state. Differences in modes of action and administration routes of the included treatments may necessitate differences in monitoring practices and resource use. The resource use is aligned with the dupilumab submission (TA534) which was accepted as relevant by the appraisal committee and validated by clinical expert opinion.[1]

Frequency of resource use is stratified by Induction (reflecting the induction period resource use), and Maintenance (annual frequency for responders). This is to ensure that the right frequency of visits or tests is captured in the appropriate period in the model. Typically, there are different requirements or frequencies during the induction period (or treatment initiation) of a treatment compared to the maintenance period of the same treatments.

Health care resource unit costs used in the model were sourced from the National Health Service Reference costs and Personal Social Services Research Unit (PSSRU) and are presented in Table 99.[84, 85]

Monitoring resource use for responders was derived from the dupilumab NICE submission (TA534) and are presented in Table 100.[1] BSC is associated with additional monitoring costs which are presented in Section B.3.5.2.

Table 99: Health care resource unit costs used in the model

Health care resource	Unit cost	Source
Dermatologist outpatient consultation (consultant led)	£114.57	NHS Reference Costs (2018–19), weighted average of WF01A–D and WF02A–C
Dermatologist nurse visit	£10.50	PSSRU
GP consultation	£39.00	PSSRU
Accident & Emergency visit	£182.58	NHS Reference Costs (2018–19), weighted average of VB06Z–VB09Z
Hospitalisation	£1,854.72	Weighted average presented in TA534 (£1,795 in the 2018 cost year) adjusted for inflation to 2020 cost year.
Day case	£433.69	NHS Reference Costs (2018–19), weighted average of JD07A–JD07K
Full blood count (FBC)	£3.00	NHS Reference Costs (2018–2019) DAPS05
Phototherapy	£103.00	NHS Reference Costs (2018–2019) JC47Z
Psychological support	£289.46	NHS Reference Costs (2018–19), weighted average of WF01A–D

Sources: PSSRU,[85] NHS Reference Costs.[84]

Table 100: Administration and monitoring health care resource use for responders

Health care resource	Baricitinib		Dupilumab		BSC		Assumption
	Induction (16 weeks)	Maintenance (annual)	Induction (16 weeks)	Maintenance (annual)	Induction (16 weeks)	Maintenance (annual)	
Dermatologist outpatient consultation (consultant led)	2.00	4.30	2.00	4.30	2.00	4.30	One visit at treatment start and response evaluation during induction, thereafter resource use is assumed to be the same as in TA534[1]
Dermatologist nurse visit	0.11	0.35	0.11	0.35	0.11	0.35	Annual resource used assumed to be the same as in TA534,[1] induction resource use calculated from annual resource use
GP consultation	1.91	6.20	1.91	6.20	1.91	6.20	
Accident & Emergency visit	0.01	0.02	0.01	0.02	0.01	0.02	
Hospitalisation	0.01	0.02	0.01	0.02	0.01	0.02	
Day case	0.00	0.00	0.00	0.00	0.00	0.00	
Full blood count (FBC)	0.00	0.00	0.00	0.00	1.23	4.00	

Source: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534).[1]

B.3.5.2 Health-state unit costs and resource use

In the BSC Non-Responder health state, patients receive treatment. In the base case, when patients have failed active treatment (baricitinib or dupilumab), their options are limited to emollients, low-to-mid potency TCS, phototherapy, psychological support and rescue therapy including higher potency TC, oral corticosteroids or TCIs. In the dupilumab NICE submission, data from a retrospective database analysis, care notes review, and clinical expert opinion were used to inform how patients who were non-responders to treatment were treating their condition. These data have been applied in this model and are presented in Table 101.

Table 101: BSC non-responder resource use and costs

Health care resource	Maintenance (annual resource use)	Assumptions (see for Table 99 sources)
Physician visits		
Dermatologist outpatient consultation (consultant led)	6.0	Assumed to be the same as in the dupilumab submission (TA534)[1]
Dermatologist nurse visit	0.46	
GP consultation	12.8	
Hospital costs		
A&E visit	0.08	Assumed to be the same as in the dupilumab submission (TA534)[1]
Hospitalisation	0.13	
Day case	0.2	
Tests and investigations		
Full blood count	4.0	Assumed to be the same as in the dupilumab submission (TA534)[1]
Concomitant medication		
All bathing products, emollients, TCS and TCI	52.0	See Table 96 for full details
Other		
Phototherapy	6%	Assumed to be the same as in the dupilumab re-submission (TA534)[1]
Psychological support	7%	

Sources: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534).[1]

B.3.5.3 Adverse reaction unit costs and resource use

The costs associated with AEs are summarised in Table 102.

Table 102: Adverse reaction unit costs

Adverse reactions	Cost	Source
Injection site reaction	£112.12	NHS Reference Costs (2018–2019), consultant-led dermatologist visit (WF01A)
Allergic conjunctivitis	£39.00	GP consultation (PSSRU)
Infectious conjunctivitis	£55.15	Ophthalmologist consultation: assumed to be £101.46, derived as weighted average of NHS Reference Costs (2018–19) WF01A–D and WF02A–C. Infectious conjunctivitis: weighted average of ophthalmologist consultation (20%) and GP consultation (£39.00, PSSRU;

		80%) with unit cost of 1% prednisolone eye drops (£3.66, MIMS)
Oral herpes	£52.96	GP consultation (PSSRU) with Zovirax 5% cream, 10mg (MIMS)
URTIs	£39.00	Unit cost of GP consultation (PSSRU)

Abbreviations: GP: general practitioner; MIMS: Monthly Index of Medical Specialties; PSSRU: Personal Social Services Research Unit; URTI: upper respiratory tract infection.
Sources: MIMS,[83] National Health Service Reference costs[84] and PSSRU.[85]

B.3.5.4 Miscellaneous unit costs and resource use

There are no further unit costs or resource use included in the model.

B.3.6 *Summary of base-case analysis inputs and assumptions*

B.3.6.1 Summary of base-case analysis inputs

A summary of the variables applied in the model in the base case analysis is provided in Table 103.

Table 103: Summary of variables applied in the economic model base case

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
Model properties			
Start age, years	37.27	• DSA: Varied ±20% of mean • PSA: Normal; standard error assumed to be 10% of the mean	Section B.3.3.1
Proportion male (%)	63.62	• DSA: Varied ±20% of mean • PSA: Beta; standard error assumed to be 10% of the mean	Section B.3.3.1
Discount rate costs (%)	3.5	• DSA: Varied to {0, 5} • PSA: Not varied	Section B.3.2.2
Discount rate benefits (%)	3.5	• DSA: Varied to {0, 5} • PSA: Not varied	Section B.3.2.2
Time horizon	Lifetime	None	Section B.3.2.2
Perspective	NHS/PSS	None	Section B.3.2.2
Include mortality	Yes	None	Section B.3.3.5
Age-adjusted utility	Yes	None	Section B.3.2.2
Definition of response	EASI50 with ΔDLQI≥4	None	B.3.2.2
Efficacy			

Response	ITC	• DSA: Varied ±20% of mean • PSA: Beta; standard error from ITC	Section B.2.9
Utilities			
Health state utilities	Baricitinib trial data	• DSA: Varied ±20% of mean • PSA: Beta; standard error assumed to be 10% of the mean	Section B.3.4.4
Costs			
Acquisition cost: baricitinib and dupilumab	Baricitinib: anticipated dosing schedule and price supplied by Lilly Dupilumab: dose and dosing schedule taken from SmPC. List price taken from MIMS.	• DSA: Varied ±20% of mean • PSA: Not varied	Section B.3.5.1
Administration and monitoring	Unit costs of resources are taken from NHS reference costs (2018/2019) and PSSRU (2018/19). Resource use was based on the dupilumab submission (TA534).	• DSA: Varied ±20% of mean • PSA: Gamma; standard error assumed to be 10% of the mean	Section B.3.5.1
Concomitant medication	Costs and resource use are based on the dupilumab submission (TA534).	• DSA: Varied ±20% of mean • PSA: Gamma; standard error assumed to be 10% of the mean	Section B.3.5.1
Flare treatment	Costs and resource use are based on the dupilumab submission (TA534).	• DSA: Varied ±20% of mean • PSA: Gamma; standard error assumed to be 10% of the mean	Section B.3.5.1
Disease management for non-responders	Costs and resource use are based on the dupilumab submission (TA534).	• DSA: Varied ±20% of mean • PSA: Gamma; standard error assumed to be 10% of the mean	Section B.3.5.2
AEs	Costs are sourced from MIMS, the PSSRU and NHS reference costs.	• DSA: Varied ±20% of mean • PSA: Gamma; standard error assumed to be 10% of the mean	Section B.3.5.3

Abbreviations: AE: adverse event; DLQI: Dermatology Life Quality Index; DSA: deterministic one-way sensitivity analysis; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; ICER: Institute for Clinical and Economic Review; ITC: indirect treatment comparison; MIMS: Monthly Index of Medical Specialties; NA: not applicable; PSA: probabilistic sensitivity analysis.

B.3.6.2 Assumptions

A list of the assumptions made in the base case analysis and their justifications is provided in Table 104. Where appropriate, the exploration of the potential impact of these assumptions in a scenario analysis is noted.

Table 104: List of assumptions for the base case analysis

Model input	Description of assumption for the base case	Justification
Response definition	Based on EASI50 and an improvement of DLQI of at least 4 at Week 16	Effective proxy method for capturing sufficient clinical benefit to justify continuing treatment after the trial period. This response definition was the implemented in the base case of the dupilumab submission (TA534) and accepted as clinically relevant in NHS practice by the appraisal committee.[1] The use of EASI75 as the definition of response is explored in a scenario analysis.
Response maintenance	Patients maintain the response they experienced at the end of the trial period until they discontinue treatment	Loss of response is assumed to result in treatment discontinuation and is reflected in the all-cause discontinuation rate as modelled in the final base case of the dupilumab submission (TA534)[1]
Conditional response rates and long-term drop-out rates	Assumed to be the same as presented in the dupilumab submission (TA534)[1]	Assumption made due to lack of data from the baricitinib trial programmes at the time of analysis. This assumption is difficult to test in the absence of comparative trial data, but may be conservative given that baricitinib, as a small molecule, will not result in anti-drug antibody development.
Discontinuation rate	Assumed that patients who responded to treatment at the initial response assessment but not at Week 52 would discontinue treatment at a continuous and constant rate between Week 16 and Week 52	It was considered that the assumption that treatment discontinuation is a random process, which would be reasonably approximated by a constant and continuous rate, was plausible.
	Assumed that patients lose response from BSC at more a rapid rate than patients discontinue baricitinib and dupilumab after one year.	Assumption is in line with the dupilumab submission (TA534)[1] and reflects that adherence to the topical regimen is unlikely to be maintained outside a trial setting.
	Assumed that after Week 52, patients discontinue treatments at the same rate for dupilumab and baricitinib, and this rate is assumed to be the same as observed in open-label extension studies[1]	Assumption is in line with the dupilumab submission (TA534)[1] and appears reasonable give the lack of direct or indirect trial data or real-world evidence to inform the model on these rates.
BSC non-responders	Assumed that patients in the BSC “non-response” state remain in that state until the end of model simulation or death.	Reflects that patients who have received BSC have failed all other treatment options or have a contraindication or intolerance to them. Therefore, it is considered reasonable to expect that these patients would not benefit substantially from treatment in clinical practice.
AE risk	Assumed to remain constant over the treatment duration	Assumption appears reasonable given the lack of long-term data from the baricitinib trial programmes at the time of analysis.

Health state utility values over time	Assumed to decline with age, with the model applying an age adjustment factor derived from Ara and Brazier, 2011, to account for this.[82]	Assumption is based on a well-established UK literature source preferred by NICE in many previous appraisals
Resource use	Assumed that baricitinib resource use is the same as in the dupilumab submission (TA534)[1]	Assumption appears reasonable given the lack of information on resource use of baricitinib.
	Assumed that resource use for concomitant treatment is decreased by 50% in responders as compared with non-responders	In line with the dupilumab submission (TA534)[1]
	BSC resource use estimates presented in the dupilumab submission (TA534)[1] are applicable to the patient population eligible for baricitinib	The data presented in the dupilumab submission (TA534)[1] were the best resource use data identified in a systematic review on UK resource use for moderate-to-severe AD and should therefore the best available evidence for the relevant population.
Treatment adherence	Adherence to treatment was not modelled separately	Reflects that compliance rates are high in both dupilumab and baricitinib trials. It is assumed that effectiveness and costs would decrease proportionally with lower compliance, thus limiting the impact on the ICER of changes to compliance.
Dupilumab administration	Assumed the patients self-administer dupilumab after an initial training session with a nurse.	In line with the dupilumab submission (TA534)[1]
Efficacy	Assumed to occur at the end of the Induction period, so only patients who enter the maintenance phase benefit from treatment.	Considered to be a conservative assumption given that novel systemics incur all costs of treatment, but not the benefits of treatment during the trial period, and the impact of this assumption is expected to be limited given the short duration of the trial period (16 weeks).

Abbreviations: AD: atopic dermatitis; AE: adverse event; BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; ICER: incremental cost-effectiveness ratio.

B.3.7 *Base-case results*

B.3.7.1 Base-case incremental cost-effectiveness analysis results

The summary of results in the base case analysis are presented in Table 105.

BSC, baricitinib 4 mg and dupilumab Q2W accumulated total costs of £[redacted], £[redacted] and £[redacted], respectively, and accumulated [redacted], [redacted], and [redacted] total QALYs, respectively. At the confidential PAS price, all ICERs in the base case population of patients who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control, were within the range considered cost-effective. In pairwise comparison of baricitinib versus BSC, the ICER was estimated at £17,941/QALY which falls below the NICE willingness-to-pay (WTP) threshold of £20,000. Versus dupilumab, baricitinib was cost-effective in the South-West quadrant, accruing considerably fewer costs and slightly fewer QALYs (ICER: 203,525/QALY foregone). The probability of cost-effectiveness at WTP thresholds of £20,000 and £30,000 is presented in Table 106 at which baricitinib had a cost-effectiveness probability of [redacted]% and [redacted]%, respectively. Net monetary benefit (NMB) as a function of willingness-to-pay is presented in Figure 41. These results demonstrate baricitinib to be a cost-effective option for the treatment of moderate-to-severe AD in the target population versus the two comparators relevant to UK clinical practice.

The deterministic sensitivity analyses (DSA) and probabilistic sensitivity analyses (PSA) conducted to test the robustness of the model to the uncertainties within the model parameters are presented in Sections B.3.8.1 and B.3.8.2, respectively. The scenario analyses undertaken to explore the uncertainty around model assumptions are presented in B.3.8.3. The clinical outcomes and disaggregated base case cost-effectiveness results are presented in Appendix J.1.

Table 105: Base case results (JAIN and JAIN-like JAIY)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,941	£17,941
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,525[b]

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LY: life years; QALYs: quality-adjusted life years.

Table 106: Probability of cost-effectiveness at a WTP threshold of £20,000 and £30,000

	WTP threshold £20,000	WTP threshold £30,000
BSC	[redacted]	[redacted]
Baricitinib	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]

Abbreviations: BSC: best supportive care; WTP: willingness-to-pay.

Figure 41: NMB as a function of WTP per incremental QALY

Abbreviations: BSC: best supportive care; NMB: net monetary benefit; QALY: quality-adjusted life year; WTP: willingness-to-pay.

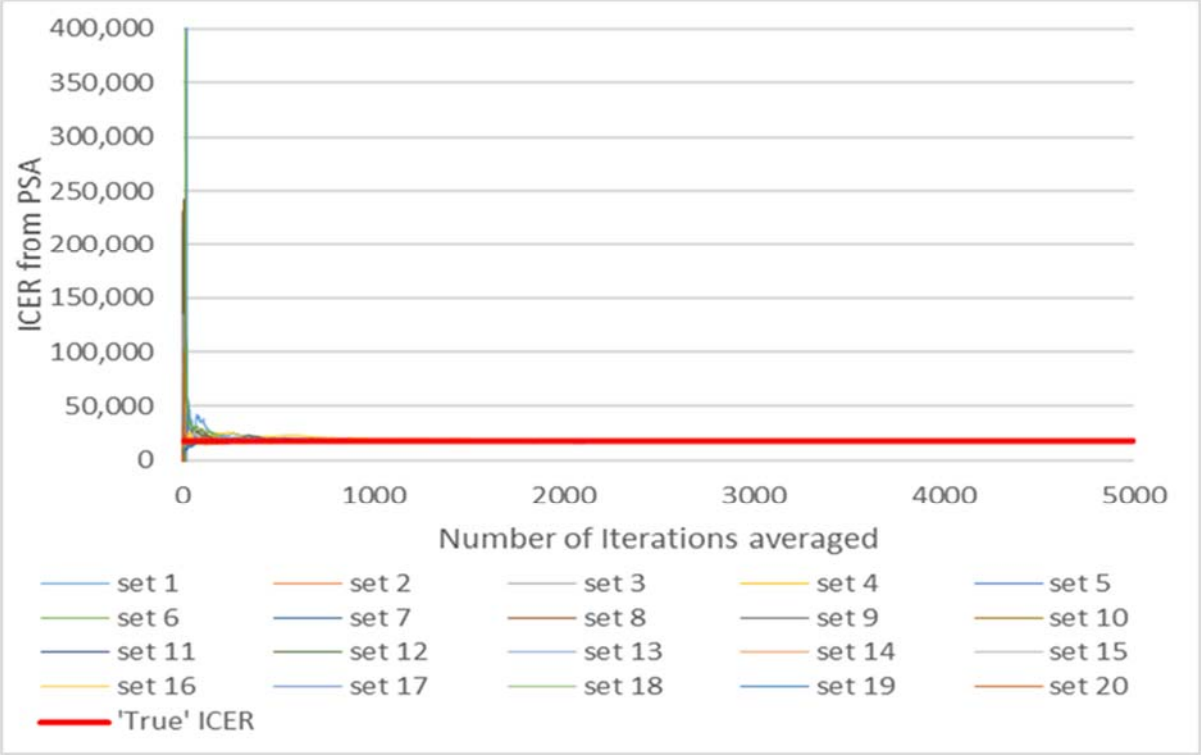
B.3.8 *Sensitivity analyses*

B.3.8.1 Probabilistic sensitivity analysis (PSA)

Probabilistic sensitivity analyses (PSAs) with 3,000 iterations were performed for each pairwise and fully incremental comparison in order to assess the uncertainty associated with model input parameters. 3,000 iterations was deemed appropriate based on the results of an ICER convergence tests, shown in Figure 42 and Figure 43 for baricitinib versus BSC and dupilumab, respectively. This testing was performed with repeated block sampling of 20 blocks of 5,000 iterations each for the model and each line in these graphs presents the change in the ICER in each block as the sample size of the block contributing to the average cost and QALY totals is increased.[86] The input parameters and distributions associated with each parameter are presented in Appendix J.2.

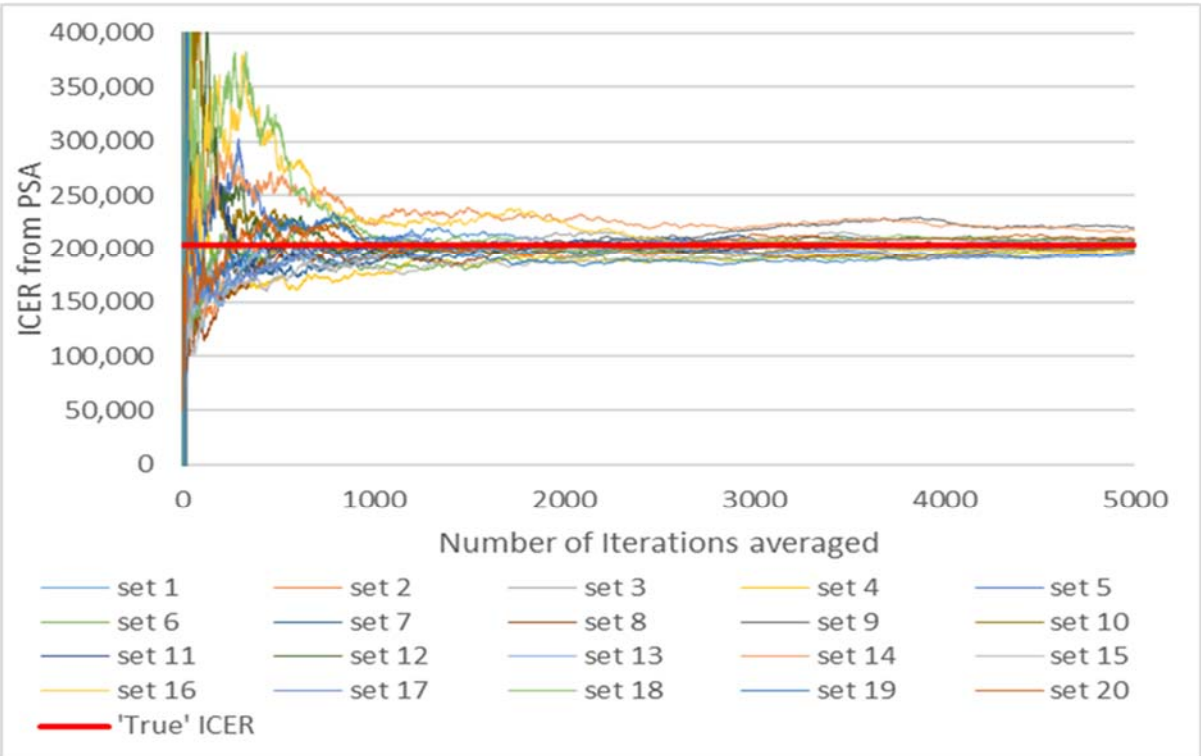
The probabilistic base case results are presented in Table 107 and the cost-effectiveness plane scatterplot and cost-effectiveness acceptability curve are presented in Figure 44 and Figure 45, respectively. Baricitinib has a higher probability of being cost-effective than both dupilumab and BSC at a willingness to pay (WTP) threshold of £20,000/QALY gained over the range of values tested in the model.

Figure 42: Probabilistic ICER convergence plot for baricitinib–BSC versus BSC



Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; PSA: probabilistic sensitivity analysis.

Figure 43: Probabilistic ICER convergence plot for baricitinib–BSC versus dupilumab–BSC



Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; PSA: probabilistic sensitivity analysis.

Table 107: Probabilistic base case results

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,853	£17,853
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£87,866	£199,001[b]

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LY: life years; QALYs: quality-adjusted life years.

Figure 44: Cost-effectiveness plane scatterplot

Generated using 3,000 iterations of the PSA.

Figure 45: Cost-effectiveness acceptability curve

Generated using 3,000 iterations of the PSA.

B.3.8.2 Deterministic sensitivity analysis (DSA)

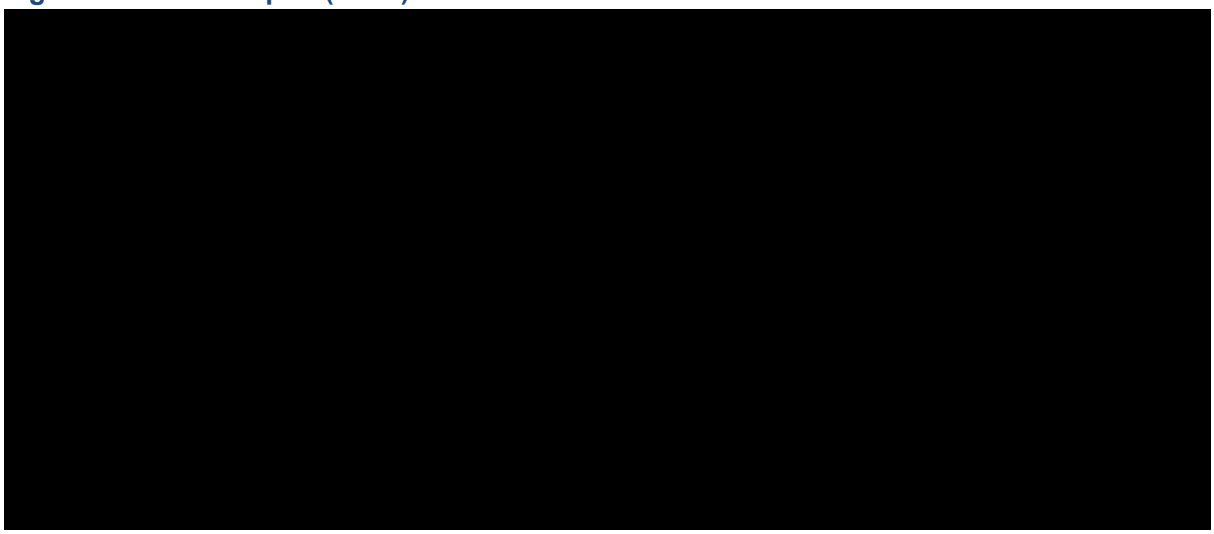
The input parameters and distributions associated with each input parameter in the DSA are presented in Appendix J.2.

The ten most influential variables in the DSA for the analysis of baricitinib versus dupilumab and baricitinib versus BSC are presented as tornado plots in Figure 46 and Figure 47, respectively. For the comparison of baricitinib versus dupilumab, the discount rate for costs had the largest impact on the ICER, with the efficacy value for the composite outcome of EASI50 + ≥4-point improvement in DLQI, the discount rate for utilities and the dupilumab pack cost also proving influential. For the comparison of baricitinib versus BSC, the discount rates for utilities and costs had the largest impact on the ICER with the EASI50 health state utility value also proving influential.

Figure 46: Tornado plot (ICER) of baricitinib-BSC versus dupilumab-BSC

Abbreviations: BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; HSUV: health state utility value; ICER: incremental cost-effectiveness ratio.

Figure 47: Tornado plot (ICER) of baricitinib-BSC versus BSC



Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; GP: general practitioner; HSUV: health state utility value; ICER: incremental cost-effectiveness ratio; TCS: topical corticosteroids.

B.3.8.3 Scenario analysis

A number of scenario analyses were explored in which model assumptions or parameters were altered. The scenario analyses carried out are presented in Table 108. The results of these scenario analyses are presented below in Table 109.

Table 108: Summary of scenario analyses

#	Scenario analysis value	Base case value	Rationale
1	Response definition: EASI75	Response definition: EASI50 with ΔDLQI≥4	To explore the impact of alternative definitions of response
2	**Censoring rule used for clinical efficacy data**		
2a	Secondary censoring of clinical efficacy data from the BREEZE-AD trials (Response definition: EASI50 with ΔDLQI≥4)	Primary censoring of clinical efficacy data from the BREEZE-AD trials (*Response definition*: EASI50 with ΔDLQI≥4)	To explore the impact of non-responder imputation modelling in interpretation according to two response definitions
2b	Secondary censoring of clinical efficacy data from the BREEZE-AD trials (Response definition: EASI75)		
3	**Population analysed**		
3a	JAIN **versus** CAFÉ (combination therapy) (*Response definition*: EASI75)	JAIN + JAIN-like JAIY **versus** CAFÉ + CAFÉ-like CHRONOS (combination therapy) (*Response definition*: EASI50 with ΔDLQI≥4)	To explore the impact of alternative populations analysed according to various response definitions
3b	JAIN **versus** CAFÉ (combination therapy) (*Response definition*: Itch NRS≥4 at Week 16)		
3c	EU population of JAIN **versus** CAFÉ (combination therapy) (*Response definition*: EASI75)		
3d	EU population of JAIN **versus** CAFÉ (combination therapy) (*Response definition*: Itch NRS≥4 at Week 16)		
3e	JAIN-like JAHL + JAIN-like JAHM **versus** CAFÉ-like SOLO (monotherapy) (*Response definition*: EASI50 with ΔDLQI≥4)		
3f	JAIN-like JAHL + JAIN-like JAHM **versus** CAFÉ-like SOLO (monotherapy) (*Response definition*: EASI75)		

Abbreviations: DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EU: European Union; NRS: numerical rating scale.

Table 109: Scenario analyses results

Scenario	Treatment	Total costs (£)	Total QALYs	Total LYs	Incremental costs (£)	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
Base case	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£17,941	£17,941
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,525[b]
1	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14,283	£14,283
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£70,873	£193,541[b]
2a	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£13,736	£13,736
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£68,392	£192,238[b]
2b	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£11,543	£11,543
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£57,463	£171,965[b]
3a	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,038	£16,038
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£69,692	£136,649[b]
3b	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£15,569	£15,569
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£79,712	£2,345,212[b]
3c	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,849	£16,849
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£70,941	£147,480[b]
	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-

3d	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,109	£16,109
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£83,520	Baricitinib dominant
3e	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£47,146	£47,146
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£234,222	£586,761[b]
3f	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£19,433	£19,433
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£95,477	£239,988[b]

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LY: life years; QALYs: quality-adjusted life years
.

B.3.8.4 Summary of sensitivity analyses results

Results of the sensitivity analyses demonstrate that the base case cost-effectiveness results exhibit little variation when the combined distributional uncertainty across model parameters is taken into account. The PSA results aligned closely with the deterministic base case results showing that baricitinib is cost-effective versus both dupilumab and BSC and indicating it to be a cost-effective use of resources in the NHS. As demonstrated by the DSA, the most influential parameters driving the model for the comparison of baricitinib with dupilumab were the discount rate for costs, the efficacy value for the composite outcome of EASI50 + ≥4-point improvement in DLQI, the discount rate for utilities and the dupilumab pack cost (which is confidential an therefore unknown); for the comparison of baricitinib versus BSC, the discount rates for utilities and costs and the EASI50 health state utility value were the most influential parameters.

Limited variation was observed in the majority of changes to the modelling approach that were explored in the scenario analyses. Across all of the scenarios conducted except one, baricitinib was associated with ICERs versus BSC of less than £30,000 per QALY gained; versus dupilumab, baricitinib was associated with a more than £30,000 saving per QALY forgone across all scenarios conducted. Altogether, these results demonstrate the robustness of the model to uncertainty.

B.3.9 *Subgroup analysis*

No further subgroup analyses were performed beyond those described above.

B.3.10 *Validation*

The model methodology was designed to align with NICE's preferred methods. The model was built to align with the NICE reference case, and used an NHS and PSS perspective and discount rates for cost and benefits of 3.5%.[52] The model used a lifetime time horizon in order to capture all costs and QALY gains associated with the interventions. In line with the NICE reference case, the EQ-5D-5L scores collected in the BREEZE-AD trials were cross-walked to EQ-5D-3L scores in line with the NICE reference case and subsequently used to generate utility index values using the UK value set by Dolan *et al*. 1997.[52, 70] The model structure is closely aligned with the model used in the dupilumab NICE submission (TA534) for the assessment of the cost-effectiveness of dupilumab in moderate-to-severe AD.[1]

B.3.10.1 Validation of cost-effectiveness analysis

Clinical validity

Given the current COVID-19 pandemic, it was not possible to pursue external confirmation of the clinical validity of the model structure and assumptions.

Internal model validity

Quality-control (QC) procedures for verification of input data and coding were performed and two checklists (for technical and stress test checks) were used to ensure that the model generated accurate results which were consistent with input data and robust to extreme values. An independent reviewer who was not involved in model development performed the technical and stress test QC checks, and the complete checklists are documented in Appendix N. As part of the technical QC, all model calculations were reviewed, including standalone formulae, equations

and Excel macros programmed in VBA. The correct functioning of the sensitivity and scenario analyses was also reviewed. The stress test ensured that the expected effect is observed when key inputs are varied in the model (e.g. when utilities for all health states and for AEs are set to 0, all QALYs should result equal to 0).

B.3.11 *Interpretation and conclusions of economic evidence*

Summary of cost-effectiveness evidence

The cost-effectiveness of baricitinib in AD was evaluated versus dupilumab and BSC, the most clinically relevant comparators for this population. In the deterministic base case, baricitinib was cost-effective in pairwise comparisons versus dupilumab and versus BSC, which are the most clinically relevant comparators for this population. The pairwise ICER for baricitinib versus BSC fell into the north-east quadrant of the cost-effectiveness plane, demonstrating an incremental QALY of [redacted], incremental costs of [redacted] and an ICER of £17,941 per QALY gained. The pairwise ICER for baricitinib versus dupilumab fell into the south-west quadrant of the cost-effectiveness plane, demonstrating that baricitinib accumulated less QALYs but also less costs compared to dupilumab). ICERs falling into this quadrant that are greater than the £20,000–£30,000 per QALY threshold may be deemed cost-effective: baricitinib versus dupilumab demonstrated an incremental QALY of [redacted], incremental costs of [redacted] and an ICER of £203,525 saved per QALY forgone.

The results of the PSA and DSA were closely aligned with the deterministic base case results, with baricitinib remaining cost-effective versus both comparators. Across scenario analyses where the definition of efficacy, the censoring rule for clinical efficacy data and the population analysed were varied, baricitinib remained cost-effective versus both comparators in all but one scenario explored. These sensitivity results demonstrate the robustness of the model to uncertainty.

Overall, the results indicate baricitinib to be a cost-effective option for the treatment of AD within the NHS versus dupilumab and BSC.

Strengths

The cost-effectiveness model developed for this submission has a number of strengths. Firstly, the model aligns with the cost-effectiveness model used in the dupilumab NICE submission (TA534) which was deemed suitable for decision making concerning the cost-effectiveness of dupilumab in moderate-to-severe AD, and incorporates key criticisms and committee preferences from TA534.[1] The efficacy of baricitinib is based largely on robust Phase III trial data derived from a population which closely aligns with the treatment pathway, and the efficacy of dupilumab is based on an ITC versus the large RCTs CAFÉ and CHRONOS with placebo as a common comparator. The model was built to align with the NICE reference case, adopting an NHS and PSS perspective, a lifetime time horizon to capture fully all costs and QALY gains associated with the interventions, and discount rates for costs and benefits of 3.5%. Finally, BREEZE-AD trial data were cross-walked from the EQ-5D-5L to the 3L, in line with the NICE position statement.

Limitations

The key limitations associated with the analysis are due to the absence of head-to-head trial data between baricitinib and dupilumab, necessitating an ITC to inform relative effectiveness in the

model, and the lack of long-term data for baricitinib. The use of indirect comparison techniques inherently results in a greater level of parameter uncertainty in the relative effectiveness estimates than head-to-head trial data. Limited data were available to inform modelling of long-term efficacy, discontinuation rates and resource use for baricitinib in the population of interest, and thus these parameters were considered to be equivalent to dupilumab within the model. Limitations were therefore addressed by use of conservative assumptions regarding response rate as well as extensive scenario analysis.

Conclusion

There remains an unmet clinical need within clinical practice for an effective, tolerable, easily-administrable treatment option for patients whose only alternative is the expensive injection-delivered biologic dupilumab or clinically-ineffective BSC. It is expected that clinicians will use baricitinib as an alternative to dupilumab following consideration of a systemic immunosuppressant agent, in line with the treatment pathway relevant to clinical practice in the NHS. Baricitinib is administered orally, removing the burden of subcutaneous injection and the common injection site reactions associated with dupilumab administration and offering the potential to simplify dramatically the treatment paradigm for patients in this setting. Baricitinib is therefore an attractive option for patients which also demonstrates robust cost-effectiveness versus both dupilumab and BSC in adult patients with moderate-to-severe AD who failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control.

B.4 References

1. National Institute for Health and Care Excellence (NICE). Dupilumab for treating moderate to severe atopic dermatitis: Technology appraisal guidance [TA534]. Available at: https://www.nice.org.uk/guidance/ta534 [Last accessed: 28th January 2020].
2. Bao L, Zhang H, Chan LS. The involvement of the JAK-STAT signaling pathway in chronic inflammatory skin disease atopic dermatitis. JAK-STAT 2013;2:e24137.
3. Lilly. How does Olumiant work? Available at: https://www.olumiant.com/hcp/moa [Last accessed: 20th January 2020].
4. National Institute for Health and Care Excellence (NICE). Baricitinib for moderate to severe rheumatoid arthritis: Technology appraisal guidance [TA466]. Available at: https://www.nice.org.uk/guidance/ta466 [Date accessed: 17th June 2020].
5. British National Formulary. Available at: www.bnf.org [Last accessed: 28th January 2020].
6. NHS. Atopic eczema. Available at: https://www.nhs.uk/conditions/atopic-eczema/ [Last accessed: 23rd January 2020].
7. Hong J, Buddenkotte J, Berger TG, et al. Management of Itch in Atopic Dermatitis. Seminars in Cutaneous Medicine and Surgery 2011;30:71-86.
8. National Institute for Health and Care Excellence (NICE). Clinical Knowledge Summary: Eczema - atopic. What is it? Available at: https://cks.nice.org.uk/eczema-atopic#!backgroundSub [Last accessed: 24th January 2020].
9. National Eczema Society. Our Skin and Eczema. Available at: http://www.eczema.org/our-skin-and-eczema/ [Last accessed: 12th June 2020].
10. Schultz Larsen F, Holm NV, Henningsen K. Atopic dermatitis: A genetic-epidemiologic study in a population-based twin sample. Journal of the American Academy of Dermatology 1986;15:487-494.
11. Barbarot S, Auziere S, Gadkari A, et al. Epidemiology of atopic dermatitis in adults: Results from an international survey. Allergy 2018;73:1284-1293.
12. National Institute for Health and Care Excellence (NICE). Resource impact report: Dupilumab for treating moderate to severe atopic dermatitis. Available at: https://www.nice.org.uk/guidance/ta534/resources [Accessed: 29th January 2020].
13. Silverberg JI, Gelfand JM, Margolis DJ, et al. Patient burden and quality of life in atopic dermatitis in US adults: A population-based cross-sectional study. Annals of Allergy, Asthma & Immunology 2018;121:340-347.
14. Langenbruch A, Radtke M, Franzke N, et al. Quality of health care of atopic eczema in Germany: results of the national health care study AtopicHealth. J Eur Acad Dermatol Venereol 2014;28:719-26.
15. Chamlin SL, Mattson CL, Frieden IJ, et al. The Price of Pruritus: Sleep Disturbance and Cosleeping in Atopic Dermatitis. Archives of Pediatrics & Adolescent Medicine 2005;159:745-750.
16. McPherson T. Current Understanding in Pathogenesis of Atopic Dermatitis. Indian J Dermatol 2016;61:649-655.
17. Vakharia PP, Chopra R, Sacotte R, et al. Burden of skin pain in atopic dermatitis. Annals of Allergy, Asthma & Immunology 2017;119:548-552.e3.
18. Ong PY, Leung DYM. Bacterial and Viral Infections in Atopic Dermatitis: a Comprehensive Review. Clinical Reviews in Allergy & Immunology 2016;51:329-337.
19. Cathcart S, Theos A. Inpatient management of atopic dermatitis. Dermatologic Therapy 2011;24:249-255.
20. Egeberg A, Skov L, Andersen YMF, et al. Ten-year mortality is increased after hospitalization for atopic dermatitis compared with the general population, but reduced compared with psoriasis. Journal of the American Academy of Dermatology 2017;76:98-105.
21. Simpson EL. Comorbidity in Atopic Dermatitis. Curr Dermatol Rep 2012;1:29-38.
22. Yano C, Saeki H, Ishiji T, et al. Impact of disease severity on sleep quality in Japanese patients with atopic dermatitis. J Dermatol Sci 2013;72:195-7.
23. Chrostowska-Plak D, Reich A, Szepietowski JC. Relationship between itch and psychological status of patients with atopic dermatitis. J Eur Acad Dermatol Venereol 2013;27:e239-42.
24. Rønnstad ATM, Halling-Overgaard A-S, Hamann CR, et al. Association of atopic dermatitis with depression, anxiety, and suicidal ideation in children and adults: A systematic review and meta-analysis. Journal of the American Academy of Dermatology 2018;79:448-456.e30.

25. Michigan Medicine (University of Michigan): Atopic Dermatitis (Eczema). Available at: https://www.uofmhealth.org/health-library/hw216104 [Last accessed: 26th January 2020].
26. de Bruin-Weller M, Gadkari A, Auziere S, et al. The Patient-reported Disease Burden in Adults with Atopic Dermatitis: A Cross-sectional Study in Europe and Canada. J Eur Acad Dermatol Venereol 2020;34:1026-1036.
27. Simpson EL, Guttman-Yassky E, Margolis DJ, et al. Association of Inadequately Controlled Disease and Disease Severity With Patient-Reported Disease Burden in Adults With Atopic Dermatitis. JAMA dermatology 2018;154:903-912.
28. Zuberbier T, Orlow SJ, Paller AS, et al. Patient perspectives on the management of atopic dermatitis. J Allergy Clin Immunol 2006;118:226-32.
29. Sanofi Genzyme. Seeing Red: Getting Under the Skin of Adult Severe Eczema, 2017.
30. The Primary Care Dermatology Society & British Association of Dermatologists: Guidelines for the management of atopic eczema (2009).
31. NHS. Hospital Admitted Patient Care Activity 2018-19: Diagnosis. Available at: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2018-19 [Last accessed: 24th January 2020].
32. Boguniewicz M, Alexis AF, Beck LA, et al. Expert Perspectives on Management of Moderate-to-Severe Atopic Dermatitis: A Multidisciplinary Consensus Addressing Current and Emerging Therapies. J Allergy Clin Immunol Pract 2017;5:1519-1531.
33. Simpson EL, Bruin-Weller M, Flohr C, et al. When does atopic dermatitis warrant systemic therapy? Recommendations from an expert panel of the International Eczema Council. J Am Acad Dermatol 2017;77:623-633.
34. National Institute for Health and Care Excellence (NICE). Atopic eczema in under 12s: diagnosis and management. Clinical guideline [CG57]. 2007. Available at: https://www.nice.org.uk/guidance/cg57. Date accessed: 5th June 2020.
35. National Institute for Health and Care Excellence (NICE). Atopic eczema in under 12s. Quality standard [QS44]. 2013. Available at: https://www.nice.org.uk/guidance/qs44. Date accessed: 5th June 2020.
36. Taylor K, Swan DJ, Affleck A, et al. Treatment of moderate-to-severe atopic eczema in adults within the U.K.: results of a national survey of dermatologists. British Journal of Dermatology 2017;176:1617-1623.
37. International Eczema Council. Understanding Eczema. Available at: https://www.eczemacouncil.org/for-patients/understanding-eczema/ [Last accessed: 26th January 2020]. Volume 2019.
38. National Institute for Health and Care Excellence (NICE). Frequency of application of topical corticosteroids for atopic eczema: Technology appraisal guidance [TA81]. Available at: https://www.nice.org.uk/guidance/ta81 [Last accessed: 28th January 2020].
39. National Institute for Health and Care Excellence (NICE). Tacrolimus and pimecrolimus for atopic eczema: Technology appraisal guidance [TA82]. Available at: https://www.nice.org.uk/guidance/ta82 [Last accessed: 28th January 2020].
40. Akdis CA, Akdis M, Bieber T, et al. Diagnosis and treatment of atopic dermatitis in children and adults: European Academy of Allergology and Clinical Immunology/American Academy of Allergy, Asthma and Immunology/PRACTALL Consensus Report. Journal of Allergy and Clinical Immunology 2006;118:152-169.
41. Eichenfield LF, Tom WL, Berger TG, et al. Guidelines of care for the management of atopic dermatitis: Section 2. Management and treatment of atopic dermatitis with topical therapies. Journal of the American Academy of Dermatology 2014;71:116-132.
42. Brunner PM, Guttman-Yassky E, Leung DY. The immunology of atopic dermatitis and its reversibility with broad-spectrum and targeted therapies. J Allergy Clin Immunol 2017;139:S65-s76.
43. Patrizi A, Raone B, Ravaioli GM. Management of atopic dermatitis: safety and efficacy of phototherapy. Clinical, cosmetic and investigational dermatology 2015;8:511-520.
44. Sidbury R, Davis DM, Cohen DE, et al. Guidelines of care for the management of atopic dermatitis: section 3. Management and treatment with phototherapy and systemic agents. J Am Acad Dermatol 2014;71:327-49.
45. EMC. Neoral Soft Gelatin Capsules: SmPC. Available at: https://www.medicines.org.uk/emc/product/1034/smpc [Last accessed: 28th January 2020].
46. EMC. Dupixent 300 mg Solution for Injection in Pre-filled Syringe: SmPC. Available at: https://www.medicines.org.uk/emc/product/8553/smpc [Last accessed: 28th January 2020].

47. Akinlade B, Guttman-Yassky E, de Bruin-Weller M, et al. Conjunctivitis in dupilumab clinical trials. British Journal of Dermatology 2019;181:459-473.
48. Guttman-Yassky E, Silverberg JI, Nemoto O, et al. Baricitinib in adult patients with moderate-to-severe atopic dermatitis: A phase 2 parallel, double-blinded, randomized placebo-controlled multiple-dose study. J Am Acad Dermatol 2019;80:913-921.e9.
49. Simpson EL, Lacour JP, Spelman L, et al. Baricitinib in patients with moderate-to-severe atopic dermatitis and inadequate response to topical corticosteroids: results from two randomized monotherapy phase III trials. Br J Dermatol 2020.
50. Beck LA, Thaçi D, Hamilton JD, et al. Dupilumab treatment in adults with moderate-to-severe atopic dermatitis. New England Journal of Medicine 2014;371:130-139.
51. Guttman-Yassky E, Bissonnette R, Ungar B, et al. Dupilumab progressively improves systemic and cutaneous abnormalities in patients with atopic dermatitis. Journal of Allergy and Clinical Immunology 2019;143:155-172.
52. Blauvelt A, Simpson EL, Tyring SK, et al. Dupilumab does not affect correlates of vaccine-induced immunity: a randomized, placebo-controlled trial in adults with moderate-to-severe atopic dermatitis. Journal of the American Academy of Dermatology 2019;80:158-167. e1.
53. Worm M, Simpson EL, Thaçi D, et al. Efficacy and safety of multiple dupilumab dose regimens after initial successful treatment in patients with atopic dermatitis: a randomized clinical trial. JAMA dermatology 2020;156:131-143.
54. I4V-MC-JAIN (BREEZE-AD4) Clinical Study Report (13th April 2020).
55. I4V-MC-JAIY (BREEZE-AD7) Clinical Study Report (25th October 2019).
56. I4V-MC-JAHL (BREEZE-AD1) Clinical Study Report (24th September 2019).
57. I4V-MC-JAHM (BREEZE-AD2) Clinical Study Report (8th October 2019).
58. Bieber T, Reich K, Paul C, et al. Efficacy and Safety of Baricitinib in Combination With Topical Corticosteroids in Patients With Moderate-to-Severe Atopic Dermatitis Who Failed, Are Intolerant to, or Have Contraindication to Cyclosporine: Results From a Randomized, Placebo-Controlled, Phase 3 Clinical Trial (BREEZE-AD4) [P0183]. Poster presented at European Academy of Dermatology and Venereology (EADV Virtual), 29-31 October, 2020.
59. I4V-MC-JAHN (BREEZE-AD3) Clinical Study Report (7th November 2019).
60. I4V-MC-JAIN(c) Clinical Protocol (9th May 2019).
61. I4V-MC-JAIN Statistical Analysis Plan Version 1 (4th December 2018).
62. I4V-MC-JAIY(a) Clinical Protocol 18th December 2018).
63. I4V-MC-JAIY Statistical Analysis Plan Version 2 (12th August 2019).
64. I4V-MC-JAHL(a) Clinical Protocol (15th December 2017).
65. I4V-MC-JAHL Statistical Analysis Plan Version 2 (16th January).
66. I4V-MC-JAHM(b) Clinical Protocol (22nd January 2018).
67. I4V-MC-JAHM Statistical Analysis Plan Version 2 (22nd January 2019).
68. Eli Lilly: Secondary Censoring Data.
69. Devlin NJ, Shah KK, Feng Y, et al. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ 2018;27:7-22.
70. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095-108.
71. de Bruin-Weller M, Thaci D, Smith CH, et al. Dupilumab with concomitant topical corticosteroid treatment in adults with atopic dermatitis with an inadequate response or intolerance to ciclosporin A or when this treatment is medically inadvisable: a placebo-controlled, randomized phase III clinical trial (LIBERTY AD CAFE). Br J Dermatol 2018;178:1083-1101.
72. Blauvelt A, de Bruin-Weller M, Gooderham M, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corticosteroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. The Lancet 2017;389:2287-2303.
73. Simpson EL, Bieber T, Guttman-Yassky E, et al. Two Phase 3 Trials of Dupilumab versus Placebo in Atopic Dermatitis. New England Journal of Medicine 2016;375:2335-2348.
74. Thaçi D, Simpson EL, Beck LA, et al. Efficacy and safety of dupilumab in adults with moderate-to-severe atopic dermatitis inadequately controlled by topical treatments: a randomised, placebo-controlled, dose-ranging phase 2b trial. The Lancet 2016;387:40-52.
75. R Package 'meta'. Available at: https://cran.r-project.org/web/packages/meta/meta.pdf [Last accessed: 29th April 2020].
76. NICE. Guide to the methods of technology appraisal 2013: The reference case. Available at: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case [Last accessed: 11/02/20].

77. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Medical decision making 1993;13:322-338.
78. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012;15:708-15.
79. EMC. Olumiant 2 mg Film-Coated Tablets: SmPC. Available at: https://www.medicines.org.uk/emc/product/2434/smpc [last accessed: 20th April 2020].
80. National Institute for Health and Care Excellence (NICE). Final Appraisal Determination: Dupilumab for treating moderate to severe atopic dermatitis. Available at: https://www.nice.org.uk/guidance/ta534/evidence [Date accessed: 18th June 2020].
81. Office for National Statistics. Population Estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2019. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2019 [Last accessed: 28th May 2020].
82. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. Value Health 2011;14:539-45.
83. Monthly Index of Medical Specialties (MIMS). 2020. Available at: https://www.mims.co.uk/. Date accessed: 9th June 2020.
84. NHS Improvement. National Cost Collection data: 2018-2019. Available at: https://improvement.nhs.uk/resources/national-cost-collection/. Date accessed: 9th June 2020.
85. PSSRU: Unit Costs of Health and Social Care (2019). Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/. Last accessed: 9th June 2020.
86. Hatswell AJ, Bullement A, Briggs A, et al. Probabilistic Sensitivity Analysis in Cost-Effectiveness Models: Determining Model Convergence in Cohort Models. Pharmacoeconomics 2018;36:1421-1426.

B.5 Appendices

Appendix C: Summary of Product Characteristics (SmPC)

Appendix D: Identification, selection and synthesis of clinical evidence

Appendix E: Subgroup analysis

Appendix F: Adverse reactions

Appendix G: Published cost-effectiveness studies

Appendix H: Health-related quality-of-life studies

Appendix I: Cost and healthcare resource identification, measurement and valuation

Appendix J: Clinical outcomes and disaggregated results from the model

Appendix K: Checklist of confidential information

Appendix L: Additional data from the BREEZE-AD1, 2, 4 and 7 trials

Appendix M: BREEZE-AD3 (JAHN)

Appendix N: Cost-effectiveness model validation checklists

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Baricitinib for the Treatment of Moderate-to-Severe Atopic Dermatitis [ID1622]

Clarification questions

August 2020

File name	Version	Contains confidential information	Date
ID1622_Baricitinib AD_Reponse to Clarification Questions_ Fully Redacted	FINAL	Yes	03/08/20

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Baricitinib clinical effectiveness

A1. A footnote in Table 4 and section B.2.11 of the company submission (CS) mention future data cuts of BREEZE-AD3 (JAHN) and BREEZE-AD4 (JAIN). Please indicate when these data cuts will be available.

The BREEZE-AD3 and -AD4 trials are ongoing. As reported in Section B.2.11 of the CS, additional data from BREEZE-AD4 may become available in October 2020, and additional data from BREEZE-AD3 in November 2020.

A2. Figures 8-17 in the CS are cut at 16 weeks, although a 24-week endpoint is reported for the outcomes reported in the Figures. Please extend Figures 8-17 to 24 weeks.

Figures 8–17 of the company submission have been extended to include the 24-week timepoint. Please note that a few significances marked with *, **, *** up to Week 16 may differ from the respective figures in the NICE dossier, due to some mistakes in the original figures (e.g. rounding), or for MMRM results, where new data up to Week 24 is included into the models (and not only until Week 16). All p-values are based on logistic regression.

Figures 8-17 up to 24 weeks.pdf

A3. Baricitinib failed to demonstrate a statistically significant benefit versus placebo in JAIN across several secondary endpoints at Week 24, despite reaching significance at Week 16. These endpoints include EASI75, IGA ≤ 1, EASI50, EASI90 and SCORAD 75 (section B.2.6.1 of the CS). Please provide an explanation as to why this is the case.

During the trial, patients who required rescue medication at any time, including at baseline, were assessed as non-responders. In addition to this non-responder imputation, the most stringent statistical analysis was applied to assess efficacy in this patient population. Despite this, a significantly higher proportion of patients receiving 4 mg baricitinib achieved clinically meaningful improvements in signs and symptoms at Week 16.

Three main reasons are likely to contribute to the loss of a statistically significant benefit associated with baricitinib beyond Week 16 for some endpoints:

- Primary censoring rule: Study JAIN evaluated the efficacy and safety of baricitinib in combination with TCS in patients who had previously failed or were contraindicated or intolerant to ciclosporin. Furthermore, all patients underwent a washout period of all AD-specific therapies before study entry. This washout period of 5 half-lives for biologic AD treatments, 4 weeks for systemic AD treatments and 2 weeks for topical AD treatments (including TCS), excluding emollients, was longer than in the CHRONOS and SOLO1/2 trials. Therefore, these patients can reasonably be considered a difficult to treat population with active disease at baseline in a condition which is inherently flaring in nature. In this context, 24 weeks is a relatively long period of time and the occurrence of flares necessitating rescue therapy in the form of higher potency TCS is to be expected. As explained in Section B.2.4 of the company submission, the primary censoring rule censored patients following the use of rescue therapy or permanent study drug discontinuation, after which patients are considered as non-responders regardless of the length of time for which rescue therapy was used. In a flaring disease, this is a conservative rule and high rates of rescue between Week 16 and Week 24 may have skewed results for categorical variables where NRI was used to account for censoring, if patients who had received rescue therapy were still benefitting from treatment with baricitinib. This is demonstrated by the more favourable results obtained when using the secondary censoring rule in which patients were censored as non-responders only following permanent study drug discontinuation (see Section B.2.6 of the company submission).
- **Sample size:** After Week 16, the placebo, 1 mg baricitinib and 4 mg baricitinib groups each had a sample size of around 90 patients, meaning that each patient accounted for more than 1% in these groups and small differences can significantly skew the data. In the 4 mg group, data for one responder were not transferred before database lock (due to a data entry error) and another responder moved to another country and thus study participation was interrupted. Therefore, both of these patients were classified as non-responders. These events did not occur in other groups.
- **Background TCS use after Week 16:** By Week 24 of this combination trial, TCS alone were able to reduce disease activity sufficiently in some patients for rescue therapy not to be needed. Therefore, it is as expected that baricitinib in combination with TCS was numerically but no longer statistically significantly superior to placebo in combination with

TCS. In order to this achieve disease control, the placebo group used more TCS than the baricitinib groups as shown in Figure 1.

Figure 1: Mean quantity of background TCS used by patients in the JAIN trial between Weeks 0 and 24

Abbreviations: gramq: gram quantity; LSM: least squares mean; TCS: topical corticosteroids.

Subgroups

A4. PRIORITY: Table 49 of the company submission provides subgroup analysis for JAIY trial. Please provide similar subgroup analysis for the JAIN study and the JAIN + JAIN-like JAIY population. If these subgroup data are not yet available, please indicate when this subgroup data will be available.

The proportion of patients achieving IGA ≤1, EASI75 or a ≥4-point improvement in Itch NRS at Week 16 for subgroups with significant interactions ($p<0.1$) is presented in Table 1 for the JAIN trial and the JAIN + JAIN-like JAIY pooled population.

In BREEZE-AD4 (JAIN), a significant interaction ($p<0.05$) was observed between gender and 4-point improvement in Itch NRS at Week 16. At Week 16 in the pooled combination therapy patients from JAIN + JAIN-like JAIY population, significant interactions ($p<0.05$) were observed between gender and IGA ≤1 and between geographical region and EASI75.

Table 1: Proportion of patients in BREEZE-AD4 [JAIN] and the pooled BREEZE-AD4 and -AD7 (JAIN and JAIN-like JAIY) population achieving IGA ≤1, EASI75 or a ≥4-point improvement in Itch NRS at Week 16 for subgroups with significant interactions (p<0.1)

Outcome	Subgroup	Category	Response at Week 16 (%)				p-value[a]
			PBO	1 mg BARI	2 mg BARI	4 mg BARI	
Combination therapy: BREEZE-AD4 (JAIN) (N=[redacted])							
EASI75	Region	Europe (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		Japan (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		ROW (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
Itch NRS improvement of 4 or more points	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
Pooled combination therapy: BREEZE-AD4 and -AD7 (JAIN and JAIN-like JAIY) (N=[redacted])							
IGA ≤1	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
EASI75	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
	Region	Europe (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		Japan (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		ROW (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	

		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
Itch NRS improvement of 4 or more points	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
	Prior systemic therapy[b]	Yes (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	
		No (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]	
		Relative risk vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	

[a] p-value shows treatment by subgroup interaction value and includes all doses of baricitinib. [b] N=[redacted] since one patient was reported as "Other".
Abbreviations: BARI: baricitinib; NC: not calculated; NRS: numerical rating scale; PBO: placebo; ROW: Rest of World; RR: risk ratio.

A5. The sensitivity analysis JAIN (European patients only) versus CAFÉ presented in section B.2.9.5 of the CS shows a much greater response in the JAIN-European patients (OR: [redacted]) compared to the full JAIN trial (OR: [redacted]). The JAIY subgroup analyses by region in Table 49 also show a similar effect. Please provide:

a) Baseline characteristics for the JAIN (European patients only) subgroup.

The baseline characteristics of the European-only patients in the JAIN trial are presented in Table 2. As compared with the baseline characteristics for all patients of the JAIN trial, presented in Section B.2.3.3 of the company submission, European JAIN patients were of a similar age, with a higher proportion of Caucasian participants, as expected. The baseline disease characteristics were broadly similar to the JAIN trial, so differences in efficacy do not appear to be driven by differences in baseline risk.

Table 2: Baseline characteristics of European patients only from BREEZE-AD4 (JAIN)

Characteristic	BREEZE-AD4 (JAIN) European patients only			
	PBO (N=[redacted])	1 mg (N=[redacted])	2 mg (N=[redacted])	4 mg (N=[redacted])
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Female, %	[redacted]	[redacted]	[redacted]	[redacted]
Race				
Caucasian, %	[redacted]	[redacted]	[redacted]	[redacted]
Asian, %	[redacted]	[redacted]	[redacted]	[redacted]
Other, %	[redacted]	[redacted]	[redacted]	[redacted]
Duration since AD diagnosis (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Weight (kg), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Body mass index (kg/m²), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Geographic region				
Europe, %	[redacted]	[redacted]	[redacted]	[redacted]
Japan, %	[redacted]	[redacted]	[redacted]	[redacted]
Rest of world, %	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 at screening Visit 1, %	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 Visit 2, %	[redacted]	[redacted]	[redacted]	[redacted]
EASI, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
SCORAD, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
BSA affected, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
POEM, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
ADSS Item 2, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
DLQI, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]

Itch NRS, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Skin Pain NRS, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
PGI-S-AD, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS depression, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
HADS anxiety and depression combined, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L VAS score, mean	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical calcineurin inhibitor use, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Prior systemic therapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Systemic corticosteroid use	[redacted]	[redacted]	[redacted]	[redacted]
Systemic immunosuppressant use	[redacted]	[redacted]	[redacted]	[redacted]
Ciclosporin use	[redacted]	[redacted]	[redacted]	[redacted]
Phototherapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; HIS: health index score; IGA: Investigator's Global Assessment; NRS: Numeric Rating Scale; PBO: placebo; PGI-S-AD: Patient Global Impression of Severity–Atopic Dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.; SD: standard deviation; TCS: topical corticosteroids; VAS: visual analogue score.

b) An explanation as to why there is a greater response in European patients.

There are two observations which together are likely to explain the difference in the Europe-only population analysed:

- **Rescue rates:** The rescue rates observed in Europe ([redacted]%) were much lower than in Japan ([redacted]%) and the rest of the world ([redacted]%). Therefore, response rates as assessed by the primary censoring rule, which censors following the use of rescue therapy or permanent study drug discontinuation, was found to be higher in the European population. While it possible that this reflects differences in the disease itself, this is not possible to determine, and it remains likely that it is instead reflective, at least in part, of differences in clinical practice and investigator's choice. For example, Japanese clinical practice favours TCS use, including high potency TCS, rather than systemic agents, while European clinical practice broadly limits the use of high potency TCS. Therefore, Japanese patients would be more likely to be rescued with higher potency TCS, leading to non-responder imputation indicating lower response rates in these patients.
- **Response rates in the PBO arm:** As compared with Europe, other regions experienced higher response rates in the PBO + TCS treatment arm, leading to a conclusion of relatively lower efficacy in these areas. While no definitive explanation for this observation is possible, it is likely to suggest that prior failure to TCS in non-European areas were associated with insufficient use or potency of TCS and that patients who received suitable

potency TCS and applied it as directed in the context of the clinical trial did then observe a clinical response. Several factors, including patient preferences or reimbursement issues could contribute to previously insufficient TCS use.

Overall, these observations suggest that the assessment of eligibility and efficacy for different patient populations is likely to be influenced by local clinical practice and assessment and this context should be considered when interpreting the results obtained. It should further be considered that the trial was not designed to investigate baricitinib efficacy in European patients compared to other patient populations and therefore definitive conclusions cannot be reached.

A6. The subgroup analysis presented in Table 49 of the CS indicates that baricitinib is [redacted]. Please provide an explanation as to why this is the case, including the potential for differences in disease pathology among such populations.

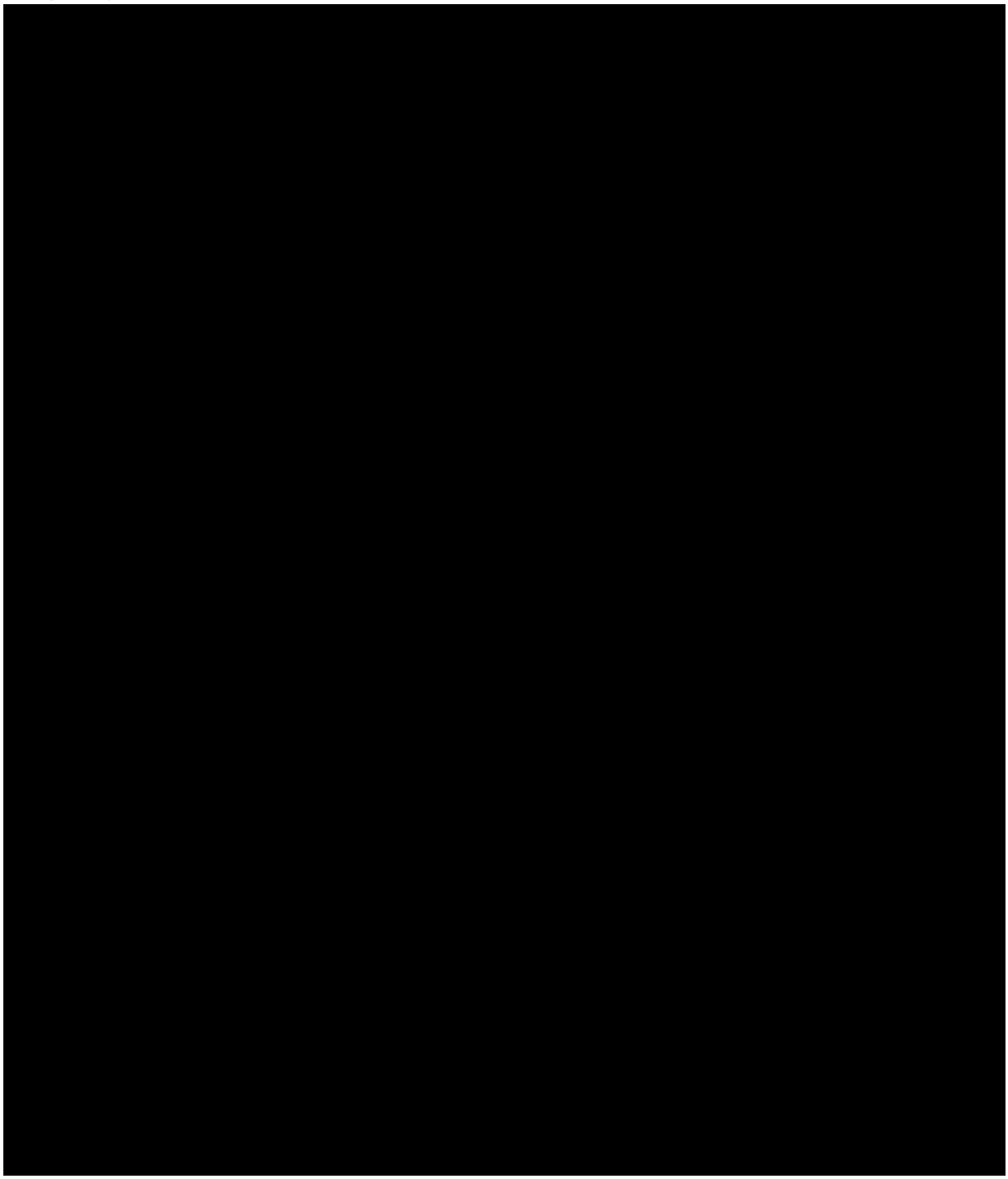
Atopic dermatitis is well characterised as a Th2-driven disease. However, it is known to encompass heterogeneous phenotypes and additional activation of Th22, Th17 and Th1 cytokine pathways may occur depending on the subtype of the disease. For example, in Asian AD patients, the Th17 axis is significantly increased as compared with European American patients.[1] In addition, there may be other geographical factors that affect efficacy of treatment, such as patterns in concomitant medication use, adherence to treatment, natural history of the disease course affected by differences in the local healthcare system as well as differences in the investigator assessment of the disease.

Nonetheless, differential clinical effect of JAK1/2 inhibition by baricitinib in different AD phenotypes has not yet been established and the trial program was not designed to investigate baricitinib efficacy in Japanese patients compared with other patient populations. The assessment of eligibility and efficacy for different patient populations is nuanced and clinicians should be aware of the way in which the disease presents for these groups.

While no definitive statements on differential clinical effect can be made, exploratory analyses for potential differences among these populations were undertaken in two monotherapy trials. Exposure-response analysis was conducted with data from Studies JAHL and JAHM (the two Phase 3 studies where PK samples were collected) up to 16 weeks of treatment. Although 5 patient factors were identified as statistically significant covariates (gender, age, disease severity, weight, and Japanese patient population), only weight and Japanese patient population were identified as significant covariates related to drug effect.

As shown in Figure 2, there was separation between the [redacted]% prediction intervals in the response-time plots for patients recruited in Japan, largely for the outcome of the IGA 0 or 1 response, but less separation was observed in the other clinically relevant endpoints analysed (EASI50, EASI75, EASI90 and 4-point improvement in Itch NRS).

Figure 2: Simulated EASI50/75/90, IGA 0 or 1, and Itch NRS response rates, with secondary censoring, for non-Japanese and Japanese patients with AD over 16 weeks at 4-mg daily dose in Studies JAHL/JAHM



Solid lines represent the median predicted response from the model prediction. The shaded area is the 90% confidence interval of the model prediction. The only between-group difference is ethnicity; other covariates were fixed to moderate disease severity (baseline IGA of 3), age of 33 years (population median), weight of 71 kg (population median) and sex as female.
Abbreviations: AD: atopic dermatitis; EASI: Eczema Area and Severity Index; IGA: Investigator's Global Assessment; ITCH: a 4-point improvement in Itch NRS; NRS: Numeric Response Scale.

Interpretation of the covariate effect for Japanese patients should be done with caution because the Japanese subpopulation was relatively small ([redacted] patients in total, or [redacted]% of the total patients included in the pharmacodynamic dataset) and differences for several baseline disease characteristics were noted between Japanese and non-Japanese patients. These differences

included EASI and BSA baseline characteristics (more severe AD in Japanese patients) and the use of TCS ([redacted]% by Week 16 in Japan compared with [redacted]% in the overall population). This increased use of TCS also suggests that in order to compare the Japanese and overall populations, the use of secondary censoring (where data after rescue are included) may be more appropriate than primary censoring (where data after rescue have been suppressed).

The exposure-response analysis considered only baricitinib-treated patients, and not the effect relative to PBO. Based on the integrated JAHL/JAHM clinical dataset, and in line with the overall results, the number of Japanese patients who achieved response for key endpoints (e.g. IGA 0, 1; EASI75; mean percent change from baseline in Itch NRS) at Week 16 with baricitinib 4 mg was consistently higher than that in PBO.

With regard to the IGA 0 or 1 response at Week 16, both PBO and treatment response rate were lower in Japan than in the overall study population, resulting in the treatment effect relative to PBO in the Japanese and overall populations being more similar. Specifically, with primary censoring, the difference between PBO and 4 mg was [redacted]% ([redacted]% versus [redacted]%) for the overall population versus [redacted]% ([redacted]% versus [redacted]%) in the Japanese population. For secondary censoring, the difference between PBO and 4 mg was [redacted]% ([redacted]% versus [redacted]%) for the overall population versus [redacted]% ([redacted]% versus [redacted]%) in the Japanese population.

Additionally, the results of regional subgroup analyses for IGA 0 or 1 in Studies JAHL and JAHM and in Study JAIY, shown in Figure 3 and Figure 4, respectively, do not support a lower response for baricitinib 4 mg in East Asian countries not including Japan, suggesting that there is not a specific effect of East Asian ethnicity.

Figure 3: Proportion of patients achieving clear or almost clear skin at Week 16 by region - pooled monotherapy Studies JAHL and JAHM with primary censoring.

Abbreviations: IGA: Investigator's Global Assessment; N = number of patients.

Figure 4: Proportion of patients achieving clear or almost clear skin at Week 16 by region - Study JAIY with primary censoring.

In Study JAIY, East Asia includes Japan.
Abbreviations: IGA: Investigator's Global Assessment; N = number of patients.

Taken together, these results do not suggest a specific effect of Asian ethnicity on the treatment effect of baricitinib, as illustrated by a lack of relevant treatment-by-subgroup interactions in the submitted studies, as well as the recently available data from Study JAIN. The overall low response rate in both PBO and baricitinib-treated patients in Japan likely reflect differences in baseline characteristics and treatment practices related to TCS rescue specific to Japan.

A7. Tables 8 to 11 of the CS suggest [redacted] were enrolled in any of the pivotal trials. Please comment on the lack of evidence to support efficacy in this population and the potential for differences in efficacy in black patients, noting any evidence for differences in disease pathology.

The trial program was not designed to investigate baricitinib efficacy in black patients compared with other patient populations and as such, the ethnicity distributions of the BREEZE-AD trials are reflective of the participating countries rather than of the occurrence of AD. Of note, patients were not recruited from the US; if they had been, it would be expected that a higher proportion of black patients would have been recruited.

The clinical effect of JAK1/2 inhibition by baricitinib in different AD phenotypes has not been established. Some evidence does exist that the pathology of AD could be distinct in black patients: mutations in the FLG gene leading to a deficiency in filaggrin have been associated with AD that is more severe and persistent than its wild type counterpart. These mutations are detected in up to 30% of individuals, but they are rarely identified in African-American populations with AD.[1] The differences in the cytokine pathways involved in atopic dermatitis across ethnic groups were noted in the dupilumab appraisal (TA534), but it was considered by the Appraisal Committee that there was insufficient evidence to determine the extent to which different cytokine pathways modify treatment effect. For this reason, the variation in cytokine expression in different ethnic groups was not considered further.[2]

A8. The NICE scope lists diseases severity as a potentially relevant subgroup; however, no clinical and economic evidence was presented for the subgroups with moderate and severe disease. The company submission states that this is because of the lack of a widely accepted classification system. However, there are published strata that allow classification by EASI score.[3, 4] As such, please provide subgroup analyses for moderate disease and severe disease using EASI score to classify patients for JAIY and the pooled JAHL and JAHM studies, to supplement what is provided in Table 49 of the CS.

As discussed in Section B.1.3.3 of the CS, the severity of AD in UK clinical practice can be classified as mild, moderate or severe based on a variety of clinical features. Whilst disease severity is a key consideration, it is not the sole consideration for treatment decisions. A steering committee consisting of a multidisciplinary group of AD experts, including 8 dermatologists, 2 allergists, and a patient advocacy group representative concluded that AD may be considered moderate-to-severe when one or more of the following criteria are met:[5]

- A minimum involvement of 10% body surface area (BSA)
- Regardless of BSA:
    - Presence of individual lesions with moderate-to-severe features
    - Involvement of highly visible areas or those important for function
    - Significantly impaired quality of life

As the ERG have highlighted, EASI score provides one classification system for AD signs. However, this measure does not capture all of the criteria listed above, and thus may not reflect all aspects of moderate or severe disease. As a result, using different measures to define severity (e.g. EASI score and IGA) may result in different disease classifications. In addition, AD is a flaring disease, and thus EASI score alone would not provide consistent classification of disease severity. The steering committee concluded that current disease severity scales, including EASI, although validated for use in clinical trials, are not practical for routine use in clinical practice.[5] Accordingly, the published strata highlighted by the ERG are recommended for use in clinical trials, not for use in clinical practice where NICE guidance is applicable.[3, 4]

In adult patients, treatment depends largely on clinician assessment of need, with over 90% of consultant-level dermatologists in a UK-based study reporting their own clinical experience

influenced or strongly influenced their choice of treatment for adult patients with moderate-to-severe AD.[6] Feedback from UK clinicians experienced in the treatment of AD confirmed that strata based on EASI score would not be used in isolation to inform treatment decisions for patients with moderate and severe disease in UK clinical practice. As such, it was not considered clinically appropriate to use EASI score to conduct subgroup analyses.

As well as these clinical considerations, no subgroup analyses for moderate versus severe AD are available for dupilumab and thus it is not feasible to conduct any efficacy comparisons with the key comparator in these populations.

For these reasons, it was considered inappropriate and infeasible to conduct these subgroup analyses. In a situation in which these analyses were considered appropriate and feasible, it would also be necessary to consider that many of the inputs underlying the cost-effectiveness model are associated with a moderate-to-severe population, including efficacy, utility and healthcare resource inputs.

Safety and discontinuation

A9. Please provide more details of the phase II JAHG study included in the safety analysis, including separate safety data.

A summary of the methodology of the JAHG trial is provided in Table 3. A summary of the adverse events and TEAEs affecting >3% of patients in the placebo and 4 mg baricitinib treatment groups of the JAHG trial are presented in Table 4 and Table 5, respectively. Additional data from the JAHG trial are available in the public domain.[7]

Table 3: Summary of JAHG trial methodology

Trial name	**JAHG[8]**
Location	Patients recruited from 10 sites in the US and 3 sites in Japan
Trial design	Phase II, double-blind, randomised, placebo-controlled study
Eligibility criteria for participants	Key inclusion criteria: • EASI≥12 at Visits 1 and 2 • ≥10% BSA involvement at Visits 1 and 2 • Diagnosed with AD at least 2 years before Visit 1 • A history of inadequate clinical response, in the opinion of the investigator, to 1 or more of the 3 treatment categories listed below (used for at least 4 weeks): o Category 1: Hydration plus topical steroids and/or antibiotics (e.g., tetracycline, trimethoprim and sulfamethoxazole, cephalosporins) and/or topical immune modulators (e.g., tacrolimus/pimecrolimus) o Category 2: Systemic steroids and/or phototherapy o Category 3: Cyclosporine and/or other immunomodulators (e.g., methotrexate, mycophenolate mofetil, and azathioprine) Key exclusion criteria: • Receiving prohibited AD therapies • A recent history of infection including active or untreated latent tuberculosis or other serious infection • Immunocompromised

	• Abnormal laboratory results • Comorbidities that increased patient risk when taking study drug
Method of study drug administration	Administered orally once daily as two tablets: 1 treatment tablet and 1 placebo tablet for the 2 mg and 4 mg treatment groups, or 2 placebo tablets for the placebo group.
Permitted and disallowed concomitant medication	All concomitant therapies for AD were prohibited throughout the trial except for: • Those outlined in the inclusion criteria above • Triamcinolone 0.1% cream • Systemic drugs required to treat an AE • Non-live seasonal vaccines and/or emergency vaccinations, such as rabies or tetanus • Topical antibiotics in the event of secondary infections and lesions (as needed) • Non-prescription shampoo (as needed) • Non-steroidal anti-inflammatory drugs (as needed) • Antihistamines (as needed) • Topical moisturisers or emollients, bath oils, oatmeal bath preparations or bleach bats if using a stable regimen prior to enrolment (as needed) • Salicylic acid preparations (as needed)
Primary outcome	To compare the proportion of patients with moderate-to-severe AD achieving EASI50 between each baricitinib dose group (2 and 4 mg) and placebo when treated daily for 16 weeks.
Secondary and exploratory outcomes	Secondary objectives: • To evaluate the absolute and percent change from baseline of the EASI with baricitinib compared to placebo • To evaluate the mean change from baseline compared to placebo for the SCORAD • To evaluate the mean change from baseline compared to placebo for the IGA • To assess QoL based on the DLQI • To assess itch based on the Itch NRS • To characterize the pharmacokinetics of baricitinib in patients with moderate-to-severe AD Exploratory objectives: • To evaluate changes in disease activity over the course of treatment • To evaluate changes in sleep quality over the course of treatment • To evaluate changes in nocturnal itch patterns over the course of treatment
Pre-planned subgroups	None
Duration of study and follow-up	The total study duration was 24 weeks, with a 4-week screening period, a 16-week treatment period and a 4-week post-treatment follow-up period.

Abbreviations: AD: atopic dermatitis; AE: adverse event; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; SCORAD: SCORing Atopic Dermatitis; IGA: Investigator's Global Assessment; NRS: numerical rating scale; QoL: quality of life; US: United States.
Source: JAHG Clinical Study Report.[8]

Table 4: Summary of adverse events in JAHG

	JAHG	
	PBO (N=[redacted])	4 mg BARI (N=[redacted])
Patients with ≥1 TEAE, n (%)	[redacted]	[redacted]
SAEs, n (%)	0 (0)	1 (3)
AEs leading to permanent discontinuation from study treatment, n (%)	[redacted]	[redacted]
AESIs, n (%)	[redacted]	[redacted]

Abbreviations: AE: adverse event; AESI: adverse event of special interest; BARI: baricitinib; PBO: placebo; SAE: serious adverse event; TEAE: treatment emergent adverse event.
Sources: JAHG Clinical Study Report,[8] ClinicalTrials.gov.[7]

Table 5: Summary of TEAEs affecting >3% of patients in the placebo and 4 mg baricitinib treatment groups in JAHG

TEAEs affecting >3% of patients, n (%)	PBO (N=[redacted])	4 mg BARI (N=[redacted])
≥1 TEAE	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Blood CPK increased	[redacted]	[redacted]
Dermatitis, atopic	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Cellulitis	[redacted]	[redacted]
Eczema	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]
Procedural pain	[redacted]	[redacted]
Somnolence	[redacted]	[redacted]
URTI	[redacted]	[redacted]
WBC count decreased	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; CPK: creatine phosphokinase; PBO: placebo; TEAE: treatment-emergent adverse event; URTI: upper respiratory tract infection; WBC: white blood cell.
Sources: JAHG Clinical Study Report.[8]

A10. In the CSRs for the BREEZE trials, discontinuation due to adverse events are classified in two categories:

a) Permanent discontinuation from study treatment due to adverse event (including death)

b) Discontinuation from study due to adverse event (including death)

Please explain the difference between the two categories of discontinuation due to adverse events.

Patients who discontinued investigational product for any reason were encouraged to remain in the study through Week 16 (Visit 8) and follow the regular visit schedule to provide the primary

efficacy and safety data including the post-treatment follow-up approximately 28 days after study drug discontinuation specified in the study protocol. These patients discontinued from the study treatment due to AE but their participation in the study continued and thus they were classified in the first category above.

In contrast, patients who permanently discontinued from the study due to AE (or other reasons, such as enrolment in another clinical trial, investigator decision, participation halted or patient decision) did not attend subsequent study visits for data collection. These patients discontinued the study in its entirety and were therefore classified in the second category above.

Indirect treatment comparisons

A11. PRIORITY: Please provide the full details of the indirect treatment comparisons (ITC) carried out, and all the electronic files required to reproduce all the ITC performed, including details of:

a) data used from each arm of each study and study subgroups (including raw data tables) to obtain the results presented in Tables 57- 70 of the CS;

b) details of methods used for data pooling within comparisons (i.e. when multiple studies of baricitinib or dupilumab were available); if possible, present forest plots of all within comparison meta-analyses.

c) the R script used to run the ITC (and any functions required), the R data and results files in electronic format – so that results can be reproduced.

The raw data tables corresponding to Tables 57–70 of the CS have been provided in the reference pack, which report the count data used from each arm of each study and study subgroups. Forest plots of all within comparison meta-analyses are also provided.

If more than one study was available for active treatment (baricitinib, dupilumab) versus placebo, then counts in active treatment and PBO arms were pooled. Meta-analysis was performed based on within study odds ratio (OR), relative risk (RR) and risk difference (RD). This included fixed effects and random effects approaches (using the Mantel-Haenszel Method). For identifying and quantifying heterogeneity, Tau^2 (the DerSimonian-Laird approach) and I^2 were calculated.[9] Q-statistics was used to test for heterogeneity. The indirect comparison between baricitinib and dupilumab via PBO was performed by an approach introduced by Bucher *et al.*, 1997.[10] Fixed effects models were preferred (see response to Question A12). Further details of the applied analyses techniques can be found in Borenstein *et al.*, 2009.[11]

The R script used to the run the ITC has also been provided. To re-run a specific analysis, count data from the respective Cheetah output tables must be entered into the two csv files that have been provided in the reference pack (one for baricitinib + placebo, one for dupilumab + placebo). The R script can then be run (after ensuring the data path matches the csv files), and the respective IC results will be printed to the console.

A12. Please justify why a fixed-effect model rather than random effect model was used for the ITC, with reference to the baseline characteristics of the included studies (Table 56 of the CS) and potential effect modifiers. The ERG recognises that there were only at most 2 studies within each comparison so statistical assessment of heterogeneity is not possible.

Too few studies were available for inclusion in the meta-analysis to produce reliable between study variations for random effects models. Fixed effects models can be used under such circumstances to describe the results.

A13. Please provide the following information regarding the baseline characteristics for the populations in the ITC, to supplement what is provided in Table 56 of the CS:

a) An additional table where the pooled data is broken down into its component trials, detailing the following:

    a. The baseline characteristics for the JAIY JAIN-like population alone: currently only the pooled JAIN + JAIY JAIN-like baseline characteristics are provided.

    b. The baseline characteristics for the JAHL JAIN-like population alone.

    c. The baseline characteristics for the JAHM JAIN-like population alone.

    d. If available, the baseline characteristics for the CHRONOS CAFÉ-like population alone.

    e. If available, the baseline characteristics for the SOLO 1 CAFÉ-like population alone.

    f. If available, the baseline characteristics for the SOLO 2 CAFÉ-like population alone.

Baseline characteristics for the JAIY JAIN-like, JAHL JAIN-like and JAHM JAIN-like populations are presented in Table 6. The baseline patient demographics and disease characteristics were broadly consistent with the JAIN trial.

Baseline characteristics for the CHRONOS, SOLO1 and SOLO2 CAFÉ-like populations were not available in the public domain.

Table 6: Baseline characteristics for JAIY, JAHL and JAHM JAIN-like populations

Intervention	JAIY JAIN-like		JAHL JAIN-like		JAHM JAIN-like	
	PBO+TCS	BARI 4 mg QD +TCS	PBO	BARI 4 mg QD	PBO	BARI 4 mg QD
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Male, %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Race						
White, %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Asian, %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other, %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline scores, mean (SD)						
EASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SCORAD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IGA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DLQI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Itch NRS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSA affected	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
POEM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Baseline characteristics have only been reported for the licensed doses of baricitinib (4 mg QD) and placebo.
[a] HADS anxiety and depression combined score presented.
Abbreviations: BSA: Body Surface Area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D: European Quality of Life-5 Dimensions; HADS: Hospital Anxiety and Depression Scale; IGA: Investigator's global assessment; NRS: Itch Numeric Rating Scale; POEM: Patient Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis; TCS: Topical corticosteroids; QD: once daily.

b) An additional table with the race, IGA, SCORAD, BSA, Pruritis NRS, POEM and HADS baseline characteristics for the JAIN + JAIY JAIN-like pooled and the JAHL/JAHM JAIN-like pooled populations. These have been reported separately for all four trials but not for the pooled populations.

Baseline characteristics for the JAIN + JAIY JAIN-like pooled and the JAHL/JAHM JAIN-like pooled populations are presented in Table 7.

Table 7: Additional baseline characteristics for JAIN + JAIY JAIN-like pooled and the JAHL/JAHM JAIN-like pooled populations

Intervention	JAIN + JAIY JAIN-like pooled		JAHL/JAHM JAIN-like pooled	
	PBO +TCS	BARI 4 mg QD +TCS	PBO	BARI 4 mg QD
N	[redacted]	[redacted]	[redacted]	[redacted]
Male, %	[redacted]	[redacted]	[redacted]	[redacted]
Race, (%)				

White	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Baseline scores, mean (SD)				
EASI	[redacted]	[redacted]	[redacted]	[redacted]
SCORAD	[redacted]	[redacted]	[redacted]	[redacted]
IGA	[redacted]	[redacted]	[redacted]	[redacted]
DLQI	[redacted]	[redacted]	[redacted]	[redacted]
Itch NRS	[redacted]	[redacted]	[redacted]	[redacted]
BSA affected	[redacted]	[redacted]	[redacted]	[redacted]
POEM	[redacted]	[redacted]	[redacted]	[redacted]
HADS[a]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D VAS	[redacted]	[redacted]	[redacted]	[redacted]

Baseline characteristics have only been reported for the licensed doses of baricitinib (4 mg QD) and placebo.
[a] HADS anxiety and depression combined score presented.
Abbreviations: BSA: Body Surface Area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D: European Quality of Life-5 Dimensions; HADS: Hospital Anxiety and Depression Scale; IGA: Investigator's global assessment; NRS: Itch Numeric Rating Scale; POEM: Patient Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis; TCS: Topical corticosteroids; QD: once daily.

c) An additional table of baseline characteristics for the JAIN + JAIN-like JAIY population stratified by region (Europe, Japan, rest of the world)

Baseline characteristics for the JAIN + JAIN-like JAIY population stratified by region are presented in Table 8.

Table 8: Baseline characteristics for the JAIN + JAIN-like JAIY population stratified by region

Intervention	Europe		Japan		Rest of the world	
	PBO+TCS	BARI 4 mg QD +TCS	PBO+TCS	BARI 4 mg QD +TCS	PBO+TCS	BARI 4 mg QD +TCS
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Male, %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Race, (%)						
White	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baseline scores, mean (SD)						
EASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SCORAD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

IGA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DLQI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Itch NRS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
POEM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D VAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Baseline characteristics have only been reported for the licensed doses of baricitinib (4 mg QD) and placebo.
Abbreviations: BSA: Body Surface Area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D: European Quality of Life-5 Dimensions; HADS: Hospital Anxiety and Depression Scale; IGA: Investigator's global assessment; NRS: Itch Numeric Rating Scale; POEM: Patient Oriented Eczema Measure; SCORAD: Scoring Atopic Dermatitis; TCS: Topical corticosteroids; QD: once daily.

A14. Page 126 of the CS states "a comparison based on IGA could not be conducted because the dupilumab and baricitinib clinical programmes used different IGA scales". Please attempt this comparison by either re-scaling one of the dupilumab or baricitinib IGA scales to make them comparable, or analysing the standardised mean difference (SMD). If an SMD analysis is possible, please use the population standard deviations on each of the measures if these are known (e.g. from external data) to standardise the IGA measures, rather than the study-estimated standard deviations.

Mean IGA scores (mean, mean change or percentage change from baseline) are not available for dupilumab from CAFÉ or the CAFÉ + CHRONOS CAFÉ-like and SOLO1/2 CAFÉ-like pooled populations; only the proportion of patients achieving IGA of 0 or 1 is reported. As such, it was not possible to conduct an indirect comparison using standardised mean differences in IGA.

A15. Separate ITC were carried out for each of the outcomes EASI 50, 75 and 90. Please confirm whether a joint (ordered categorical) model was considered for these measures considered (see NICE DSU technical support document 2, section 3.6 and example 6).

No joint (ordered categorical) analysis of EASI was considered. From a medical point of view, EASI75 is considered to be the most important EASI outcome (since it is the most sensitive in clinical practice). There were also few cases of EASI90, which may compromise the results if combined with other EASI measures. Therefore, only standalone results are presented.

A16. Please comment on the suitability of the assumption of a common relative treatment effect of baricitinib and dupilumab vs best supportive care (BSC) across the 3 cut-points (50, 75 and 90). If this assumption is reasonable, please carry out the ITC using this model and comment on model fit and the precision of the estimated relative effect of baricitinib compared to dupilumab. Please provide all code and data (including initial values) used to carry out this analysis.

An analysis combining categorical EASI endpoints (EASI50, 75 and 90) may be useful to maximise use of data in a situation where the underlying distribution of EASI scores is unknown. In this case, the mean change from baseline in absolute EASI score is available from the BREEZE-AD trials, and as such the underlying distribution of EASI scores is known. As a result, there is no additional value in conducting an analysis where categorical EASI measures are combined.

A17. Please explain the key differences in the results from the two ITC performed for combination therapy in sections B.2.9.4 and B.2.9.5. Specifically, whether results from the JAIN versus CAFÉ or results from the pooled JAIN + JAIN-like JAIY versus CAFÉ + CAFÉ-like CHRONOS patients are more relevant and whether there are any meaningful differences. Please also comment on how any differences in results should be interpreted for these analyses.

A summary of the results from the JAIN + JAIN-like JAIY versus CAFÉ + CAFÉ-like CHRONOS and JAIN versus CAFÉ ITCs is presented in Table 9. Overall, the results for EASI50 and EASI75 were comparable between both analyses: no reversal of treatment effect was observed, with [redacted] in all comparisons. In some comparisons, [redacted] in the pooled JAIN + JAIN-like JAIY versus CAFÉ + CAFÉ-like CHRONOS comparison [redacted] in the JAIN versus CAFÉ comparison, and this is likely attributable to the increased statistical power in the pooled comparison due to the larger population size.

The key differences between the results is the availability of outcomes: the composite outcome of EASI50 with DLQI ≥4-point improvement was not available for the CAFÉ trial alone, and EASI90 data were not available for the pooled CAFÉ + CAFÉ-like CHRONOS population. In alignment with TA534 where its use was based on clinical expert advice to the committee, the composite outcome was considered the most clinically relevant for use in the base case. This is further reflective of the consideration of patient quality of life alongside clinical signs and symptoms during treatment decision marking (discussed further in Sections A8 and B2).[2]

Table 9: Relative treatment effect of pairwise comparisons expressed as OR, RR and RD for JAIN + JAIN-like JAIY versus CAFÉ + CAFÉ-like CHRONOS and JAIN versus CAFÉ comparisons

	JAIN + JAIN-like JAIY versus CAFÉ + CAFÉ-like CHRONOS	JAIN versus CAFÉ
EASI50 and DLQI ≥4-point improvement		

OR (95% CI)	[redacted]	[redacted]
RR (95% CI)	[redacted]	[redacted]
RD (95% CI)	[redacted]	[redacted]
EASI50		
OR (95% CI)	[redacted]	[redacted]
RR (95% CI)	[redacted]	[redacted]
RD (95% CI)	[redacted]	[redacted]
EASI75		
OR (95% CI)	[redacted]	[redacted]
RR (95% CI)	[redacted]	[redacted]
RD (95% CI)	[redacted]	[redacted]
EASI90		
OR (95% CI)	[redacted]	[redacted]
RR (95% CI)	[redacted]	[redacted]
RD (95% CI)	[redacted]	[redacted]

* indicates statistical difference favouring dupilumab. All p-values were derived from fixed-effects model.
Abbreviations: CI: confidence interval; DLQI: dermatology life quality index; EASI: Eczema Area and Severity Index; OR: odds ratio: RD: risk difference; RR: risk ratio.

Section B: Clarification on cost-effectiveness data

Population

B1. The patient population in the company's cost-effectiveness analysis is patients with moderate-to-severe AD who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This is narrower than the population in the final scope issued by NICE, which is for patients who are candidates for systemic therapy but who have not necessarily failed on immunosuppressants. Page 22 of the CS describes the poor safety profiles of current systemic immunosuppressants, and as an oral drug it is possible that baricitinib may be preferred by clinicians. Please provide further justification as to why patients who have not failed on systemic immunosuppressants are not a relevant population for baricitinib.

As discussed in Section B.1.3.3 of the CS, the expected eligible patient population for baricitinib in UK clinical practice is adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This population is in line with the clinical positioning of baricitinib in current UK practice and the eligibility criteria for the BREEZE-AD4 (JAIN) trial.

The marketing authorisation for dupilumab is in line with baricitinib: "moderate to severe atopic dermatitis in adults who are candidates for systemic therapy". However, according to a panel of

clinical experts during an advisory board held in September 2017 reported in TA534, it was expected that dupilumab would be used in "moderate-to-severe patients previously optimised on topical treatments and for whom current systemic immunosuppressants had failed because of inadequate control due to contraindication, intolerance or they were otherwise medically inadvisable", based on the greatest unmet need in the treatment pathway, and this was confirmed by two clinical experts to the Committee.[2] It is anticipated that baricitinib will be used as an alternative to dupilumab in UK clinical practice, not as an alternative to first-line systemic immunosuppressants.

Furthermore, as discussed in the response to Question B18, there is limited evidence for the efficacy of systemic immunosuppressants in AD, and in the absence of available data, conservative assumptions regarding efficacy may have been required to facilitate a comparison. In such a scenario demonstrating cost-effectiveness versus systemic immunosuppressants would be challenging.

Response

B2. Please provide the clinical rationale supporting the use of a combination of secondary trial outcomes (i.e. EASI50 and DLQI ≥4) to define clinical response in the model, referring in particular to the significance of this combination of outcomes in clinical decision making.

The composite endpoint of EASI50 with a DLQI improvement of four or more points (ΔDLQI ≥4) was the preferred option of the Appraisal Committee during the dupilumab NICE appraisal (TA534).[2]

In TA534, the composite endpoint was used for the economic analysis based on what clinicians considered to be clinically meaningful changes in outcomes, while the CAFÉ and CHRONOS trial endpoints were dictated by the requirements of regulatory agencies. The clinical experts to the Committee explained that EASI75 and IGA 0/1, the endpoints of the clinical trials, are difficult to achieve in practice, and that the composite endpoint was more sensitive to changes in treatment outcomes and more clinically relevant.

As such, the composite endpoint was chosen as the response endpoint for the base case analysis. A scenario analysis was conducted where response was defined based on achieving EASI75, and results were similar to the base case analysis.

B3. PRIORITY: Please provide details on how the response rates presented in Tables 83 and 84 of the CS were calculated from the results of the ITC. Specifically, please include details of the source of all values and calculations used to apply the relative treatment effect estimated by the ITC.

An adjusted dupilumab response is derived by applying the RD for dupilumab versus placebo to the placebo response from the BREEZE-AD trial data. For example, in the base case analysis of EASI50 plus ΔDLQI≥4 in the JAIN plus JAIY JAIN-like pooled population, the response for placebo was 31.25% (as reported in Table 83 of the CS, and shown in Table 10). The RD for the comparison of dupilumab versus placebo for the CAFÉ plus CHRONOS CAFÉ-like population

was [redacted]%. Thus, the response rate for dupilumab used in the model was 79.25% (31.25% + [redacted]%). For the respective SE the same precision as in the dupilumab trial is assumed.

Please note that the SE for dupilumab was incorrectly reported in the CS and in the original model submitted, and has been corrected in Table 10. Please ensure that the correction to this value is included in the ERG base case and any scenario analyses.

Table 10: Response rates for EASI50 + ΔDLQI≥4 employed in the base case analysis

	Response probability, % (SE%)
Baricitinib	48.99 (4.09)
Dupilumab	79.25 (4.07)
BSC	31.25 (3.86)

Abbreviations: BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; SE: standard error.

B4. PRIORITY: The draft SmPC supplied suggests that initial assessment of response should be carried out at 12 weeks. This does not align with the bulk of the trial evidence presented in the CS.

a) Please explain this difference and comment on your expectation of when response on baricitinib will be assessed in clinical practice.

b) Please present key results (IGA, EASI50, EASI 75, EASI50/DLQI ≥4) for the JAIN and JAIN-like JAIY population assuming a 12-week assessment period.

c) Please present a scenario analysis (and add model functionality) using a 12-week assessment period for baricitinib.

As per usual UK clinical practice, it is expected that the majority of clinical assessments will be carried out at 16 weeks. This expectation was confirmed to be valid by a panel of expert dermatologist advisors.

The draft SmPC does state that consideration should be given to discontinuing treatment in patients who show no evidence of therapeutic benefit before this time. This possibility was discussed with UK expert dermatologists at a recent advisory board meeting, and the advisors felt that an early clinical assessment of efficacy would risk discontinuing treatment in patients who would go on to respond. Therefore, they concluded that they would be evaluating patients at Week 16 as per their usual clinical practice.

Discontinuation

B5. PRIORITY: Please clarify why conditional response rates at 52 weeks for the EASI50 response criteria scenario were assumed to be equal to the EASI75

rate, when separate data are available from CHRONOS?[2] (Table 86 of CS). Please present a scenario analysis using the EASI50 data.

This was not intentional; the model has been updated to include the data specific to EASI50 (Week 52 response rate conditional on response at Week 16: [redacted]; annual discontinuation rate: [redacted]). Fully incremental results for the base case population (JAIN + JAIY JAIN-like patients) are presented in Table 11. The impact of this update is to make baricitinib slightly more cost-effective versus both BSC and dupilumab.

Table 11: Scenario analysis using EASI50 discontinuation rates for EASI50 response

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£15,247	£15,247	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£75,187	£224,395[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

B6. PRIORITY: Please clarify why the 16 to 52-week discontinuation rate for BSC was based on continued EASI75 response in the base-case analysis, rather than EASI50 + DLQI≥4. If this was intentional please provide justification, and provide a scenario in which the response criteria for BSC are the same as for baricitinib and dupilumab.

This was not intentional, and the model has been updated to include a conditional response rate of [redacted] for BSC for the composite outcome. Fully incremental results for the base case population (JAIN + JAIY JAIN-like patients) are presented in Table 12. The impact of this update is negligible for baricitinib versus both BSC and dupilumab and did not affect the conclusions presented in the company submission.

Table 12: Scenario analysis using conditional response rates of BSC for EASI50 + DLQI≥4

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,996	£17,996	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£89,048	£203,968[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; DLQI: Dermatology Life Quality Index; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

B7. PRIORITY: In recognition of the proportion of patients who discontinued treatment before Week 16 due to adverse events and other reasons in the trial, please provide model functionality to allow patients to discontinue treatment before Week 16.

The proportion of patients who discontinued treatment prior to Week 16 in the placebo and 4 mg baricitinib arms in the JAIN + JAIY JAIN-like, JAIN and JAIN (Europe only) populations are displayed in Table 13. Corresponding data for dupilumab are not available.

Table 13: Discontinuation rates to Week 16 in the placebo and 4 mg baricitinib arms

Treatment disposition to Week 16, n (%)	JAIN + JAIY JAIN-like		JAIN		JAIN (Europe only)	
	PBO+TCS	BARI 4 mg QD +TCS	PBO	BARI 4 mg QD	PBO	BARI 4 mg QD
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Completed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; PBO: placebo; QD: once daily; TCS: topical corticosteroids.

The structure of the model is not equipped to process discontinuation prior to response assessment. Response estimates used in the model calculations consider those who did not remain on treatment until the Week 16 assessment as non-responders, so no impact to the proportion of patients in the "maintenance" state would be seen if such a change were made in the model calculations. The impact of modifying the model to accommodate discontinuation prior to the Week 16 assessment on the comparison of baricitinib versus BSC is structurally limited to the 16 week period prior to response assessment. The marginal costs for patients in baricitinib and BSC treatment arms who discontinue prior to Week 16 is presented in Table 14. Assuming a constant rate of discontinuation, the patients who discontinue would receive a different cost of active treatment and follow-up care (baricitinib patients would move to BSC with lower total costs [£[redacted] per week], BSC patients would move to non-response with higher total costs [£[redacted] per week]) for on average 8 weeks. The model applies the same baseline utility to patients during the induction phase, so including early discontinuation is not expected to impact total accrued utility or marginal utility.

Table 14: Marginal costs for patients in baricitinib and BSC treatment arms who discontinue prior to Week 16

Change in health state due to discontinuation	BSC -> non-response	Baricitinib -> BSC
Annual cost baricitinib	[redacted]	[redacted]
Annual cost BSC	[redacted]	[redacted]
Annual cost non-response	[redacted]	[redacted]
Change in annual cost	[redacted]	[redacted]
Change in cost per week	[redacted]	[redacted]
Change in cost for 8 weeks	[redacted]	[redacted]

Abbreviations: BSC: best supportive care.

Therefore, the maximum potential change in accrued and incremental costs can be estimated by comparing the cost patients receive in induction versus the cost they would receive if they

discontinue prior to the end of induction. Patients starting on baricitinib would switch to a lower cost treatment (BSC) while patients starting on BSC would switch to a higher cost in non-response, which would benefit baricitinib in the comparison. Based on the proportions of patients who discontinue prior to 16 weeks, and the marginal treatment/follow-up costs for those patients who discontinue, the incremental cost comparing baricitinib and BSC is expected to be reduced if the discontinuation prior to Week 16 is added to the model, as shown in Table 15. Given the impact of this change is small for comparison of baricitinib and BSC, and this functionality could not be added for the dupilumab treatment arm given the lack of data, the functionality to account for discontinuation prior to Week 16 was not added to the model.

Table 15: Net change to accrued and incremental cost

Trial population	JAIN+JAIY-JAIN like		JAIN		JAIN EU	
Treatment group	BSC	Baricitinib	BSC	Baricitinib	BSC	Baricitinib
Discontinued Treatment before Week 16: %	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Estimated maximum change in accrued cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Net change to incremental cost	[redacted]		[redacted]		[redacted]	

Abbreviations: BSC: best supportive care.

B8. PRIORITY: Please present a scenario (and model functionality) where treatment discontinuation to Week 52 is based on the rate observed between weeks 16 and 24 in the JAIN and JAIN-like JAIY patients.

No Week 24 discontinuation data are available for the JAIN-like JAIY patients. While some discontinuation data to Week 24 in the JAIN trial are available, their use in the model would necessitate extrapolation to Week 52 dependent on one single time point (i.e. 16–24 weeks). It is reasonable to consider this extrapolation to be unreliable and likely misleading with respect to the true discontinuation rate associated with baricitinib to Week 52, and would likely result in an overestimate of the discontinuation rate for baricitinib. From an economic standpoint, overestimating the discontinuation rate from Week 16 to 52 in the model risks underestimating the total cost of baricitinib treatment and biasing the model in favour of baricitinib. As such, it was considered more appropriate to use an assumption of equivalence to dupilumab for the discontinuation rate up to Week 52 (i.e. based on the conditional probability of response at Week 52 given response at Week 16 in the dupilumab submission [TA534]).[2]

B9. PRIORITY: The ERG notes that long-term data on the effectiveness of baricitinib and adherence to treatment is available from the JAHN extension study. Please justify why this study was not used and the CHRONOS data were favoured. To allow comparison with the JAIN and JAIN-like JAIY patients, please provide data from the JAHN trial on discontinuation between T_0 and T_8 (16 to 24 weeks of treatment).

The long-term data currently available from the JAHN extension study are from monotherapy-treated patients whereas the economic model is informed by combination therapy patients. Given

these differences in the intervention received, the company does not consider the population currently available in the JAHN extension study to be relevant for the target population in the economic model, rendering a comparison of the study populations unfeasible. Therefore, the company has not considered the JAHN extension study as a source of the long-term baricitinib efficacy and adherence data.

B10. PRIORITY: Please provide the following information from the JAHN extension study.

a) **The proportion of patients on baricitinib 4mg who achieved EASI 50/DLQI ≥4 response at T_0 (16 weeks of treatment). Please stratify data according to response status at T_0 (as defined in the contributing trials).**

b) **The proportion of patients on baricitinib 4mg who maintained EASI 50/DLQI ≥4 response at T_{36} (52 weeks of treatment). I.e. the conditional probability of maintaining response. Please stratify data according to response status at T_0 (as defined in the contributing trials).**

c) **The proportion of patients on baricitinib 4mg who discontinued treatment between T_0 and T_{36}, providing reasons for discontinuation. Please stratify data according to response status at T_0 (as defined in the contributing trials).**

d) **The proportion of patients on placebo who achieved EASI 50/DLQI ≥4 response at T_0 (16 weeks of treatment), stratifying according to response status at T_0 (as defined in the contributing trials) and the arm patients were re-randomized to (non-responders only).**

e) **The proportion of patients on placebo 4mg who maintained EASI 50/DLQI ≥4 response at T36 (52 weeks of treatment). I.e. the conditional probability of maintaining response. Please stratify according to response status at T0 and the arm patients were re-randomized to (non-responders only).**

f) **The proportion of patients on randomised to placebo in the contributing studies) who discontinued treatment between T_0 and T_{36}, providing reasons for discontinuation. Please stratify according to**

response status at T0 and the arm patients were re-randomized to (non-responders only).

g) The proportion of patients reporting symptom flares and/or using rescue medication between T0 and T36. If known, please also provide details of the type of rescue medications used, broken down into the categories in Table 97 of the CS.

The JAHN study is still ongoing and at the current time, insufficient results from patients recruited following completion of BREEZE-AD7 (JAIY) are available to inform conclusions on the long-term efficacy of baricitinib in combination-treated patients. Data from combination trials BREEZE-AD4 (JAIN) and BREEZE-AD7 (JAIY) will be available for later data cuts and are expected to become available in November 2020. The data discussed below are based on a data cut-off date of 13th December 2019 and include monotherapy patients from the BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) trials only. Therefore, these data will be subject to change following later data cuts from the ongoing JAHN trial.

(a), (b), (d) and (e): The proportion of patients on baricitinib 4 mg and placebo who achieved EASI50 + DLQI ≥4 response at Week 0 of Study JAHN (i.e. 16 weeks of treatment) and who maintained this response to Week 36 (i.e. 52 weeks of treatment) are presented in Table 16, stratified according to response status at Week 0. For non-responders, data were stratified according to response status at Week 0 and the arm patients were re-randomised to.

(c) and (f): The proportion of patients on baricitinib 4 mg and placebo who discontinued treatment between Week 0–36 including reasons for discontinuation are presented in Table 17, stratified according to response status at Week 0. For non-responders, data were stratified according to response status at Week 0 and the arm patients were re-randomised to.

(g) The proportion of patients reporting symptom flares and/or using rescue medication between Week 0–36 was not collected.

Table 16: EASI50 and DLQI≥4 response at Week 0 (Week 16 of originating studies) and Week 36 (Week 52 of originating studies) conditional on Week 0 in the JAHN study for responders/partial responders and for non-responders (monotherapy patients from JAHL/JAHM only)

Responder	Outcome	BARI 4 mg to BARI 4 mg	PBO to BARI 2 mg	PBO to BARI 4 mg	PBO to PBO
Y	IGA Response Week 0, N	[redacted]	[redacted]	[redacted]	[redacted]
	Composite endpoint Week 0: Yes, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
	Composite endpoint at Week 36: Yes, conditional on Week 0: Yes, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
N	IGA Response Week 0, N	[redacted]	[redacted]	[redacted]	[redacted]
	Composite endpoint Week 0: Yes, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
	Composite endpoint at Week 36: Yes, conditional on Week 0: Yes, n (%)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; IGA: Investigator's Global Assessment; N: no; PBO: placebo; Y: yes.

Table 17: Discontinuation before or at Week 36, for patients in the JAHN study, responders/partial responders and for non-responders (monotherapy patients from JAHL/JAHM only)

Responder	Outcome	BARI 4 mg to BARI 4 mg	PBO to BARI 2 mg	PBO to BARI 4 mg	PBO to PBO
Y	IGA Response Week 0, N	[redacted]	[redacted]	[redacted]	[redacted]
	Discontinuation before or at Week 36 (Week 52 original study), n (%)	[redacted]	[redacted]	[redacted]	[redacted]
	Adverse event	[redacted]	[redacted]	[redacted]	[redacted]
	Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]
	Lost to follow-up	[redacted]	[redacted]	[redacted]	[redacted]
	Physician decision	[redacted]	[redacted]	[redacted]	[redacted]
	Withdrawal by subject	[redacted]	[redacted]	[redacted]	[redacted]
N	IGA Response Week 0, N	[redacted]	[redacted]	[redacted]	[redacted]
	Discontinuation before or at Week 36 (Week 52 original study), n (%)	[redacted]	[redacted]	[redacted]	[redacted]
	Adverse event	[redacted]	[redacted]	[redacted]	[redacted]
	Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]
	Lost to follow-up	[redacted]	[redacted]	[redacted]	[redacted]
	Physician decision	[redacted]	[redacted]	[redacted]	[redacted]
	Withdrawal by subject	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; IGA: Investigator's Global Assessment; N: no; PBO: placebo; Y: yes.

HRQoL

B11. In line with the ICER report identified as part of the cost-effectiveness review,[12] please present a scenario (and add model functionality) in which separate utilities are modelled for patients with moderate and severe disease.

The anticipated indication for baricitinib is for the treatment of patients with moderate-to-severe atopic dermatitis. As discussed in Question A8, it was not considered clinically appropriate to use a single endpoint such as EASI score to conduct subgroup analyses, since these measures may not reflect all aspects of moderate or severe disease. As such, while separate utilities could be generated for subgroups defined using these measures, they may not accurately reflect utility for patients with moderate and severe disease in clinical practice. Additionally, no efficacy data are available for dupilumab in the moderate versus severe subgroups and this split was not performed in the TA534 appraisal.[2] Finally, even if appropriate utility values could be generated, the company considers that a scenario in which only utilities differ between moderate and severe subgroups would be inadequate as it would be expected that response assessment inputs, resource use inputs, flare treatment costs, adverse event frequencies, and estimated costs of non-response could all differ between moderate and severe patients, if such a firm distinction into subgroups could reliably be made. Therefore, the company considers the separation of the

model population into moderate and severe to be both infeasible and not relevant to the decision problem.

B12. The ERG is not familiar with the approach taken by the company to age adjust utilities. Please describe how the formula used to implement age adjustment to utility was derived (1.0708-0.0044*age), and explain why this was used instead of the standard method of implementing age related decrements from Ara and Brazier.[13]

A declining health utility with age of 0.004 per year was used to deduct a constant decrement in health utility per year. The equation applied in the model was estimated using data from the general UK population, as presented by Ara et al., to which a linear trend adjusted for age-specific weights was fitted. The formula was added as a multiplicative approach as recommended by NICE DSU guidance (TSD 12) and its application was in alignment with the approach taken in the dupilumab submission (TA534).[2]

B13. PRIORITY: It appears from the revised company submission that only the JAIN and JAIY JAIN -like patients were used to estimate the utility values used in the model. The company, however, also provides supplemental values for JAIN-like patients in the JAHL and JAHM studies.

a) **Please comment on why these data were not also used to generate utility values.**

The utility data in the model are derived from the population analysed: the utility values applied in the analysis of the JAIN-only population are derived from JAIN-only patients, and in the JAIN + JAIY JAIN-like population utilities derived from the JAIN + JAIY JAIN-like population are applied.

Patients in the JAHL and JAHM trials received monotherapy and based on this difference in intervention, it was considered appropriate for separate analyses to be run for this population. Results from the JAHL/JAHM JAIN-like population have been presented as separate scenario analysis.

b) **Please provide an additional scenario analysis in which utility values are generated from all JAIN-like patients i.e. those from JAIN, JAIY, JAHL and JAHM.**

In the base case analysis, baricitinib is modelled to be used in combination with TCS, as this is considered to represent typical AD management in the UK.[14] As such, the most relevant evidence to the decision problem for the efficacy and safety of baricitinib is the JAIN and JAIY JAIN-like pooled population. The JAHL and JAHM trials provide supportive evidence, but included baricitinib as monotherapy only, which is not in line with the expected use of baricitinib in clinical practice. However, for transparency, an analysis was conducted using data from all JAIN-like patients across the BREEZE-AD trials (i.e. those who had a history of intolerance or inadequate response to ciclosporin). The analysis was conducted in line with the methodology presented in Section B.3.4.4 of the CS, and the results are presented in Table 18.

Table 18: EQ-5D-3L utility score at baseline and Week 16 for all JAIN-like patients by EASI50 with ΔDLQI ≥4 response category at Week 16

EASI50 with ΔDLQI ≥4 (Week 16)	Baseline EQ-5D-3L[a]			Change in EQ-5D-3L (baseline to Week 16)[b]	
	Number of patients	Mean	Standard deviation	LS Mean	95% CIs
Overall	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Observed values. [b]from mixed model. Number of observations used = 3275.
Abbreviations: CI: confidence interval; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; LS: least squares; SD: standard deviation.

The results for the scenario analysis using pooled utilities from all JAIN-like patients are presented in Table 19. As in the base case analysis, the baseline utility value was applied in the induction and non-response states and the utility value for those with an EASI50 with ΔDLQI ≥4 response at Week 16 (i.e. baseline plus LS mean change to Week 16) is applied in the maintenance state. The effect of this change is to increase ICERs due to a decrease in incremental QALYs. However, the conclusion regarding the cost-effectiveness of baricitinib versus BSC and dupilumab is unaffected in this comparison as both comparisons remain cost-effective at a willingness-to-pay threshold of £30,000.

Table 19: Scenario analysis applying pooled utilities for all JAIN-like patients

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£25,092	£25,092	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£124,256	£284,654[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

B14. PRIORITY: Please provide further details on the mixed model used to generate the regression coefficients reported in Table 90, including the number of observations at each time point. Please also provide an explanation of how the values reported in Table 90 are used to calculate the values reported in Table 91.

The SAS procedure MIXED was used to generate the utility models, including EASI50 with ΔDLQI ≥4 at Week 16 as a fixed effect. Further fixed effects were week, age and baseline EQ-5D values. Parameters in Table 90 of the CS were provided to show which of the main effects were different from zero and does not contain all information required to calculate the LS-mean results given in Table 91 of the CS.

It is not a trivial task to derive LS means as provided by SAS. Input would be needed, which is calculated in intermediate steps of the SAS procedure. The LSMEANS statement of the MIXED procedure computes least-squares means (LS-means) of fixed effects. LS-means are *predicted*

population margins, that is, they estimate the marginal means over a balanced population. In a sense, LS-means are to unbalanced designs as class and subclass arithmetic means are to balanced designs. Each LS-mean is computed from the coefficient matrix associated with the least-squares mean estimate of the fixed-effects parameter vector.[15]

The numbers of patients included in the utility model at timepoints between Week 1 and Week 16 are shown in Table 20. All observed values across patients receiving all baricitinib dose groups and placebo were included in the analysis.

Table 20: EQ-5D-3L (health state score) change from baseline summary statistics by visit

Visit	JAIN plus JAIY JAIN-like patients, n
Week 1	[redacted]
Week 2	[redacted]
Week 4	[redacted]
Week 8	[redacted]
Week 12	[redacted]
Week 16	[redacted]

Abbreviations: EQ-5D: EuroQol 5 dimensions.

As the outcome was change from baseline, Week 0 is not presented. The number of patients informing baseline utility is reported in Table 91 of the CS.

B15. Priority: The coefficients reported in Table 91 suggest non-responders have superior HRQL at week 16 compared with responders. Please explain this apparent anomaly or provide details of why this is not the case.

This observation by the ERG of the pattern in the data is correct. The sample size in the responder group (N=[redacted]) was much lower than the non-responder group (N=[redacted]). As such, since the analysis was post-hoc and thus non-randomised, these groups may differ in some potentially unmeasured respects that could have led to biased results, resulting in a higher utility value for non-responders. For this reason, and as discussed in Section B.3.4.4 of the CS, non-responders were assigned baseline utility in the economic model based on advice from clinical experts and in line with the assumptions in the US ICER model, given that it was not deemed clinically plausible for non-responders to have superior utility compared with responders. A utility benefit for those with EASI50 and ΔDLQI≥4 response was accepted in the dupilumab appraisal (TA534).[2]

B16. The BSC discontinuation rate modelled for Years 2-5+ was based on rescue therapy frequency + study withdrawal for BSC in CHRONOS (57.0%), but all-cause discontinuation for baricitinib/dupilumab, which was much lower than rescue.

a) Please comment on the clinical validity of assuming permanent loss of efficacy when BSC patients are rescued, but continuing response on dupilumab and baricitinib.

The annual probability of discontinuation for the second and subsequent years for dupilumab and for baricitinib was based on all-cause discontinuation for consistency with TA534.

While the clinical benefit of BSC is unknown, it is assumed that BSC consists of treatment modalities which have previously failed, given that the population of patients modelled are contraindicated or intolerant to systemics. Additionally, it is unlikely that results for BSC from a clinical trial would be replicable outside the trial setting as patients are more likely to show compliance with a topical treatment regimen in a clinical trial setting than outside of it. Therefore, it is reasonable to assume limited clinical efficacy of BSC in these patients and that the efficacy rates in usual clinical practice would be lower than those observed in clinical trials. Following consideration of these factors, 57% was considered to be the best estimate available for BSC discontinuation rate. This higher rate of discontinuation was considered to represent a loss of efficacy, with a transition to the "non-response" state within the model. Long term discontinuation rates are applied to all competing treatments with a similar intent: *i.e.*, a state transition associated with discontinuation represents "loss of efficacy" in the long term and associated loss in incremental utility. The model assumes patients will not continue to use baricitinib or dupilumab if efficacy has been lost, rather than a case where patients persist in using a medication which no longer provides benefit.

b) In TA534, the probability of a sustained HRQoL response was modelled for Years 2-5+ based on time to rescue/stopping study projections from CHRONOS. Please present a scenario (and add model functionality) in which the probability of a sustained HRQoL response is modelled for BSC.

In an approach supported by clinical experts at a recent advisory board, and in alignment with the ICER model, a discontinuation rate was applied instead of a HRQoL benefit reduction from treatment. The company considered it to be a reasonable assumption that patients will stop complying with an arduous topical treatment regimen, thus losing HRQoL, and that they would not continue with the regimen following loss of the HRQoL benefit. For this reason, the discontinuation rate was considered the more reasonable approach for BSC.

However, scenarios where the probability of a sustained HRQoL response is modelled for BSC has been explored in line with in TA534; one where the discontinuation rate for BSC is equal to that of dupilumab and utility decreases over time for all treatments, and another where the discontinuation rate for BSC is equal to that of dupilumab and utility remains constant over time. The results of these analyses for the base case settings (JAIN + JAIY JAIN-like population for the composite endpoint) are presented in Table 21 and Table 22, respectively. Details of the modifications made to the model are presented in the Excel file provided in the reference pack. The impact of these scenarios is small for baricitinib versus both BSC and dupilumab and did not affect the conclusions presented in the company submission.

Table 21: Scenario analysis with lower discontinuation probability for BSC and loss of utility applied over time in maintenance for all treatments

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£20,005	£20,005	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£96,267	£220,020[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

Table 22: Scenario analysis with lower discontinuation probability for BSC and constant utility over time in maintenance for all treatments

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£20,475	£20,475	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£98,162	£222,989[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

Comparators

B17. PRIORITY: On page 136 of the CS, the company comments that treatment sequences were not considered relevant to the decision problem and UK practice.

a) Please provide a fuller explanation of why the company considers treatment sequences involving both baricitinib and dupilumab irrelevant including reference to the clinical plausibility of using both agents in a sequence.

b) Please present appropriate fully incremental analysis considering treatment sequences in which both baricitinib and dupilumab appear.

Treatment sequences were not explored as part of the company submission because they are not in line with the anticipated positioning for baricitinib in UK clinical practice, nor the population included in the key trial exploring the efficacy and safety of baricitinib (BREEZE-AD4 [JAIN]). Feedback from clinical experts indicates that baricitinib would be used as a fifth-line therapy in adult patients with moderate-to-severe AD who are candidates for systemic therapy and who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. Baricitinib would be used as an alternative to dupilumab or as an alternative to BSC for those patients for whom dupilumab is not recommended or contraindicated.

Treatment sequences are not relevant in patients for whom dupilumab is not recommended or contraindicated, whose only remaining treatment option is BSC. For patients who are eligible to receive dupilumab, baricitinib would not be considered after dupilumab in the treatment pathway, since there is very limited evidence from the BREEZE clinical trials for the efficacy and safety of baricitinib in patients who have received prior dupilumab; across all treatment arms, [redacted] ([redacted]) patients in the BREEZE-AD4 trial, [redacted] ([redacted]) patients in BREEZE-AD7, [redacted] ([redacted]) patients in BREEZE-AD1 and [redacted] ([redacted]) patients had received prior dupilumab. Similarly, there is no

evidence for the efficacy and safety of dupilumab in patients who received prior baricitinib, and it would not be appropriate to assume that the response to dupilumab for patients who had discontinued baricitinib (e.g. due to inadequate disease control) would be the same as the response rates reported in the CAFÉ and CHRONOS trials. As such, no further incremental analysis considering treatment sequences has been presented.

B18. PRIORITY: The NICE scope lists immunosuppressive systemic agents as a relevant comparator, these are however, not considered in the company submission.

a) Please justify this decision. While the modelled population is patients who have failed at least one systemic agent, the availability of several systemic agents implies that these could be used in this population.

As discussed in Section B.1.3.3 of the CS, baricitinib is being positioned as a fifth-line therapy; the expected eligible patient population for baricitinib in UK clinical practice is adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control, in line with TA534. This population is in line with the clinical positioning of baricitinib in current UK practice and the eligibility criteria for the BREEZE-AD4 (JAIN) trial, reflecting the highest unmet need in UK clinical practice.

The only systemic immunosuppressant therapy currently licensed for AD in the UK is ciclosporin.[6] However, other systemic therapies may also be used off-label in UK clinical practice, such as methotrexate, azathioprine and mycophenolate mofetil. Accordingly, in TA534, the ERG's clinical expert noted that azathioprine or methotrexate may be tried if ciclosporin fails, despite the fact that they are not licenced for this condition. However, a clinical expert to the Committee explained that, in practice, patients are unlikely to be offered every fourth-line treatment option available before being offered dupilumab given the toxicity risks of systemic therapies, but that patients were likely to have had at least 1 systemic therapy. The committee concluded that it would appraise dupilumab for moderate to severe atopic dermatitis, compared with best supportive care, after other systemic therapies.

The positioning of baricitinib is in line with dupilumab in TA534, and thus the relevant comparators considered in this appraisal are dupilumab or BSC in patients for whom use of dupilumab is not recommended or contraindicated.

b) Please present scenario analyses (and include model functionality) in which systemic agents are considered as a potential comparator.

As discussed in the response to Question B18a, systemic agents are not considered relevant comparators for this appraisal. As discussed in Section B.2.1 of the CS, an SLR was conducted to identify all relevant clinical evidence on the efficacy and safety of baricitinib and potential comparators for the treatment of adults with moderate-to-severe AD. 6 studies were identified that investigated the efficacy of ciclosporin, methotrexate or azathioprine in patients with moderate-to-severe AD. However, none of these studies included a common comparator for the BREEZE-AD trials, and all were excluded from the ITC (Table 23). Given the lack of evidence for

the relative efficacy of systemic therapies in the target population for this submission, a scenario analysis could not be conducted where systemic agents were considered as a potential comparator.

Table 23: Summary of studies investigating systemic therapies included and excluded from the ITC

Trial name	Patient population	Interventions	Patients randomised	Phase	Included (Yes/No)	Reason for exclusion
Wahlgren 1990[16]	-	• Ciclosporin A (5 mg/kg/day) • Placebo	10	-	No	Short treatment duration
Zurbriggen 1999[17]	Severe AD	• Ciclosporin A (Sandimmun) • Ciclosporin A microemulsion (Neoral)	14	II	No	Short treatment duration
Czech 2000[18]	Severe AD	• Ciclosporin A Neoral (150 mg/day) • Ciclosporin A Neoral (300 mg/day)	106	-	No	Duration only 8 weeks, including a dose reduction in case of response at 2 weeks; after 8 weeks there was an open-label follow-up of 4 weeks
Granlund 2001[19]	AD	• Ciclosporin A Neoral (1–4 mg/kg/d) • UVAB	71	-	No	Intermittent administration: treatment period of 8 weeks (treatment phase) followed by a period of only topical treatment (remission phase), no common comparator
Goujon 2017[20]	Moderate to severe AD	• Ciclosporin A (2.5–5 mg/kg/d) • Methotrexate (oral) (15–25 mg/wk)	97	III	No	After 8 weeks of treatment, these doses were, respectively, increased to 25 mg/wk and 5 mg/kg of body weight/d for 16 more weeks in the patients who did not achieve 50% reduction in the SCORAD index (SCORAD 50), no common comparator
MAcAD[21, 22]	Severe AD	• Methotrexate (10–22.5 mg/wk) • Azathioprine (1.5–2.5 mg/kg/d)	42	Follow-up phase	No	Short treatment duration, not connected

Abbreviations: AD: atopic dermatitis; ITC: indirect treatment comparison; SCORAD: SCORing Atopic Dermatitis.

Adverse events (AE)

B19. PRIORITY: Please confirm that the source of AE rates used in the model is the JAIN and JAIN-like JAIY population, and detail how adverse event rates were selected for inclusion in the model.

It is correct that the rates of adverse events (AEs) have been derived from the respective population – for example, in the base case, they have been derived from the JAIN and JAIY JAIN-like pooled population. Adverse event rates were selected based on their frequency and cost relevance from the baricitinib trials.

B20. PRIORITY: Please explain the differences between the incidence of allergic conjunctivitis reported in Table 74 of the CS for baricitinib and placebo, and the rates used in the economic model (Table 88). If appropriate, please provide a revised scenario analysis in which the rate of adverse events is consistent with the observed trial data.

All serious adverse events (SAEs), including those not marked as on treatment, are included in Table 74 of the CS, whereas Table 88 considers only AEs on treatment. Initial data output from the populations analysed in the model did not include the frequency of allergic conjunctivitis because it was not flagged in the data analysis as a "treatment emergent event." Revised TEAE probabilities for the populations included in the model are presented in Table 24.

Table 24: Revised TEAEs between baseline and Week 16

TEAE, n (%)	JAIN + JAIY JAIN-like			JAIN			JAIN (Europe only)		
	PBO (N=[redacted])	4 mg BARI (N=[redacted])	RR vs PBO	PBO (N=[redacted])	4 mg BARI (N=[redacted])	RR vs PBO	PBO (N=[redacted])	4 mg BARI (N=[redacted])	RR vs PBO
Injection site reaction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Allergic conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infectious conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral herpes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
URTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; PBO: placebo; RR: relative risk; TEAE: treatment-emergent adverse event; URTI: upper respiratory tract infection.

Annual probabilities for AEs were then recalculated for baricitinib as per Section B.3.3.4 of the CS; the AE probabilities for the base case population (JAIN + JAIY JAIN-like) are presented in Table 25.

Table 25: Revised annual AE probabilities for baricitinib (JAIN + JAIY JAIN-like)

Event	Original model		Revised	
	Induction	Annual	Induction	Annual
Injection site reaction	[redacted]	[redacted]	[redacted]	[redacted]
Allergic conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]
Infectious conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]
Oral herpes	[redacted]	[redacted]	[redacted]	[redacted]
Upper respiratory tract infections	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; URTI: upper respiratory tract infection.

A scenario analysis was conducted using these revised inputs for baricitinib AEs, and produces results that are not significantly different from the base case (Table 26).

Table 26: Scenario analysis applying revised AE probabilities for baricitinib

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,897	£17,897	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,596[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone vs. baricitinib; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: AE: adverse event; BSC: best supportive care; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

B21. PRIORITY: Related to the above, please justify why the rate of allergic conjunctivitis, infectious conjunctivitis and oral herpes for BSC [redacted] baricitinib combination therapy given that combination therapy includes BSC.

Sources for the treatment specific data report different rates for these events. In the version of the model originally submitted, the frequency of adverse events for the BSC group were taken from the dupilumab submission, while frequencies for baricitinib were taken from the baricitinib trial data. As an attempt to standardise the sources, a scenario has been conducted using the original submission data for the frequency for BSC and dupilumab (i.e. from TA534), and revised AE frequencies for baricitinib calculated by applying the relative risk versus BSC based on the observed TEAEs for the JAIN+JAIY-JAIN trial population (see Table 24). The adjusted AEs for baricitinib are presented in Table 27.

Table 27: Adjusted AE frequencies for baricitinib based on relative risk versus placebo frequencies from TA534

Frequency of Event (n with event / N observed)	JAIN + JAIY JAIN-like			Placebo value from model (TA534)	Adjusted baricitinib value
	Placebo	Baricitinib	RR (baricitinib versus placebo)		
Injection site reaction	[redacted]	[redacted]	[redacted]	0%	[redacted]
Allergic conjunctivitis	[redacted]	[redacted]	[redacted]	5.8%	[redacted]

Infectious conjunctivitis	[redacted]	[redacted]	[redacted]	1.0%	[redacted]
Oral herpes	[redacted]	[redacted]	[redacted]	3.4%	[redacted]
Upper respiratory tract infections	[redacted]	[redacted]	[redacted]	0%	[redacted]

[a] Assumed equivalent to placebo for these events, [b] No RR can be calculated as placebo had 0 events, frequency calculated as [redacted].
Abbreviations: AE: adverse event; RR: relative risk.

The results of this scenario for the JAIN+JAIY JAIN-like population are presented in Table 28.

Table 28: Scenario analysis applying adjusted AE probabilities for baricitinib based on relative risk versus placebo frequencies from TA534

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,948	£17,948	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,513[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone vs. baricitinib; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: AE: adverse event; BSC: best supportive care; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

B22. PRIORITY: On the assumption that short term data from the JAIN-like JAIY population is the source of the AE rates, please justify why the JAHN study was not considered as a source of long-term adverse event rates. Please present a scenario analysis where the post-16-week adverse event rates are based on the JAHN study.

The long-term AE rates currently available from the JAHN extension study are from monotherapy-treated patients whereas the base case analysis in the economic model considers patients receiving combination therapy. Given this difference in the intervention received, the company does not consider the population in the JAHN extension study to be relevant for the target population in the economic model. Additionally, the JAHN extension study does not include a comparable placebo arm to act as a control. Furthermore, enrolment from originating studies into JAHN may have selected specific types of patient (i.e. there is a risk of unmeasured confounding), and this may introduce further bias.

However, an additional analysis was conducted summarising AEs in the baricitinib 4 mg arm from the originating JAHL/JAHM studies combined with that from the respective follow up period up to Week 36 in the JAHN study, and results of this analysis can be found in Table 29.

Table 29: Proportion of patients with selected TEAEs between baseline (JAHL/JAHM) and Week 52 (JAHN Week 36)

TEAE, n (%)	4 mg BARI (N=[redacted])
Oral herpes	[redacted]
Upper respiratory tract infection	[redacted]
Allergic conjunctivitis	[redacted]

Infectious conjunctivitis	[redacted]
Injection site reaction	[redacted]

Abbreviations: BARI: baricitinib; TEAE: treatment-emergent adverse event.

A scenario analysis was conducted where these data were used to inform AEs for baricitinib in the model. Fully incremental results for the base case population (JAIN + JAIY JAIN-like patients) are presented in Table 30. The impact of this update is negligible for baricitinib versus both BSC and dupilumab and did not affect the conclusions presented in the company submission.

Table 30: Scenario analysis where AEs for baricitinib are based on JAHN

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)	Quadrant
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,880	£17,880	NE
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,622[b]	NE

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone vs. baricitinib; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; NE: North East; QALYs: quality-adjusted life years.

Other

B23. PRIORITY: Please explain how the standard errors used in the PsA were generated, as it appears all values are 10% of the mean. Please amend the PsA so that the standard errors reflect the uncertainty in the data they are drawn from.

The standard error estimates for response inputs were derived from the ITC presented in Section B.2.9 of the company submission and are implemented in the model with conditional formulae responsive to the selected population source. All other estimates are assumed to be 10% of the mean based on a lack of additional input data describing the distribution other than the mean. Response rate data were derived from the BREEZE-AD trial, and thus standard errors were derived directly from the trial data and used in the model. Whilst utility data were derived from the BREEZE-AD trials, it was considered appropriate to use a conservative assumption of 10% of the mean in the PSA, since regression outputs might not accurately reflect the uncertainty in the trial data. For example, the LS mean change from baseline in EQ-5D (used to calculated the utility value for the maintenance state) represents an adjusted output from a regression, and thus the error terms calculated for this adjusted value may artificially limit the variability fed into the model.

However, an analysis has been conducted where the uncertainty for utility values was estimated based on the output of the regression models. For the baseline utility value, the SE was calculated from the observed data, based on the sample size and the standard deviation (reported in Table 91 of the CS). For the change from baseline in EQ-5D (used to calculate the utility value for the maintenance state), the SE was estimated based on the 95% CIs generated in the output of the MMRM analysis (i.e. by dividing the CI reported in Table 91 of the CS by 1.96, assuming a normal distribution). The revised probabilistic base case results are presented

in Table 31 and the revised cost-effectiveness plane scatterplot and cost-effectiveness acceptability curve are presented in Figure 5 and Figure 6, respectively. Baricitinib had a higher probability of being cost-effective at a willingness to pay (WTP) threshold of £20,000/QALY and £30,000/QALY compared with the analysis presented in the original submission. Please note that this analysis was conducted incorporating the correction to the SE for the dupilumab response rate, as highlighted in the response to Question B3.

Table 31: Revised probabilistic base case results

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER vs baseline (£/QALY)	Incremental ICER[a] (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,965	£17,965
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£89,879	£208,938[b]

[a] ICER of baricitinib versus comparator. [b] ICER represents cost saving per QALY foregone vs. baricitinib; ICER >£30,000 per QALY may be considered cost-effective.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life years.

Figure 5: Revised cost-effectiveness plane scatterplot

Generated using 3,000 iterations of the PSA.

Figure 6: Revised cost-effectiveness acceptability curve

Generated using 3,000 iterations of the PSA.

B24. Please provide functionality in the cost-effectiveness model to validate the following scenario analyses presented in Table 109 of the CS.

a) Scenario 2a: secondary censoring of clinical efficacy data (response definition: EASI50 and ΔDLQI≥4)

b) Scenario 2b: secondary censoring of clinical efficacy data (response definition: EASI75)

c) Scenario 3e: JAIN-like JAHL + JAIN-like JAHM versus CAFÉ-like SOLO (response definition: EASI50 and ΔDLQI≥4)

d) Scenario 3f: JAIN-like JAHL + JAIN-like JAHM versus CAFÉ-like SOLO (response definition: EASI75)

These scenarios were run manually due to the formatting of the cost-effectiveness model presented. For this reason, and in alignment with the approach agreed during the clarification call on 21st July, an Excel file is provided in the reference pack which contains all of the data necessary alongside guidance on how to run these scenarios.

Section C: Textual clarification and additional points

Systematic Literature Review (SLR)

C1. The search strategies used to identify studies for the clinical effectiveness SLR (Tables 1, 2, 3 Appendix D, p 9-14) contain the search terms “atopic eczema” and

“flexural eczema” but do not include the search term “eczema” alone. Please explain why this search term was excluded, commenting on the potential for relevant studies to be missed by its exclusion.

The decision problem presented by NICE in the final scope referred to atopic dermatitis only.[23] As the ERG have highlighted, the scope of the search performed was widened to include the terms “atopic eczema” and “flexural eczema”. This was judged to be appropriate given the use of the term “atopic eczema” in the background section of the final scope document. However, the suggested term “eczema”, used unqualified, can refer to a large number of other conditions which are not relevant to the NICE decision problem (or the anticipated baricitinib licensed indication).[24]

To ensure maximum rigour and that no relevant publications were excluded, conference proceedings and ongoing trials were searched and the dupilumab submission (TA534), which also only included atopic dermatitis in their eligibility criteria, was cross-checked.[2] No publications included in the clinical SLR for TA534 were omitted from our clinical search. Therefore, we have no reason to suspect that searching for eczema, beyond atopic or flexural eczema, would identify publications that are relevant to the assessment of baricitinib in patients with moderate-to-severe atopic dermatitis beyond those already identified within the search performed.

Textual clarifications

C2. In section B.3.2.1 of the CS it is stated that “Whilst not interchangeable, it can be assumed that ciclosporin is broadly comparable to azathioprine, methotrexate, and mycophenolate mofetil […]”. Please explain what the interventions are “broadly comparable” in (e.g. clinical effectiveness, adverse events, costs, mode of action or other).

We agree that these treatments may exhibit distinct benefit-risk profiles however, long-term comparative analyses have not been carried out. Additionally, these treatments are recommended for use at the same point in the treatment pathway and have similar cost and administration requirements. Upon consultation, expert dermatologists have advised that these treatments all have a poor benefit-risk profile. Together, these similarities make these treatments broadly comparable.

Additional Items from Lilly

C3. [redacted]

References

1. Brunner PM, Guttman-Yassky E, Leung DY. The immunology of atopic dermatitis and its reversibility with broad-spectrum and targeted therapies. J Allergy Clin Immunol 2017;139:S65-s76.
2. National Institute for Health and Care Excellence (NICE). Dupilumab for treating moderate to severe atopic dermatitis: Technology appraisal guidance [TA534]. Available at: https://www.nice.org.uk/guidance/ta534 [Last accessed: 28th January 2020].
3. Chopra R, Vakharia PP, Sacotte R, et al. Severity strata for Eczema Area and Severity Index (EASI), modified EASI, Scoring Atopic Dermatitis (SCORAD), objective SCORAD, Atopic Dermatitis Severity Index and body surface area in adolescents and adults with atopic dermatitis. Br J Dermatol 2017;177:1316-1321.
4. Leshem YA, Hajar T, Hanifin JM, et al. What the Eczema Area and Severity Index score tells us about the severity of atopic dermatitis - an interpretability study. The British Journal of Dermatology 2015;172:1353-7.
5. Boguniewicz M, Alexis AF, Beck LA, et al. Expert Perspectives on Management of Moderate-to-Severe Atopic Dermatitis: A Multidisciplinary Consensus Addressing Current and Emerging Therapies. J Allergy Clin Immunol Pract 2017;5:1519-1531.
6. Taylor K, Swan DJ, Affleck A, et al. Treatment of moderate-to-severe atopic eczema in adults within the U.K.: results of a national survey of dermatologists. British Journal of Dermatology 2017;176:1617-1623.
7. ClinicalTrials.gov. A Study of Baricitinib (LY3009104) in Participants With Moderate-to-Severe Atopic Dermatitis. Avaialble at: https://clinicaltrials.gov/ct2/show/results/NCT02576938 [Last accessed: 3rd August 2020].
8. I4V-MC-JAHG Clinical Study Report (20th October 2017).
9. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials 1986;7:177-188.
10. Bucher HC, Guyatt GH, Griffith LE, et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. Journal of Clinical Epidemiology 1997;50:683-691.
11. Borenstein M, Hedges LV, Higgins JP, et al. Fixed-effect versus random-effects models. Introduction to Meta-analysis 2009;77:85.
12. Institute for Clinical and Economic Review (ICER). Dupilumab and crisaborole for atopic dermatitis: effectiveness and value: Institute for Clinical and Economic Review, 2017.
13. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. Value Health 2011;14:539-45.
14. EMC. Olumiant 2 mg Film-Coated Tablets: SmPC. Available at: https://www.medicines.org.uk/emc/product/2434/smpc [last accessed: 20th April 2020].
15. SAS/STAT® 14.3 User's Guide. The MIXED procedure. Available at: https://v8doc.sas.com/sashtml/stat/chap30/sect16.htm [Last accessed: 3rd August 2020].
16. Ong PY, Leung DYM. Bacterial and Viral Infections in Atopic Dermatitis: a Comprehensive Review. Clinical Reviews in Allergy & Immunology 2016;51:329-337.
17. Egeberg A, Skov L, Andersen YMF, et al. Ten-year mortality is increased after hospitalization for atopic dermatitis compared with the general population, but reduced compared with psoriasis. Journal of the American Academy of Dermatology 2017;76:98-105.
18. Czech W, Bräutigam M, Weidinger G, et al. A body-weight–independent dosing regimen of cyclosporine microemulsion is effective in severe atopic dermatitis and improves the quality of life. Journal of the American Academy of Dermatology 2000;42(4):653-59.
19. Cork MJ, Danby SG, Ogg GS. Atopic dermatitis epidemiology and unmet need in the United Kingdom. Journal of Dermatological Treatment 2019:1-9.
20. Goujon C, Viguier M, Staumont-Sallé D, et al. Methotrexate versus cyclosporine in adults with moderate-to-severe atopic dermatitis: a phase III randomized noninferiority trial. The Journal of Allergy and Clinical Immunology: In Practice 2018;6(2):562-69. e3.

21. Schram ME, Roekevisch E, Leeflang MM, et al. A randomized trial of methotrexate versus azathioprine for severe atopic eczema. The Journal of allergy and clinical immunology 2011;128(2):353-9.
22. Gerbens LAA, Hamann SAS, Brouwer MWD, et al. Methotrexate and azathioprine for severe atopic dermatitis: a 5-year follow-up study of a randomized controlled trial. The British journal of dermatology 2018;178(6):1288-96.
23. National Institute for Health and Care Excellence (NICE). Final Scope: Baricitinib for treating moderate to severe atopic dermatitis [ID1622]. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ta10562/documents [Date accessed: 28th July 2020].
24. NHS. Atopic eczema. Available at: https://www.nhs.uk/conditions/atopic-eczema/ [Last accessed: 23rd January 2020].

NICE National Institute for Health and Care Excellence

Professional organisation submission

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**British Association of Dermatologists, University of Edinburgh, NHS Lothian**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	**British Association of Dermatologists.**
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] If so, please state the name of manufacturer, amount, and	[redacted] has participated in a Lilly advisory board on baricitinib.

purpose of funding.	
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	Baricitinib is designed to ameliorate symptoms and signs of atopic dermatitis and thus improve quality of life.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	A reduction in EASI score of 75% or a fall in IGA of 2 points.

8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. An additional oral agent for treating AD is needed.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	For patients with moderate to severe eczema, uncontrolled with topical agents, phototherapy or systemic treatments are generally required. Phototherapy is limited in supply and generally involves frequent time consuming visits to hospital. Existing systemic agents have a significant side effect profile, and require careful monitoring. Of the conventional systemic agents, only ciclosporin has a license for use in eczema and this only for 8 weeks, in inadequate length of time for a chronic condition such as eczema. Dupilumab has been a step change in treatment of eczema for patients not responding to, or being intolerant of, existing systemic agents. Unfortunately, not all patients respond to dupilumab, some develop problematic conjunctivitis, and others are fearful of injections.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	SIGN guidelines on Atopic Eczema in primary care (revised 2014).
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please	Well defined pathway of care in Scotland following SIGN guidelines, with local referral guidelines following these (e.g. Refhelp in Lothian)

state if your experience is from outside England.)	
• What impact would the technology have on the current pathway of care?	An alternative treatment to dupilumab for patients not responding to existing systemic agents and phototherapy.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes - an additional oral treatment for eczema, but probably following failure/intolerance of one or more of ciclosporin/methotrexate/azathioprine/MMF.
• How does healthcare resource use differ between the technology and current care?	No significant difference in resource use than existing systemic agents. Screening investigations will need to be performed before initiation and then occasional blood monitoring once treatment has started. This is similar to e.g. methotrexate/azathioprine use. Dupilumab does not require so much monitoring, but patients have to be taught to self-inject
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care. Outpatient treatment. Dermatology specialist service.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Education in safety profile and monitoring requirements for drug.

11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes - in that subset of patients who do not get benefit on existing systemic AD treatments.
• Do you expect the technology to increase length of life more than current care?	No.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No, but ongoing stratification studies on eczema patients may identify factors predicting best response to different drugs.
The use of the technology	
13. Will the technology be	Easier than dupilumab (no injections) and ciclosporin (less blood monitoring). Similar to MTX and

easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	azathioprine.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Presumably this will be more expensive than existing systemic treatments and safety profile less well understood. Thus guidelines will be required for starting criteria (e.g. failure of 1+ existing systemic agents) and a stop/go decision to be made by supervising clinician at a defined time after starting treatment, based on clinical response/adverse effects.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are	No.

unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	An additional treatment option for patients wishing/needing to treat their eczema with an oral agent.
• Is the technology a 'step-change' in the management of the condition?	No - not in the way dupilumab was, but a useful additional treatment option I hope.
• Does the use of the technology address any particular unmet need of the patient population?	An oral treatment for resistant eczema and for those intolerant of dupilumab due to ocular side effects.
17. How do any side effects or adverse effects of the	Side effect profile appears different from that of existing systemic Rx. As adverse profile of systemic drugs often determines which is used (e.g. hypertension/renal impairment a C.I for ciclosporin, liver dysfunction a

technology affect the management of the condition and the patient's quality of life?	CI for MTX) a drug with a different SE profile gives more options.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes.
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	EASI 75, DLQI, Pruritus score and IGA. All of these were measured.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Not relevant.
• Are there any adverse effects that were not	Not that I am aware of.

apparent in clinical trials but have come to light subsequently?	
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
20. How do data on real-world experience compare with the trial data?	I am not aware of any real-world data on baricitinib for atopic dermatitis.
Equality	
21a. Are there any potential equality issues that should be taken into account when considering this treatment?	Effects on different skin type (e.g. BAME skin).
21b. Consider whether these issues are different from issues with current care and why.	

Key messages
22. In up to 5 bullet points, please summarise the key messages of your submission. • First oral Jak inhibitor treatment for atopic dermatitis • First of a new class of drugs for treatment of atopic dermatitis • An alternative treatment option for patients intolerant of/not responding to conventional systemic agents • •

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Patient organisation submission

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	National Eczema Society
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	National Eczema Society is the UK charity for people of all ages living with eczema and those who care for them. We support people with information and advice about eczema and its management and treatment, which we deliver through our website, social media platforms, publications and nurse-supported Helpline. We are the campaigning voice for people with eczema and raise awareness of the needs of people with eczema with healthcare professionals, teachers and the government. We are funded by membership fees, donations from the public and organisations, and our corporate partners (pharmaceutical and emollient companies that sell products or services for people with eczema). We have approximately 2,600 members.
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.]	The manufacturer Eli Lilly has been a Corporate Member of National Eczema Society since May 2019, and the corporate membership agreement complies with the ABPI code of practice. The annual Corporate Membership fee paid by the company is £10,000 plus VAT. The Corporate Membership Scheme allows company partners to demonstrate public support for the important work of the Society. The funding helps pay for the charity's core operating costs with the purpose of helping the Society achieve its overall objective of supporting people living with eczema. In 2019, National Eczema Society also supported the company in conducting a small patient workshop exploring the experiences of people living with eczema. National Eczema Society currently has eight corporate members including Eli Lilly.

If so, please state the name of manufacturer, amount, and purpose of funding.	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No.
5. How did you gather information about the experiences of patients and carers to include in your submission?	National Eczema Society operates a nurse-supported Helpline service, responding to telephone and email enquiries from people affected by eczema who are seeking advice either on their own behalf or for a loved one. The calls and emails we receive give us a valuable insight into the experiences of people living with eczema and the many challenges they face. We also gain insights from the conversations and comments shared by people with eczema on our busy social media platforms.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Atopic eczema is a chronic dry skin condition. Its major symptom is itchiness, which can be intense and unbearable. Constant scratching causes the skin to split and bleed, and leaves it open to infection. Even when the eczema is mild to moderate (as opposed to severe), when it is not well-controlled it can have a significant impact on quality of life. In the UK, one in five children and one in twelve adults has eczema. Constant itchiness is one of the most challenging aspects of eczema; it can result in reduced social interaction and inability to work and study. In addition to the pain and discomfort brought about by scratching, itchiness often makes sleeping extremely difficult. Lack of sleep can compromise people's ability to concentrate at work and school/university and carry out everyday tasks effectively. It also damages personal relationships - as can itchiness alone. Eczema can have a significant negative impact

	on the whole family. People who are constantly itchy and/or have eczema on visible areas of their body can feel extremely self-conscious about their condition and appearance, and reluctant to leave their home. Eczema management is time-consuming. In addition to applying topical treatments at least twice a day, and every few hours when the skin is very dry, people who scratch a lot overnight may have to wash their bedding every day to remove blood and skin flakes. People who have a mental health condition (e.g. depression) as a result of their eczema, or in addition to it, often find it difficult to manage both conditions effectively. Even people who haven't been diagnosed with a mental health condition can find daily eczema management onerous and dispiriting. Caring for a child or adult with eczema can be time-consuming and exhausting, both physically and emotionally. Carers may need to apply topical treatments to the person in their care multiple times a day, try to distract them when they are itchy, provide emotional support and take them to regular GP or hospital appointments. Carers' ability to sleep is compromised when the person in their care is unable to sleep because of itchiness. Carers often need to get up several times during the night to apply emollient and comfort the person for whom they are caring. Lack of sleep for carers, as for people with eczema, can lead to their experiencing a diminished ability to concentrate at work and other activities, and carry out tasks effectively.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Many patients and carers consider the current treatments available for eczema on the NHS to be limited in number and effectiveness. Many patients are reluctant to use topical corticosteroids on a routine basis to control their symptoms because of concerns about adverse effects, notably skin thinning. Access to topical calcineurin inhibitors is limited, being prescribed for areas of delicate skin only. Current second-line treatments for eczema include phototherapy, oral steroids, immunosuppressant drugs (azathioprine, ciclosporin, methotrexate and mycophenelate mofetil) and a biologic drug (dupilumab). Second-line treatments can be effective for many people with eczema. However, a large proportion of people with eczema and their families have serious concerns about the potential for significant long-term harm through severe adverse side effects associated with immunosuppressant drugs. These concerns

	have been further highlighted with the Coronavirus pandemic. Dupilumab has fewer potential side effects than immunosuppressant drugs, but is only available to people who have tried and failed on at least one immunosuppressant drug and those who would not be eligible to take them. In addition, it is not effective for everyone who tries it, or suitable for people with certain co-morbidities.
8. Is there an unmet need for patients with this condition?	People with moderate to severe eczema are currently faced with the choice of managing as best they can with topical treatments, in great pain and discomfort, or starting phototherapy (which is not universally available) or dupilumab if they are eligible, or immunosuppressant drugs of uncertain efficacy with the potential for significant long-term harm through severe adverse side effects. Even if baricitinib is made available only in the same circumstances as dupilumab (i.e. for people who have tried and failed on at least one immunosuppressant drug, and those for whom immunosuppressant drugs are contraindicated), it will constitute a valuable additional treatment option for people with severe eczema, increasing the likelihood that patients will find a treatment that works effectively for them. Eczema is a heterogeneous disease requiring a variety of treatment options to meet patient need. Baricitinib also has the potential to reduce the need for topical steroid treatment, which people with severe eczema desperately want and deserve.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	The advantages of baricitinib are that it has been shown to improve the debilitating symptoms of eczema (itchiness, skin pain, sleep disturbance) and to do so rapidly. BREEZE trial data results are impressive in terms of symptom improvement and rapidity of symptom improvement. Baricitinib appears to work more quickly than azathioprine, methotrexate and mycophenelate mofetil, making it potentially more suitable than these drugs for acute eczema flares. ‘Efficacy And Safety of Baricitinib in Moderate to Severe Atopic Dermatitis: Results Of Two Phase 3 Monotherapy Randomized, Double-Blind, Placebo controlled 16-Week Trials (BREEZE-AD1 and BREEZE-AD2)’ (2019) by Eric L. Simpson et al shows baricitinib’s rapid onset of action. Significant improvement in itch was achieved as early as Week 1 for 4-mg and Week 2 for 2-mg. Improvements in

	night-time awakenings, skin pain, dermatology life quality index, and Patient-Oriented Eczema Measure were observed by Week 1 for both 4-mg and 2-mg. 'Efficacy and Safety of Baricitinib in Combination with Topical Corticosteroids in Moderate to Severe Atopic Dermatitis: Results of a Phase 3 Randomized, Double-Blind, Placebo-Controlled 16-week Trial (BREEZE-AD7)' (2019) by Kristian Reich et al also shows rapid clinically meaningful improvements in the patient-reported outcomes of itch, skin pain and sleep disturbance. Adverse events in these trials were mainly mild and moderate, and the safety profile was consistent with earlier findings. Many patients are likely to prefer an oral medication (such as baricitinib) over an injection. Baricitinib does not remain in the body for as long as some other current eczema treatments after you stop taking it, so people are able to regain their full ability to fight infection quickly if needed.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	One disadvantage of the technology is that it is unlikely to work effectively for everyone eligible to use it. Some patients may start treatment and not receive sufficient benefit to warrant continuing, which would be incredibly demoralising and result in a longer period of poorly controlled symptoms. We understand that the most common adverse events in patients treated with baricitinib were colds and headaches.

Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Patients with moderate to severe eczema for whom topical treatments are insufficiently effective and who must progress to second-line treatments would benefit from the introduction of a new second-line treatment option. Baracitinib has a different mode of action and safety profile and will benefit some people who currently have extremely poor symptom control. Patients with moderate to severe eczema who are concerned about the potential side effects of immunosuppressant drugs would benefit from the introduction of a new second-line treatment option, particularly a new type of treatment (i.e. a Janus kinase (JAK) inhibitor).
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	N/A

Other issues	
13. Are there any other issues that you would like the committee to consider?	N/A

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- The treatment options for eczema currently available on the NHS are limited and insufficient. The introduction of baricitinib has the potential to broaden patient choice, and would increase the likelihood that patients with moderate to severe eczema would find a treatment that is effective for them.
- Many people with eczema and their families have serious concerns about the potential for significant long-term harm through severe adverse side effects associated with immunosuppressant drugs. Adverse events in baricitinib trials were mainly mild and moderate.
- Trial data results show that baricitinib can not only improve, but rapidly improve, the symptoms of eczema that most people with the condition report as being the most debilitating (itchiness, skin pain, sleep disturbance).

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Clinical expert statement

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

Thank you for agreeing to give us your views on baricitinib and its possible use in the NHS.

You can provide a unique perspective on baricitinib in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Richard Weller**
2. Name of organisation	**NHS Lothian and University of Edinburgh**

3. Job title or position	**Honorary Consultant Dermatologist and University Reader**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with moderate-to-severe atopic dermatitis? ☐ a specialist in the clinical evidence base for moderate-to-severe atopic dermatitis or baricitinib? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for moderate-to-severe atopic dermatitis	
7. What is the main aim of treatment? (For example, to reduce/relieve symptoms (pruritus and dermatitis), prevent flares, or improve quality of life.)	To reduce severity of eczema, particularly symptoms of pruritus, and signs of inflamed skin. To improve quality of life.
8. What do you consider a clinically significant treatment response? (For example, achieving a certain EASI or IGA score, or a certain level of improvement from baseline.)	Reducing severity of eczema to mild (EASI <6, IGA 0 or 1)
9. In your view, is there an unmet need for patients and healthcare professionals in moderate-to-severe atopic dermatitis?	Not so much an unmet need, as a need for improvement- greater choice of agents used to treat this spectrum of eczema

What is the expected place of baricitinib in current practice?	
10. How is moderate-to-severe atopic dermatitis currently treated in the NHS?	Patients with mild and mild to moderate eczema are usually treated with topical agents. If these are ineffective phototherapy can be considered as a next step. Patients with moderate and severe eczema will usually have failed these topical and phototherapy treatments and rely on systemic treatment's. Current systemic treatments are in most cases either cyclosporin, Azathioprine or methotrexate. These are inexpensive and well-studied but have a significant side effect profile which often limits their use and not all patients can either tolerate or benefit from one or more of these agents.
• Are any clinical guidelines used in the treatment of moderate-to-severe atopic dermatitis, and if so, which?	No. guidelines are available in Englnad and Scotland for primary care management of eczema, but not moderate to severe which is generally treated in secondary care.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Pathway of care fairly well defined, but relatively poorly evidence based. The A*STAR study is collecting observational data on the moderate to severe eczema cohort to improve this evidence base, particularly in relation to existing systemic medication.
• What impact would baricitinib have on the current pathway of care?	It would be an additional choice for patients with mod to severe eczema

11. Will baricitinib be used (or is it already used) in the same way as current care in NHS clinical practice?	Probably- an alternative to existing systemic treatments. Although these are inexpensive they require fairly extensive monitoring and have a not insubstantial side effect profile.
• How does healthcare resource use differ between baricitinib and current care?	Probably less monitoring required. Time will tell, but side effect profile may be less- or at least different from existing drugs. Specific side effects tend to prevent the use of existing systemic agents in particular patients. For example obese patients might have abnormal liver function tests due to fatty livers and this would preclude the use of methotrexate. Patients with hypertension are less able to take cyclosporin. Azathioprine and methotrexate can cause nausea in a significant number of patients which prevents its use
• In what clinical setting should baricitinib be used? (For example, primary or secondary care, specialist clinics.)	Secondary care, and probably specialist eczema clinics where available.
• What investment is needed to introduce baricitinib? (For example, for facilities, equipment, or training.)	Training of prescribers.
12. Do you expect baricitinib to provide clinically meaningful	Yes.

benefits compared with current care?	
• Do you expect baricitinib to increase length of life more than current care?	no
• Do you expect baricitinib to increase health-related quality of life more than current care?	yes
13. Are there any groups of people for whom baricitinib would be more or less effective (or appropriate) than the general population?	Patients unable to take existing systemic treatments because of side effect profile- eg. Hypertensives, the obese.
The use of baricitinib	
14. Will baricitinib be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical	Easier than existing systemic treatments as less monitoring required.

implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with baricitinib? Do these include any additional testing?	Response -or its absence- at a given time point
16. Do you consider that the use of baricitinib will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No

17. Do you consider baricitinib to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes.
• Is baricitinib a 'step-change' in the management of moderate-to-severe atopic dermatitis?	Not a step change in the way that Dupilumab has been, but a significant improvement.
• Does the use of baricitinib address any particular unmet need of the patient population?	Patients unable to tolerate existing systemic treatments
18. How do any side effects or adverse effects of baricitinib affect the management of moderate-to-severe atopic	Don't know

dermatitis and the patient's quality of life?	
Sources of evidence	
19. Do the clinical trials on baricitinib reflect current UK clinical practice?	yes
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials	

but have come to light subsequently?	
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
21. Are you aware of any new evidence for the comparator treatment since the publication of NICE technology appraisal guidance [TA534]?	No
22. How do data on real-world experience compare with the trial data?	
Equality	
23a. Are there any potential equality issues that should be taken into account when	No

considering moderate-to-severe atopic dermatitis?	
23b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
24. The company have limited their submission to adults with moderate-to-severe atopic dermatitis who are candidates for systemic therapy and have failed at least one systemic immunosuppressant. Do you consider adults with moderate-to-severe atopic dermatitis who are candidates for systemic therapy and have not failed at least one systemic immunosuppressant to be a	No.

relevant patient population for this appraisal?	
25. Do you consider systemic immunosuppressants to be a relevant comparator in adults with moderate-to-severe atopic dermatitis who are candidates for systemic therapy and have failed at least one systemic immunosuppressant? (i.e. would patients be potentially offered more than one line of therapy with systemic immunosuppressants?)	yes
26. Do you consider alitretinoin to be a relevant comparator for this appraisal?	No- only licensed for hand eczema
Key messages	

25. In up to 5 bullet points, please summarise the key messages of your statement.

- A highly prevalent disease, with a limited number of treatments at present
- Existing treatments for mod-severe eczema need extensive monitoring
- Baricitinib a valuable addition to the existing limited treatment armementarium
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Patient expert statement

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

Thank you for agreeing to give us your views on baricitinib and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Alice Lambert**
2. Are you (please tick all that apply):	☐ a patient with moderate-to-severe atopic dermatitis? ☐ a carer of a patient with moderate-to-severe atopic dermatitis? ☒ a patient organisation employee or volunteer?

	☐ other (please specify):
3. Name of your nominating organisation	National Eczema Society
4. Did your nominating organisation submit a submission?	☒ yes, they did ☐ no, they didn't ☐ I don't know
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☒ yes
7. How did you gather the information included in your statement? (please tick all that apply)	☐ I have personal experience of moderate-to-severe atopic dermatitis ☐ I have personal experience of baricitinib ☐ I have other relevant personal experience. Please specify what other experience: ☐ I am drawing on others' experiences. Please specify how this information was gathered:
Living with moderate-to-severe atopic dermatitis	
8. What is it like to live with moderate-to-severe atopic dermatitis? What do carers experience when caring for someone with this condition?	

Current treatment of moderate-to-severe atopic dermatitis in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS? Are there any disadvantages of current treatments?	
10. Is there an unmet need for patients with moderate-to-severe atopic dermatitis?	
Advantages of baricitinib	
11. What do patients or carers think are the advantages of baricitinib compared to current treatments for moderate-to-severe atopic dermatitis? If there is more than one advantage, which is the most important and why?	

Disadvantages of baricitinib	
12. What do patients or carers think are the disadvantages of baricitinib?	
Patient population	
13. Are there any groups of patients who might benefit more or less from baricitinib than others? If so, please describe them and explain why.	
Equality	
14. Are there any potential equality issues that should be taken into account when considering moderate-to-severe atopic dermatitis and baricitinib?	

Other issues	
15. Are there any other issues that you would like the committee to consider?	

Key messages
16. In up to 5 bullet points, please summarise the key messages of your statement: • • • • •

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

.....

Your privacy

NICE National Institute for Health and Care Excellence

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

CONFIDENTIAL UNTIL PUBLISHED
Evidence Review Group's Report
Baricitinib for the Treatment of Moderate-to-Severe Atopic Dermatitis

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD
Authors	Robert Hodgson, Research Fellow, CRD
	Sahar Sharif, Research Fellow, CRD
	Lucy Beresford, Research Fellow, CRD
	Sumayya Anwer, Research Fellow, CRD
	Matthew Walton, Research Fellow, CRD
	Miriam Wittmann, Associate Professor in Inflammatory Skin Diseases, University of Leeds
	Melissa Harden, Information Specialist, CRD
	Claire Khouja, Research Fellow, CRD
	Sofia Dias, Professor of Health Technology Assessment, CRD
Correspondence to	Sofia Dias, Professor of Health Technology Assessment, CRD, University of York, York YO10 5DD
Date completed	28/08/2020 (corrected 24/11/2020)

Source of funding

This report was commissioned by the NIHR Systematic Reviews Programme as project number 131739.

Declared competing interests of the authors

None.

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:

Hodgson R, Sharif S, Beresford L, Anwer S, Walton M, Wittmann M, Harden M, Khouja C, Dias S. Baricitinib for the Treatment of Moderate-to-Severe Atopic Dermatitis: A Single Technology Appraisal. CRD and CHE Technology Assessment Group, University of York, 2020.

Contributions of authors

Sumayya Anwer, Sofia Dias and Sahar Sharif wrote the clinical effectiveness sections of the report. Lucy Beresford, Robert Hodgson and Matthew Walton wrote the cost effectiveness sections and conducted the ERG economic analyses. Miriam Wittmann provided expert clinical advice and commented on drafts of the report. Melissa Harden wrote the search strategy sections. Claire Khouja checked data and commented on drafts of the report. Sofia Dias took overall responsibility for the report.

Note on the text

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

Copyright statement

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

AD	Atopic dermatitis
AE	Adverse event
AESI	Adverse events of special interest
BAME	Black, Asian and Minority Ethnic
BSA	Body surface area
BSC	Best supportive care
CHMP	Committee for Human Medicinal Products
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
DLQI	Dermatology life quality index
EASI	Eczema Area and Severity Index
EMA	European Medicines Agency
eMIT	Electronic market information tool
ERG	Evidence review group
FE	Fixed effect
FLG	Filaggrin
HES	Hospital episode statistics
HRQL	Health related quality of Life
ICER	Incremental cost-effectiveness ratio
IGA	Investigators Global Assessment
IPD	Individual patient data
ITC	Indirect treatment comparison
ITT	Intention to treat
JAK	Janus Kinase
MCFB	Mean change from baseline
MMRM	Mixed model for repeated measures
NES	National Eczema Society
NICE	National Institute for Health and Care Excellence

NRS	Numeric rating scale
PAS	Patient access scheme
PASLU	Patient access scheme liaison unit
PFC	Points for clarification
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
QALY	Quality adjusted life-year
RCT	Randomised controlled trial
ROW	Rest of the world
SAE	Serious adverse event
SCORAD	SCORing Atopic Dermatitis
SLR	Systematic literature review
SmPC	Summary of product characteristics
SMR	Standardised mortality ratio
TCI	Topical calcineurin inhibitors
TCS	Topical corticosteroids
TE	treatment-emergent
TEAE	Treatment-emergent adverse events
WTP	Willingness-to-pay

1 EXECUTIVE SUMMARY

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

Population

The population considered in the submission was adults with moderate to severe atopic dermatitis (AD) who are candidates for systemic therapy and who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This is different to the NICE scope, which states the population is adults with moderate-to-severe AD who are candidates for systemic therapy that had an inadequate response or intolerance to existing topical treatments. Clinical advice is that the population in the company submission (CS) is restrictive as, in practice, baricitinib is likely to be used at the same point in the treatment pathway as other immunosuppressants, that is, prior to dupilumab. Therefore, the population addressed in the CS may not be the most relevant and fully representative population for this indication (Table 4).

The inclusion criteria for the clinical trials presented in the submission (JAIN, JAIY, JAHL and JAHM) specified an Eczema Area and Severity Index (EASI) score ≥ 16, Investigators Global Assessment (IGA) score ≥ 3 and body surface area (BSA) involvement ≥ 10%. Several published strata for the EASI score, state that moderate disease is associated with EASI scores as low as 6 ranging up to 22.9.[1] This includes patients with EASI scores far below the trial cut off of 16. This issue also applies to the clinical trials (CAFÉ and CHRONOS) supporting dupilumab, the main comparator considered in the appraisal, where the eligibility criteria included an EASI score of ≥ 20. Therefore, patients on the lower end of the moderate scale may be excluded, from the evidence presented for baricitinib and dupilumab, biasing the trial populations towards more severe disease.[1] The ERG therefore considers that the population presented in the clinical evidence may not represent all moderate to severe patients in the NHS population (Section 3.2.2).

Comparators

The company state that the use of baricitinib in the UK is expected to be as 5th line therapy following failure or contraindication of topical therapies, phototherapy and systemic immunosuppressant agents, with the comparator primarily being dupilumab. However, clinical advice to the ERG is that baricitinib would be given after topical treatment and phototherapy, when systemic immunosuppressants are considered. As a Janus kinase (JAK) 1/JAK2 inhibitor, baricitinib has a more targeted mode of action as compared with other systemic immunosuppressants, but a less targeted mode of action compared with dupilumab. (Table 4). After failure to respond to a treatment, patients tend to move on to other available treatments before best supportive care (BSC). Therefore, comparators to baricitinib should also include systemic immunosuppressants. Additionally, clinical advice is that currently dupilumab is favoured above a "second" systemic agent and thus started early

in the pathway due to its low side effect profile and less frequent monitoring requirements. This indicates that there may also be scope for treatment with baricitinib after dupilumab.

Outcomes

The draft summary of product characteristics (SmPC) for baricitinib suggests that initial assessment of response should be carried out at 12 weeks. However, this does not align with the trial outcomes presented in the CS, which are reported at 16 weeks. In response to the points for clarification (PFC), the company stated that as per usual UK clinical practice, it is expected that the majority of clinical assessments will be carried out at 16 weeks. This expectation was confirmed to be valid by a panel of expert dermatologist advisors, who felt that an early clinical assessment of efficacy would risk discontinuing treatment in patients who would go on to respond. However, the ERG does not agree that an early clinical assessment of efficacy would risk discontinuing patients' treatment. This is supported by the clinical data presented in the CS (Section 3.2.4.2), which reports peak response before week 12 in multiple outcomes.

Subgroups

Data were not available to conduct subgroup analyses for skin colour subgroups, although this was specified in the NICE scope and was a pre-planned subgroup in all four baricitinib trials. [redacted] were enrolled in any of the baricitinib trials, making subgroup analysis by skin colour difficult. The company stated that the trial program was not designed to investigate baricitinib efficacy in Black patients compared with other patient populations, but noted that there is some evidence that the pathology of AD could be more severe and persistent in Black patients. Without data on this cohort, the efficacy of baricitinib in this population is uncertain (Section 3.2.4.3).

Subgroups of people with moderate dermatitis and those with severe dermatitis were not presented, although this subgroup was specified in the NICE scope. In all four trials presented in the CS, baseline disease severity by IGA was a pre-planned subgroup. The CS has presented subgroup analyses for the JAIY, JAHL and JAHM studies by IGA (3 or 4). There are also published strata which allow classification by EASI score. In response to the PFC, the company stated that EASI does not reflect all aspects of moderate or severe disease and it does not provide consistent classification of disease severity. However, clinical advice to the ERG is that EASI is widely accepted and considers all the relevant aspects of the clinical signs of AD. The ERG considers that although there are limitations to using one severity classification, it would have been possible and beneficial to present separate subgroups of moderate and severe AD (Section 3.2.4.3).

Equality considerations

The ERG have identified that treatment efficacy may differ in people with different skin colours, particularly Black, Asian and Minority Ethnic (BAME) patients, which could be a potential equality issue. Although there is data on Japanese and East Asian patients, there is no data reported on Black patients. For this reason, subgroup analyses on skin colour were not conducted, which means that it is not possible to establish baricitinib efficacy in this population. Furthermore, the British Association of Dermatologists (BAD) state that effects on different skin type (e.g. BAME skin) should be considered as an equality issue for this indication (Table 4).

1.2 Summary of the key issues in the clinical effectiveness evidence

Generalisability of trial populations

All four baricitinib trials were reasonably good quality and the results are likely to be reliable. However, the mean age of patients in all four trials was [redacted] years old, which is higher than would be expected in the NHS population. The ERG also notes that published EASI strata for severe AD ranges from 21.1 to 50[1, 2] and the mean EASI scores in all four trials ranges from [redacted], which represents severe disease (see Section 3.2.2). Therefore, in terms of age and disease severity, the ERG considers that the population in the clinical evidence presented may not represent all moderate to severe patients in the NHS population.

Censoring of trial results

In the JAIN and JAIY trials, patients treated with 4 mg baricitinib were more likely to achieve EASI50, EASI75 and Itch NRS ≥ 4-point improvement compared with placebo at 16 weeks. There was also a significant ≥ 4-point improvement at week 24 for Itch NRS but not for EASI50, EASI75 or EASI90 in the JAIN trial. Patients in the 4 mg baricitinib arm were more likely to achieve a ≥ 4-point improvement in dermatology life quality index (DLQI) compared to those in placebo at week 16 in JAIN and JAIY. These results were consistent when using primary and secondary censoring (see Section 3.2.4). Clinical advice to the ERG is that the secondary censoring rule, where patients are not censored if they use topical corticosteroids (TCS) as rescue therapy, is more likely to reflect clinical practice as it is expected that rescue medication will be used concomitantly with baricitinib. For patients on baricitinib and dupilumab, flare is not considered an indicator of loss of response or grounds to discontinue treatment. A flare is considered a short event which can be managed with rescue therapy, with the patient either continuing on with treatment or resuming treatment after stopping the rescue therapy. Thus, data should not be censored after the initiation of rescue therapy (see Section 3.2.4.1).

Significant subgroup analyses by region

The greater response seen in European patients compared with non-European and Japanese patients in the JAIY trial and JAIN + JAIN-like JAIY population may be due to differences in clinical practice, particularly the use of rescue TCS and baseline severity of AD. In Europe, clinical practice broadly limits the use of high potency TCS, whereas Japan favours it. Additionally, differences in baseline EASI score and BSA were noted, indicating that Japanese patients have more severe disease (see Section 3.2.4.3). Notably, the results of regional subgroup analyses do not support a lower response for baricitinib 4 mg in East Asian countries not including Japan, suggesting there is not a specific effect of East Asian ethnicity but rather the difference in response may be due to other characteristics of the recruited Japanese patients. However, this is a source of uncertainty, indicating that the trial populations are not fully generalisable to the NHS population, which should be considered when interpreting the results.

Trial differences included in the indirect treatment comparisons (ITC)

Inevitably the trials included in the ITC vary by design and patient characteristics. There was a substantial difference in the proportion of Asian patients between the JAIN ([redacted]%) and CAFÉ (2%) studies. As noted, geographic region may be an effect modifier and therefore this is potentially a source of inconsistency between these trials (see Section 3.3). There was also a difference in the baseline severity of the patients included in the trials. The eligibility criteria for the baricitinib trials was an EASI score of ≥ 16, whereas for the CAFÉ and CHRONOS trials it was as EASI score of ≥ 20. This indicates that patients in the CAFÉ and CHRONOS trials are likely to be more severe than those included in the baricitinib trials, which is reflected in the baseline EASI scores of the trials (Section 3.3).

Limitations of the ITC

The ITC results using primary censoring report that dupilumab is more effective than baricitinib in achieving EASI50 + ≥ 4-point improvement in DLQI (odds ratio [OR]: [redacted] 95% confidence interval [CI]: [redacted]) and [redacted] (OR: [redacted]) at week 16 using primary censoring. These results were similar using secondary censoring. For the full JAIN vs CAFÉ populations, there was a [redacted] of achieving EASI50 with dupilumab than baricitinib (OR: [redacted]). However, there is considerable uncertainty in most ITC results due to wide confidence intervals (Section 3.4.2). Patient reported outcomes, such as skin pain NRS and ADSS were not included in the ITC due to the outcomes not being available from the CAFÉ and CHRONOS trials. Itch NRS ≥ 4-point improvement was only available for the full JAIN vs CAFÉ population, which reported no significant differences between groups. Clinical advice to the ERG is that these patient-reported outcomes are very important due to the effect they have on the quality of life of patients, particularly itch, as it is correlated with flares and lack of sleep in

patients with AD. Additionally, the ITC was only available for outcomes at 16 weeks, therefore, the long-term efficacy of baricitinib compared with dupilumab is uncertain. There was also no ITC for adverse events carried out, even though this could have been conducted for adverse events that were reported by the CAFÉ trial, including for patients with ≥ 1 treatment-emergent adverse events (TEAE), ≥ 1 treatment-emergent (TE) serious adverse events (SAE), death, diarrhoea, abdominal pain, back pain and asthma.

Adverse events of baricitinib

In JAIN, at 16 weeks, a higher proportion of patients in the 4 mg baricitinib group (75.0%) experienced at least one TEAE compared to the placebo group (53.8%). Between 16 weeks and 24 weeks, an additional [redacted] TEAEs were experienced in the 4 mg baricitinib group. In the integrated analysis, [redacted] experienced at least 1 TEAE in both groups. In both JAIN and the integrated analysis, the most common AESIs were treatment emergent (TE) infections ([redacted]% and [redacted]%, respectively), and in particular herpes simplex ([redacted]% and [redacted]%, respectively). In the CAFÉ and CHRONOS trials the most common adverse events with dupilumab treatment were infections and infestations (45.8% and 57%, respectively), particularly nasopharyngitis (20.6% and 23%, respectively). However, eye disorders (19.6% and 31%, respectively) were also observed as adverse effects with dupilumab. [redacted] were observed [redacted] of the trials for the 16-week duration, or up to 24 weeks for JAIN (see Section 3.2.5).

Flare suppression

An integral part of managing AD is the control of flares, as AD is episodic in nature. Flares are typically treated using high potency TCS. Reducing flares and TCS use is a priority to patients and clinicians due to the adverse effects associated with using TCS. Although flare is not an outcome presented in the submission, receipt of rescue can be considered a proxy for a flare. In the JAIN trial, a similar number of people were rescued in both the placebo and 4 mg baricitinib arms (n = [redacted] and [redacted], respectively) at week 16 and more patients were rescued in the baricitinib 4 mg arm (n = [redacted]) compared with the placebo arm (n = [redacted]) at week 24 (see Section 3.2.5.6). However, in the CHRONOS trial, dupilumab reported greater flare suppression when compared with placebo (16% vs 52% respectively), significantly reducing the need for rescue therapy. This indicates that baricitinib treatment may not be effective at reducing flares (Section 3.4.2).

1.3 Summary of the key issues in the cost effectiveness evidence

The ERG identified structural uncertainties associated with the company's approach that limit the value of the analysis. A number of potentially substantive issues in data selection and analysis methodology were inadequately explored by the company. Unfortunately, in many cases the company

chose not to provide data to the ERG in order to resolve these issues, meaning that a great deal of uncertainty remains regarding the predicted cost-effectiveness of baricitinib.

Model structure

Structure does not reflect the disease course

The model does not account for the waxing and waning nature of AD, nor how treatment is currently used to address patterns of disease. An important consequence of this omission is that patients who don't respond, or lose their response to treatment, are assumed to remain in a state of chronic and severe AD until death. This is inconsistent with clinical reality and misrepresents the effectiveness of BSC (see Section 4.2.1).

Use of one response health state

The company took a comparatively simplistic approach to modelling patients' response to treatment, using a single health state to represent all patients responding to treatment rather than splitting this into categories indicating magnitude of response. While this approach was adopted in TA534, it is less precise and potentially biases the model in favour of less efficacious treatment options. Further, in TA534 treatment specific utilities were adopted potentially justifying the use of a single response health state (see Section 4.2.1).

Response assessment period

Response to baricitinib was assumed to be assessed at Week 16; however, the ERG notes this is contrary to recommendations made in the draft SmPC for baricitinib, which outlines that response should initially be assessed at week 12. This reflects the fact that peak response rates are achieved well before 12 weeks of treatment. Given a 16 week assessment period is already in use for dupilumab it is uncertain whether response assessment in NHS clinical practice for baricitinib would be at 12 or 16 weeks (see Section 4.2.1).

Meaningfulness of company's composite response definition

It is not clear whether the *post hoc* composite of EASI50 and ΔDLQI $\geq$ 4 outcomes to define response would be recognised or treated as clinically meaningful in practice. Whilst the committee in TA534 concluded that the post hoc composite of EASI50 and ΔDLQI $\geq$ 4 outcomes to define response was appropriate for decision-making, the ERG notes that the primary trial outcome was EASI75, which was also considered a clinically significant improvement by the British Association of Dermatologists in their submission. There ERG further notes that there is no correlation between response and HRQL in the company's regression analysis of JAIN and JAIN-like JAIY patients, which may suggest the response criteria do not reflect the benefits of treatment (see Section 4.2.5.1).

Modelled population

Generalisability of modelled population: disease severity

The baricitinib trials (and thus the modelled population) limited inclusion to patients with more severe disease than would be expected in the moderate-to-severe AD population seen in NHS practice. Inclusion of patients with less severe disease may impact on cost-effectiveness due to the potential for differential effectiveness in these patients as well differences in the costs and benefits associated with these patients (see Section 4.2.2).

Generalisability of modelled population: Ethnicity

Ethnicity and skin colour may represent important treatment effect modifiers in AD, as has been observed across other inflammatory disorders, and in limited subgroup analyses presented by the company. The ERG found it particularly concerning that there were [redacted] black patients included in the evidence base comprising the company's ITC, given the greater prevalence and severity of AD in the Black British population. This issue pertains to both baricitinib and comparator trials. Given differences in disease pathology and treatment efficacy across ethnic groups, it is questionable whether it is appropriate to assume the efficacy results observed in white patients are transferrable to other ethnicities unrepresented in the trial evidence (see Section 4.2.2).

Intervention and comparators

Company did not consider the use of treatment sequences

The ERG was concerned that the company's strategy for the positioning of baricitinib in the treatment pathway may limit other treatment options available to patients. The company wishes NICE to consider baricitinib for use only in patients naïve to dupilumab, and the model does not consider the potential for sequential treatment using these therapies. This was contrary to advice received by the ERG, which suggested that clinicians would be very keen to have both treatment options available to patients, as is the case with newer therapies in psoriasis. Furthermore, a substantial proportion of the eligible population will be dupilumab-experienced, which would preclude patient access to baricitinib if sequences are not permitted (see Section 4.2.3).

Omission of comparators listed in the NICE Scope

The company positions baricitinib as a comparator primarily to dupilumab and therefore focuses on a population who have failed one or more immunosuppressants. As such, immunosuppressants are excluded as a comparator in economic analysis. The ERG, however, considers that there is scope for further immunosuppressant use in many of these patients given the availability of several different immunosuppressants and the potential for patients to be re-inducted. Further, the mode of action of baricitinib, potentially places it as a more natural comparator to the immunosuppressants than dupilumab. This is because as a JAK1/JAK2 inhibitor, baricitinib is more broadly immunosuppressive

than dupilumab which has a more focused mode of action distinct from both baricitinib and immunosuppressants (see Section 4.2.3).

Treatment effectiveness

Improper calculation of response rates from the ITC

The response rates applied in the model were derived using absolute measures of the treatment effect rather than relative effects. The use of absolute treatment effects can result in bias where the response rates in the common comparator (placebo + TCS) differ across studies. It also departs from the analysis suggested in NICE DSU TSD5,[3] which recommends that absolute response rates be pooled on the log-odds scale. Following the approach recommended by the DSU would also allow the company to use standard errors for the response probabilities in their probabilistic analysis (see Section 4.2.5.1).

Validity of long-term efficacy assumption

The majority of health benefit generated by baricitinib in the company's model is based on the assumption of equivalence with dupilumab in terms of long-term efficacy. The ERG has a number of substantive issues with this assumption, and considers such benefits unlikely to be realised in practice. Specifically, the ERG notes that these technologies have vastly different mechanisms of action and modes of administration. The available clinical evidence from JAIN and JAHN also does not support this assumption and suggests substantial differences in adherence, and fewer patients retaining response to treatment (see Section 4.2.5.2).

Inconsistencies and bias in discontinuation rates applied beyond Week 52

The source of discontinuation rates for BSC appears to be biased strongly in favour of baricitinib and dupilumab, as patients requiring rescue therapy on BSC permanently lose response, while those on baricitinib and dupilumab do not. The use of rescue therapy is not a good indication of patients losing response. Symptom control may be overcome by a trigger factor in AD, resulting in a flare. However, clinicians would expect that, in the case of such a flare, control can be re-established on the same medication following rescue, thus flare should not be conflated with loss of response and permanent loss of any HRQL gain. The rapid rate of discontinuation modelled beyond Week 52 for BSC represents permanent loss of disease control, which does not represent the ERG's view of the course and management of AD (see Section 4.2.5.3).

Unsupported claim of flare suppression

The ERG does not agree that the company used the most plausible available estimates of flare frequency. In the CHRONOS trial, dupilumab demonstrated flare suppression over long-term continuous use versus placebo (16% vs 52% respectively), and thus a significantly reduced need for rescue therapy. The company assumed that baricitinib is equally effective as dupilumab with regards

to flare control. However, in the JAIN trial, more patients in the baricitinib arm required rescue therapy for flares than did those on BSC, implying no substantial flare control is associated with baricitinib treatment (see Section 4.2.5.4).

Health related quality of life

Impact of response status on HRQL

The regression analysis performed by the company on a large sample of HRQL data found no significant difference in the utility of patients classified as responders and non-responders. In the context of the model, this implies that there are no health gains from treatment as valued by EQ-5D. The company disregarded this result in the economic model, and instead applied the baseline utility to non-responders health state. This baseline utility was substantially lower than the average utility measured at the same time point in responders. The ERG does not consider this an appropriate method for estimating the relationship between response and non-response on HRQL (see Section 4.2.6).

Source of utility data

The company base-case uses the utility values from the JAIN trial as well as JAIN-like patients in the JAIY trial. While this population is consistent with the effectiveness data used in the model, it ignores available data on JAIN-like patients recruited to other pivotal trials (JAHL and JAHM). It is the ERG's view that utility should be drawn from the largest possible sample and should include all JAIN-like patients, particularly given the small number of responders providing data (see Section 4.2.6). However, it should be noted that, unlike the patient population considered in the economic model, patients recruited to the JAHL and JAHM trials did not receive concomitant TCS.

Resource use and costs

Composition of BSC

Within the economic model BSC is modelled as a blended comparator consisting of topical mometasone (TCS), topical tacrolimus (TCI) and oral Prednisolone (a corticosteroid). Several elements of BSC (mometasone and tacrolimus) are however, also included as part of concomitant treatment received by all patients. This leads to the model double counting the costs of BSC and is inconsistent with the approach adopted in TA534. The ERG also notes several inconsistencies in the dosing of elements of BSC depending upon where they are applied in the model (see Section 4.2.7.1).

Other resource issues

The ERG identified several minor issues relating to the composition of concomitant treatments and health state costs. These specifically related to the inclusion of bathing products which are no longer used in practice and the omission of blood monitoring tests for baricitinib which are likely to be required due to increases in blood creatine kinase and lipids. Elevation in lipids may also require some patients to take statins (see Section 4.2.5.5 and 4.2.7).

1.4 Summary of exploratory and sensitivity analyses undertaken by the ERG

The scenario analyses conducted by the ERG are summarised in Table 1. The ERG also modelled the cost-effectiveness of baricitinib when used as part of the treatment sequence.

Table 1. Summary of ERG Scenario Analyses

Scenario 1	Discontinuation from first-line BSC removed.
Scenario 2a	Primary censoring, using rate of response in the JAIN-like patients in placebo arm, applying relative effects used in ITC.
Scenario 2b	Secondary censoring, using rate of response in the JAIN-like patients in placebo arm, applying relative effects used in ITC.
Scenario 2c	Primary censoring, using rate of response in the JAIN-like and CAFÉ-like patients in placebo arm, applying relative effects used in ITC.
Scenario 2d	Secondary censoring, using rate of response in the JAIN-like and CAFÉ-like patients in placebo arm, applying relative effects used in ITC.
Scenario 3a	Conditional probability of retaining response based on JAHN
Scenario 3b	Conditional probability of retaining response and post 52 week discontinuation based on JAHN
Scenario 3c	Week 16 to 52 and post week 52 to rates set to all-cause discontinuation rates from JAHN and CHRONOS.
Scenario 4	Changing the flare rate in baricitinib to equal that of BSC.
Scenario 5a	Apply comparative utilities estimated from JAHN and JAIY JAIN-like patients.
Scenario 5b	Applying the dupilumab utilities to all treatment arms.
Scenario 6a	Removing the drug acquisition costs from BSC.
Scenario 6b	Altering the drug acquisition costs for BSC.
Scenario 7	Removing bathing products from concomitant therapy.
Scenario 8	Adding blood tests to the monitoring.
Scenario 9	Revising number of dupilumab injections in drug acquisition costs.

The results of the ERG scenario analyses are presented in Table 2. These results are presented inclusive of the PAS available for baricitinib but exclude the PAS discount for dupilumab. Results including the PAS discount are presented in a confidential Appendix.

Table 2. Fully incremental deterministic Results of ERG scenario analyses

Analysis	Intervention	Discounted Costs	Discounted QALYS	Fully incremental ICER	Change from Base Case
ERG Correction of Model Errors	BSC	[redacted]	[redacted]	-	
	Baricitinib	[redacted]	[redacted]	£18,003	+£7
	Dupilumab	[redacted]	[redacted]	£204,046	+£78
Treatment Sequencing	BSC	[redacted]	[redacted]		-
	Baricitinib	[redacted]	[redacted]	£18,003	-
	Baricitinib + Dupilumab	[redacted]	[redacted]	£90,446	-
	Dupilumab	[redacted]	[redacted]	Dominated	-
	Dupilumab + Baricitinib*	[redacted]	[redacted]	£3,597,452	-
Scenario 1: No discontinuation on first-line BSC	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£29,595	+£11,599
	Dupilumab	[redacted]	[redacted]	£291,428	+£87,460
Scenario 2a) ITC Relative Effect, Primary Censoring, JAIN-like population	BSC	[redacted]	[redacted]	-	
	Baricitinib	[redacted]	[redacted]	£18,009	£13
	Dupilumab	[redacted]	[redacted]	£205,062	£1094
Scenario 2b) ITC Relative Effect, Primary Censoring, JAIN/ CAFÉ population	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£17,959	-£37
	Dupilumab	[redacted]	[redacted]	£253,917	£49,949
Scenario 2c) ITC Relative Effect, Secondary Censoring, JAIN-like population	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£18,046	£50
	Dupilumab	[redacted]	[redacted]	£182,592	-£21,376
Scenario 2d) ITC Relative Effect, Secondary Censoring, JAIN/ CAFÉ population	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£17,954	-£42
	Dupilumab	[redacted]	[redacted]	£237,490	£33,522
Scenario 3a) Conditional Response JAHN	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£18,413	£447
	Dupilumab	[redacted]	[redacted]	£144,144	-£59,824
Scenario 3b) Conditional Response and Discontinuation Rates JAHN	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£21,465	£3,499
	Dupilumab	[redacted]	[redacted]	£98,746	-£105,222
Scenario 3c) JAHN discontinuation rates	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£20,543	£2,577
	Dupilumab	[redacted]	[redacted]	£100,909	-£103,059
Scenario 4: Flare rates for baricitinib based on BSC	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£18,070	£74

	Dupilumab	[redacted]	[redacted]	£203,938	-£30
Scenario 5a: Utilities based on Company's regression analysis	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	Dominated	-
	Dupilumab	[redacted]	[redacted]	Dominated	-
Scenario 5b Utilities based on dupilumab (TA534) submission.	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£33,451	£15,455
	Dupilumab	[redacted]	[redacted]	£352,831	£148,863
Scenario 6a: BSC drug costs amended	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£22,808	£4,812
	Dupilumab	[redacted]	[redacted]	£208,619	£4,651
Scenario 6b: BSC drug costs not included	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£20,223	£2,257
	Dupilumab	[redacted]	[redacted]	£206,159	£2,191
Scenario 7: Remove the costs of bathing products	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£17,330	-£636
	Dupilumab	[redacted]	[redacted]	£203,407	-£561
Scenario 8: Monitoring costs for baricitinib	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£18,078	£112
	Dupilumab	[redacted]	[redacted]	£203,925	-£43
Scenario 9: Correction of number of dupilumab injections	BSC	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	£18,003	£37
	Dupilumab	[redacted]	[redacted]	£203,056	-£912

*This ICER is estimated relative to the sequence baricitinib + dupilumab as the sequence including dupilumab alone is strongly dominated.

1.5 Summary of ERG's preferred assumptions and resulting ICER

The ERG's preferred base case applies Scenarios 1, 2d, 3c), 4, 5b, 6b, 7, 8 and 9. The ERG's preferred base case also applies treatment sequencing. Results are presented in Table 3. These results are presented inclusive of the PAS available for baricitinib but exclude the PAS discount for dupilumab. Results including the PAS discount are presented in a confidential Appendix.

Table 3 Fully incremental Deterministic ICER resulting from ERG's preferred assumptions

Analysis	Intervention	Discounted Costs	Discounted QALYS	ICER	Net Monetary Benefit	
					£20,000 WTP threshold	£30,000 WTP threshold
ERG Base Case: Sequencing & Scenarios 1, 2d, 3, 4, 5b, 6b, 7, 8 and 9	BSC	[redacted]	[redacted]			
	Baricitinib	[redacted]	[redacted]	£64,710	[redacted]	[redacted]
	Baricitinib + Dupilumab	[redacted]	[redacted]	£174,071	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	Dominated	[redacted]	[redacted]
	Dupilumab + Baricitinib	[redacted]	[redacted]	£334,999	[redacted]	[redacted]

EVIDENCE REVIEW GROUP REPORT

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

The company submission (CS) includes an appropriate and relevant summary of the underlying health problem.

Atopic dermatitis (AD) is a chronic inflammatory skin disease, with significant impact on life quality due to intense pruritus (itch) impacting on sleep and daily life. The typical course of AD is episodic with repeated flare ups. The course can be continuous for long periods, however the defining feature of AD is flares.[4] In adolescence and adulthood, flexural areas such as antecubital area are typically affected but the disease can extend all over the body often also affecting the face, hands and feet, which impairs function.[5] Acute AD lesions are red, oozing and painful, chronic AD shows a thickened, dark red skin (lichenification). An underlying feature in most AD patients is very dry skin and as mentioned an intense and uncomfortable itch, leading to sleep loss and substantial impairments in quality of life.[6]

AD has in the past been considered as mainly a paediatric disease; however, AD also affects a significant number of adults. The symptoms of AD may begin at any age and it can be a life-long condition, however little is known about the variability of disease activity over the long-term.[7] The CS states that the prevalence of AD in adults in the UK has been reported as 2.5%, which equates to roughly 1.2 million people.[8] The ERG notes that of these patients, more than half (53% to 68%) have moderate to severe disease.[9, 10] These estimates can differ depending on the scale of severity measurement used. There are several scales used to measure the severity of AD: Eczema Area and Severity Index (EASI), Investigators Global Assessment (IGA), SCORing Atopic Dermatitis (SCORAD) and body surface area (BSA), which often lead to inconsistent classification of the severity of AD.[11] Recently, also supported by the HOME initiative (harmonising outcome measures in Eczema) EASI has become the most commonly used outcome measure[12].

2.2 Background

Overall, the CS provides a generally appropriate summary of the current service provision for patients with moderate to severe atopic dermatitis.

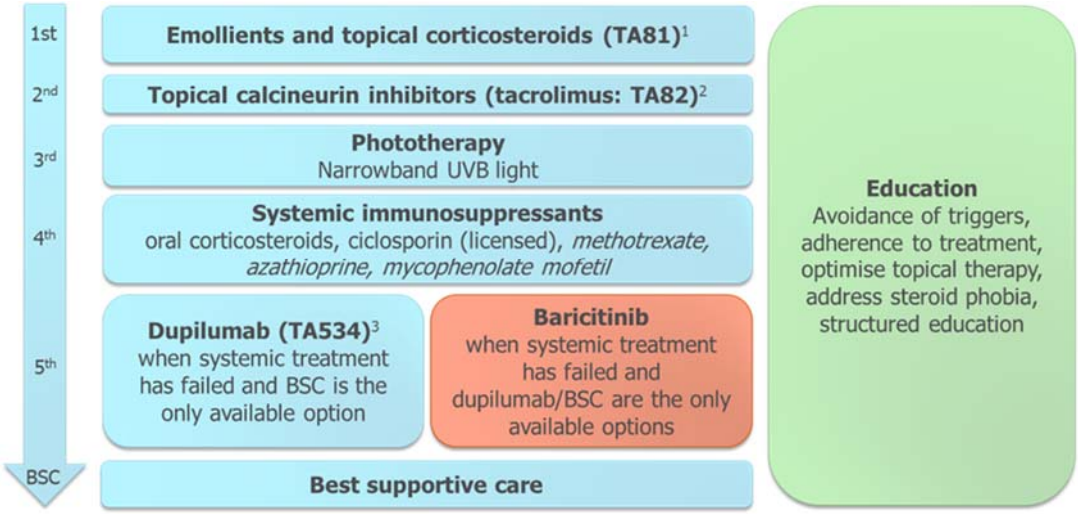
The only available NICE guidelines for the treatment and management of AD in the UK is for patients under the age of 12 years, which the CS states contributes to clinical practice being highly individualised.[13] The CS states that emollients are recommended as first line treatment alongside anti-

inflammatory topical corticosteroids (TCS) to treat active disease or prevent a relapse of symptoms. However, clinical advice is that emollients and interval use of topical corticosteroids/calcineurine-antagonists for the prevention of trans epidermal water loss and thus ease of pruritus caused by dry skin is recommended at all treatment stages and is indeed common practise.[14] If symptoms persist following emollient treatment, topical calcineurin inhibitors (TCI) are recommended as second line treatment. Mild-to-moderate disease is often managed with emollients, TCS and TCI.[15] However, the CS states that short-term use of TCS is best, due to the increased risk of adverse events including skin atrophy, skin bleaching and skin infections.[16] Phototherapy is recommended, as third line treatment, where non-pharmacological and topical measures have failed. Although due to the need for frequent applications in specialised centres, it is not feasible for everyone.[17]

Fourth line treatment constitutes systemic immunosuppressants, which include oral corticosteroids, ciclosporin, methotrexate, azathioprine and mycophenolate mofetil.[18] The only systemic immunosuppressant therapy currently licensed for AD in the UK is ciclosporin. However, clinical advice is that other systemic therapies, particularly methotrexate, are often used off-label instead of ciclosporin, due to its poor safety profile. Ciclosporin is not used for longer than a year due to the increased risk of renal insufficiency, tremor, hypertension and malignancy, particularly of the skin.[19] As fifth line treatment, dupilumab has been recommended by NICE for adults with severe-to-moderate AD who experience failure with, are intolerant to or have contraindication to at least one systemic therapy.[20] However, the ERG's clinical advisor states the current clinical reality is that dupilumab is favoured above a "second" systemic and thus started early in the pathway due to its low side effect profile and less frequent monitoring requirements. On the other hand, dupilumab is given as an injection, which can be difficult for some patients. Best supportive care (BSC) which generally includes low-to-mid potency topical corticosteroids, phototherapy, psychological support, rescue therapy, higher potency topical or oral corticosteroids or topical calcineurin inhibitors and extensive use of emollients is used as last stage treatment.[14]

As shown in Figure 1, the CS positions baricitinib as an alternative to dupilumab in adult patients with moderate-to-severe AD who are candidates for systemic therapy and who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This positioning is more restrictive than the licensed population, which is for adult patients who are candidates for systemic therapy. This is discussed in more detail in Table 4.

Figure 1 Treatment pathway for patients with moderate to severe atopic dermatitis (from CS, Figure 3)



The CS states that baricitinib is an oral Janus Kinase (JAK) inhibitor, which acts selectively and reversibly to inhibit the JAK family of protein tyrosine kinases, specifically JAK1 and JAK2. These enzymes mediate pathways involved in the inflammatory processes underlying AD. By inhibiting this signalling, baricitinib modulates the intracellular signalling of multiple cytokines involved in AD.[21] Clinical advice to the ERG and the British Association of Dermatologists (BAD) suggest that baricitinib treatment will require frequent blood monitoring, which is similar to administering systemic immunosuppressants.

2.3 Critique of company's definition of decision problem

Table 4 compares and critiques the company's decision problem with the final NICE scope

2.3.1 Population

No definition of moderate to severe AD is specified in the NICE scope and there is no gold standard for defining moderate to severe AD.[11] The CS specifies that AD may be considered moderate to severe when one or more of the following criteria are met:

- A minimum involvement of 10% body surface area
- Presence of individual lesions with moderate to severe features
- Involvement of highly visible areas or those important for function
- Significantly impaired quality of life

However, clinical advice to the ERG is that at least one of the first two criteria, or both last criteria are generally required in practice to classify a moderate to severe patient. Most clinical measures to assess

severity have never been validated and are largely heterogeneous in nature.[1] The IGA is a five-point scale that provides a global clinical assessment of AD severity ranging from 0 to 4, where 3 is moderate, and 4 indicates severe. Whereas EASI assesses both the extent and severity of AD, where the indicated categories from two published sources are 6 to 22.9 for moderate AD and 21.1 to 50 for severe AD.[1, 2, 18] Table 4 notes that the inclusion criteria of the baricitinib trials regarding AD severity may indicate that the trial populations are not representative of the NHS population.

2.3.2 Intervention

The intervention presented in the CS is baricitinib (Olumiant®), which matches the NICE scope (Table 4). The recommended posology is 4 mg once daily. Topical corticosteroids can be given alongside baricitinib. The CS states that an optional down-titration dose of 2 mg is appropriate for some patients such as those aged 75 years or older, or patients with a history of recurrent infections. Clinical advice to the ERG is that over 75s are a small proportion of patients with AD. Due to co-morbidities, phototherapy and topical treatments are often the preferred treatment options for over 75s and it may seem more appropriate to suggest dose-reduction depending on co-morbidities rather than age.

The marketing authorisation for baricitinib is expected between [redacted] and positive opinion from the Committee for Human Medicinal Products (CHMP) is expected on [redacted].

2.3.3 Outcomes

The draft SmPC for baricitinib suggests that initial assessment of response should be carried out at 12 weeks. However, this does not align with the trial outcomes presented in the CS, which are reported at 16 weeks. In response to the PFC, the company stated that as per usual UK clinical practice, it is expected that the majority of clinical assessments will be carried out at 16 weeks. This expectation was confirmed to be valid by a panel of expert dermatologist advisors, who felt that an early clinical assessment of efficacy would risk discontinuing treatment in patients who would go on to respond. Therefore, they concluded that they would be evaluating patients at week 16 as per their usual clinical practice. However, the ERG does not agree that an early clinical assessment of efficacy would risk discontinuing patients' treatment. This is supported by the clinical data presented in the CS (Section 3.2.4.2), which reports peak response before week 12 in multiple outcomes. Given a 16 week assessment period is already in use for dupilumab it is uncertain whether response assessment in NHS clinical practice for baricitinib would be at 12 or 16 weeks

The primary endpoint in the BREEZE-AD4 (JAIN) trial was the proportion of patients in the intention-to-treat (ITT) population achieving EASI75 at week 16. The primary endpoint in the

BREEZE-AD7 (JAIY), BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) studies was the proportion of patients in the ITT population achieving IGA of 0 or 1 with a ≥ 2-point improvement at week 16.

Table 4 Summary of decision problem (adapted from CS Table 1)

	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 moderate-to-severe atopic dermatitis (AD) who are candidates for systemic therapy that had an inadequate response or intolerance to existing topical treatments.	Adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one systemic immunosuppressant due to intolerance, contraindication or inadequate disease control.	The population considered in this submission is most relevant to UK clinical practice as it is expected that clinicians will use baricitinib after considering a systemic immunosuppressant agent. It reflects the highest unmet clinical need for patients whose only treatment options are dupilumab or BSC. The eligibility criteria for the BREEZE-AD4 (JAIN) trial aligns with this patient population and is a subgroup of the full licensed population.	Clinical advice is that the population in the CS is restrictive as, in practice, baricitinib is likely to be used at the same point in the treatment pathway as other immunosuppressants, prior to dupilumab. Dupilumab targets the atopic/allergic inflammatory responses and provides more targeted immunomodulation. Whereas, baricitinib acts in a similar manner to other systemic immunosuppressants such as methotrexate and ciclosporin in targeting a broader range of cellular processes and mediators than dupilumab.[22] Therefore, in practice, it is expected that baricitinib would be given after topical treatment fails, when systemic immunosuppressants are considered. The National Eczema Society (NES) also states that patients with moderate to severe eczema for whom topical treatments are insufficiently effective and who must progress to second-line treatments would benefit from the introduction of a new second-line treatment option. Therefore, the population addressed in the CS may not be

				the most relevant and fully representative population for this indication. The ERG also notes that the inclusion criteria for the JAIN, JAIY, JAHL and JAHM clinical trials presented in the submission specified an EASI score ≥ 16, IGA score ≥ 3 and BSA involvement ≥ 10%. Although the IGA inclusion criteria ≥ 3 covers moderate to severe patients, the EASI inclusion criteria of ≥ 16 may exclude patients on the lower end of the moderate scale and bias the trial populations towards more severe disease.[1] [2]Therefore, in terms of disease severity, the ERG considers that the population in the clinical evidence presented may not represent all moderate to severe patients in the NHS population. This is discussed in more detail in Section 2.3.1
Intervention	Baricitinib with and without corticosteroids	Baricitinib with and without corticosteroids	N/A – in line with the NICE scope	N/A
Comparator(s)	• Phototherapy including ultraviolet B (UVB) radiation or psoralen-ultraviolet A (PUVA) • Systemic immunosuppressive therapies (azathioprine, ciclosporin, methotrexate	• Dupilumab • BSC (emollients, low-to-mid potency topical corticosteroids, phototherapy, psychological support, and rescue therapy including higher	The use of baricitinib in the UK is expected to be 5th line therapy following failure or contraindication of topical therapies, phototherapy and systemic immunosuppressant agents. This makes dupilumab and BSC the relevant	Clinical advice to the ERG is that baricitinib would be given after topical treatment and phototherapy, when systemic immunosuppressants are considered. As a JAK1/JAK2, baricitinib has a more targeted mode of action as compared with other systemic

	and mycophenolate mofetil) • Alitretinoin (in people with AD affecting the hands) • Dupilumab • Best supportive care (BSC)	potency topical or oral corticosteroids or topical calcineurin inhibitors)	comparators in UK clinical practice. Alitretinoin is not a relevant comparator based on its licensed indication and place in therapy for the treatment of severe hand eczema.	immunosuppressants, but a less targeted mode of action as compared with dupilumab. After failure to respond to a treatment, patients tend to move on to other available treatments. Therefore, comparators to baricitinib should not be restricted to dupilumab but should also include systemic immunosuppressants. Additionally, clinicians may move on to dupilumab as soon as possible due to the low side effect profile and the reduced need for monitoring. This indicates that there may also be scope for treatment with baricitinib after dupilumab. Alitretinoin is not relevant to the population addressed in the scope and therefore the ERG considers that the company's rationale for excluding alitretinoin is acceptable.
Outcomes	The outcome measures to be considered include: • Measures of disease severity • Measures of symptom control • Disease-free period/maintenance of remission • Time to relapse/prevention of relapse	The outcome measures to be included in the submission include: • Measures of disease severity and symptom control (including IGA, EASI scores, Itch NRS, Skin pain NRS) • Adverse effects of treatment (including AEs, SAEs, AESIs)	Whilst data for time-to-relapse and disease-free period are not explicitly available, evidence for maintenance of response is available for the population of interest from JAIN.	The ERG is satisfied with the outcomes considered and the reason for not reporting the time-to-relapse and disease-free period outcomes.

	• Adverse effects of treatment • Health-related quality of life	• Health-related quality of life (including EQ-5D-5L, DLQI, POEM, HADS, ADSS, WPAI-AD) • Maintenance of response (including IGA, EASI scores, Itch NRS, Skin pain NRS and HRQL outcomes)		
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. 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.	As per scope	N/A	N/A
Subgroups	If the evidence allows, the following subgroups will be considered. These include: • skin colour subgroups, • people with moderate dermatitis and those with severe dermatitis • people who are ciclosporin naïve and those who have previously received ciclosporin.	The subgroups specified in the NICE final scope were not considered in this submission.	Data were not available to conduct subgroup analyses for skin colour subgroups. The patient population considered in the submission will be adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to	It is evident that [redacted] patients were enrolled in any of the trials, making subgroup analysis by skin colour difficult to do. However, in all four trials (JAIN, JAIY, JAHL and JAHM) race was a pre-planned subgroup. In response to the points for clarification, the company stated that the trial program was not designed to investigate baricitinib efficacy in Black patients compared with

			intolerance, contraindication or inadequate disease control. As such, all patients can be considered to have moderate-to-severe AD, since systemic therapies are not considered until failure of topical treatments, phototherapy and photochemotherapy (psoralen-ultraviolet A [PUVA]). However, the clinical classification systems used to define AD severity are not consistent, with patients often receiving highly individualised treatment, and therefore defining separate subgroups of moderate AD and severe AD was not considered plausible or possible. In the patient population considered in the submission who have experienced failure with or are intolerant to or have contraindication to at least 1 systemic therapy, the vast majority of these patients will have received prior ciclosporin as ciclosporin is currently the only licensed systemic immunosuppressant for AD. Therefore, subgroup analyses based on ciclosporin-naivety was not considered relevant to the submission.	other patient populations and as such, the ethnicity distributions of the BREEZE-AD trials are reflective of the participating countries rather than of the occurrence of AD. The CS presented data on race including, Caucasian, Asian and other. Therefore, limited subgroups on race could have been presented. Furthermore, subgroups by region (Europe and Japan) were presented in Section 3.2.4.3, which may be considered a reasonable proxy for ethnicity. They reported a significant interaction, which indicates that outcomes for patients with different skin types are not the same. However, the evidence provided in the company's response to clarification suggests that the differences are not driven primarily by ethnicity, but rather by differences in the characteristics of the recruited patients and treatment practices. The ERG agrees that there is no consensus on defining severity for AD and measures can be inconsistent, however there are several widely accepted and commonly used classification systems. In all four trials presented in the CS baseline disease severity by IGA was a

				pre-planned subgroup. The CS has presented subgroup analyses for JAIY, JAHL and JAHM by IGA (3 or 4). There are published strata which allow classification by EASI score. In response to the PFC, the company stated that EASI does not reflect all aspects of moderate or severe disease and it does not provide consistent classification of disease severity. However, clinical advice to the ERG is that EASI is widely accepted and considers all relevant aspects of the clinical signs of AD. The ERG considers that although there are limitations to using one severity classification, it would have been plausible and beneficial to present separate subgroups of moderate and severe AD. The company's base case is relevant to the subgroup of people who have previously received ciclosporin as the population in the submission are patients who have experienced failure with or are intolerant or have contraindication to at least one systemic therapy. The majority of these patients will have received ciclosporin as it is the only licensed immunosuppressant. Therefore,

				the ERG agree that subgroup analyses based on ciclosporin-naivety are not relevant.
Special considerations including issues related to equity or equality	None identified	N/A – in line with the NICE final scope.	N/A – in line with the NICE final scope.	The ERG have identified that treatment efficacy may differ in people with different skin colours, particularly BAME patients, which could be a potential equality issue. Although there is data on Japanese and East Asian patients, there is no data reported on Black patients. For this reason, subgroup analyses on skin colour were not conducted and the efficacy of baricitinib in this population could not be established. The British Association of Dermatologists state that effects on different skin type (e.g. BAME skin types) should be considered as an equality issue for this indication.

AD: atopic dermatitis; BSC: best supportive care; CS: company submission; ERG: evidence review group; EASI: Eczema area and severity index; IGA: investigators global assessment; BSA: body surface area; N/A: not applicable; NRS: numeric rating scale; AE: adverse events; SAE: serious adverse events; AESI: adverse event of special interest; EQ-5D-5L: 5-level EuroQoL 5 dimensions; DLQI: dermatology life quality index; POEM: patient-orientated eczema measure; HADS: hospital anxiety depression scale; ADSS: atopic dermatitis sleep scale; WPAI-AD: work productivity and activity impairment questionnaire; HRQL: health related quality of life

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The CS describes a systematic literature review (SLR) of the clinical effectiveness and safety of baricitinib as well as relevant comparators for the treatment of adult patients with moderate to severe atopic dermatitis. Details of the SLR methods are presented in Appendix D of the CS.

3.1.1 Searches

The search strategies reported in the company submission appear to be appropriate to identify relevant trials of baricitinib and comparator therapies for adults with moderate-to-severe AD. Some weaknesses have been identified by the ERG, outlined in Table 5, which could have impacted on the comprehensiveness of the search.

Table 5 ERG appraisal of evidence identification for the clinical effectiveness review

Topic	ERG response	Note
Is the report of the search clear and comprehensive?	Yes	
Were appropriate sources searched?	Partly	Sources searched: MEDLINE, Embase, CENTRAL, ClinicalTrials.gov, conference abstracts, HTA agency websites. - WHO International Clinical Trials Registry Platform (ICTRP) was unavailable due to the current pandemic. - EU Clinical Trials register was not searched. - Reference checking of relevant reviews or included studies was not undertaken.
Was the timespan of the searches appropriate?	Yes	The database searches covered the period from database inception to 10th March 2020.
Were appropriate parts of the PICOS included in the search strategies?	Yes	Atopic dermatitis (P) AND RCTs (S)
Were appropriate search terms used?	Partly	Population terms could have been expanded to include the broader term eczema given the lack of standard terms for AD.[23] This may have ensured more comprehensive retrieval of relevant studies. This approach has been used in the searches for a recent living systematic review and NMA of treatments for atopic dermatitis.[10, 24]

Were any search restrictions applied appropriate?	Partly	Conference abstracts were removed from the search results in Embase.
Were any search filters used validated and referenced?	Unclear	Retrieval was restricted to RCTs in MEDLINE and Embase, however the source of the RCT study design search filters was not referenced, therefore it was unclear if the filters used were validated.

ERG response = YES/NO/PARTLY/UNCLEAR/NOT APPLICABLE

3.1.2 Inclusion criteria

Full eligibility criteria for the clinical SLR are presented in Table 6 of the CS Appendices. RCTs and systematic reviews that compared at least two of baricitinib and other relevant comparators, including placebo, in adult patients with moderate to severe atopic dermatitis, were included in the review. Studies were not restricted by outcomes reported and there were no date limits applied. Only English and German language publications were included.

The study selection methods described by the CS are appropriate. Two independent reviewers screened titles and abstracts using the inclusion criteria stated above. The full texts were then screened for inclusion, before decisions were compared, and any disagreements or queries were referred to a third independent reviewer.

A PRISMA flow diagram and a list of studies excluded from the systematic review, with reasons for exclusion, are included in the CS appendix D (Figure 1 and Table 9, respectively). The SLR included 40 unique studies. Two phase III randomised controlled trials (RCTs) of baricitinib (JAHL and JAHM)[25] were identified and a phase II RCT[26] of baricitinib was also identified in the SLR. However, this was not considered further due to the availability of more relevant data from phase III trials. Two further phase III RCTs (JAIN[27] and JAIY)[28] and one long-term extension study BREEZE-AD3 (JAHN)[29], which have not yet been published are presented in the submission, providing evidence for the efficacy and safety of baricitinib.

3.1.3 Critique of data extraction

The methods of data extraction are described on page 17 of Appendix D. Information for each included article was extracted by a single individual, in the first instance, checked against the publication and validated by a second independent reviewer. The ERG considers the methods to be appropriate and sound.

There are sufficient data from the four phase III trials: JAIN, JAIY, JAHL and JAHM presented in the submission. Study details, baseline characteristics and outcomes of JAHN are presented in Appendix M.

3.1.4 Quality assessment

Quality assessment of the trials was performed using a method adapted from the York Centre for Reviews and Dissemination Handbook.[30] The checklist covered randomisation, concealment of treatment allocation, similarity of baseline characteristics, blinding, imbalances in dropouts, completeness of outcome reporting and intention-to-treat analysis. Results of quality assessment of the JAHL, JAHM, JAIN and JAIY trials, included in the indirect treatment comparisons (ITC), are presented in Table 11 of Appendix D. This is discussed further in Section 3.2.3 of this report. The four baricitinib trials were considered to be of relatively good quality with low risk of bias. However, full justifications for the risk of bias decisions are not provided and there is no information given on how many reviewers undertook quality assessment.

3.1.5 Evidence synthesis

Results of the full ITT population of the four baricitinib trials are presented separately in section B.2.6 and as pooled analyses for the ITC in section B.2.8 of the CS.

The patient population in JAIN, and a sub-population of the JAIY trial, are in line with the relevant population for this submission; combination therapy (baricitinib plus topical corticosteroids) for patients with moderate to severe AD who are candidates for systemic therapy and who experience failure with, are intolerant to or have contraindication to ciclosporin. Therefore, a pooled population of JAIN + JAIN-like JAIY patients informs the base-case economic model. Sub-populations of patients who had a history of intolerance or inadequate response to ciclosporin in the JAHL and JAHM trials are pooled for evidence on baricitinib as monotherapy. These are used as scenario analyses in the economic model. The JAHN 52-week long-term extension study, which is also used in the economic model, is presented in Appendix L of the CS.

Safety data are presented for JAIN alone, and as an integrated safety analysis, which included pooled safety data from JAHL, JAHM, JAIY and the phase II study JAHG from week 0 to week 16.[31] Long-term safety data from week 0 to week 52 of the extension study JAHN are presented in Appendix M. An ITC was conducted to assess the clinical effectiveness of baricitinib versus dupilumab, which is described in Section 3.4 of this report.

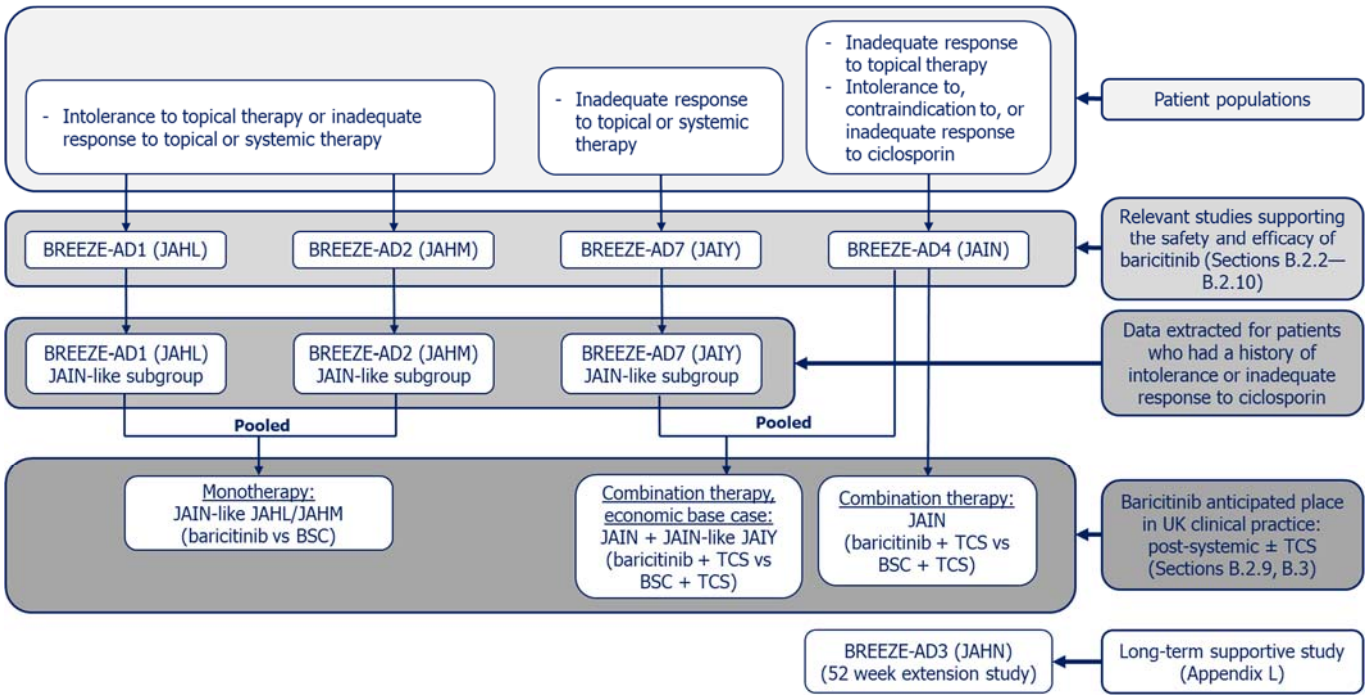
3.1.6 Ongoing studies

The JAHN and JAIN trials are currently ongoing. Additional data from JAIN may become available in October 2020 and in November 2020 from JAHN.

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

The company included two published trials: BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM), and two unpublished trials: BREEZE-AD4 (JAIN) and BREEZE-AD7 (JAIY) as well as an unpublished long-term extension study BREEZE-AD3 (JAHN). All four original trials were international multicentre, double blind, placebo-controlled phase III trials. The trials aimed to determine the efficacy and safety of baricitinib either in combination with TCS (JAIY and JAIN) or as a monotherapy (JAHL and JAHM) in adults with moderate-to-severe AD. The two monotherapy trials (JAHL and JAHM) are not as relevant as the combination trials because baricitinib is most likely to be given alongside TCS in practice. Therefore, the results of the JAIN and JAIY trials are discussed in more detail. The relationship between all five BREEZE-AD studies and how they inform the decision problem is described schematically in Figure 2.

Figure 2 A schematic representation of the BREEZE-AD trials informing the decision problem (from CS, Figure 4)



TCS: topical corticosteroids; BSC: best supportive care

The patient population in JAIN and a subgroup in JAIY are consistent with the population and intervention of interest for the submission, i.e. adults with moderate-to-severe AD who have failed at least one systemic immunosuppressant. For the indirect treatment comparison (ITC), data were extracted from the JAIY, JAHL and JAHM trials for the "JAIN-like" subgroups of patients who had a history of intolerance or inadequate response to ciclosporin. To maximise sample sizes, data were

pooled to produce relative treatment effects for baricitinib monotherapy (JAHL + JAHM JAIN-like) and baricitinib +TCS (JAIN + JAIY JAIN-like) compared to placebo and placebo + TCS, respectively.

3.2.1 Trial Designs and Methods

Details of the design and methodology of all trials are reported in Section 2.3.1 and Appendix D of the CS.

In all trials, randomisation was preceded by a screening period where patients were required to washout systemic and topical AD therapies in order to minimise possible confounding effects due to background treatment. However, patients were required to use emollients daily during the 14 days prior to randomisation and throughout the study.

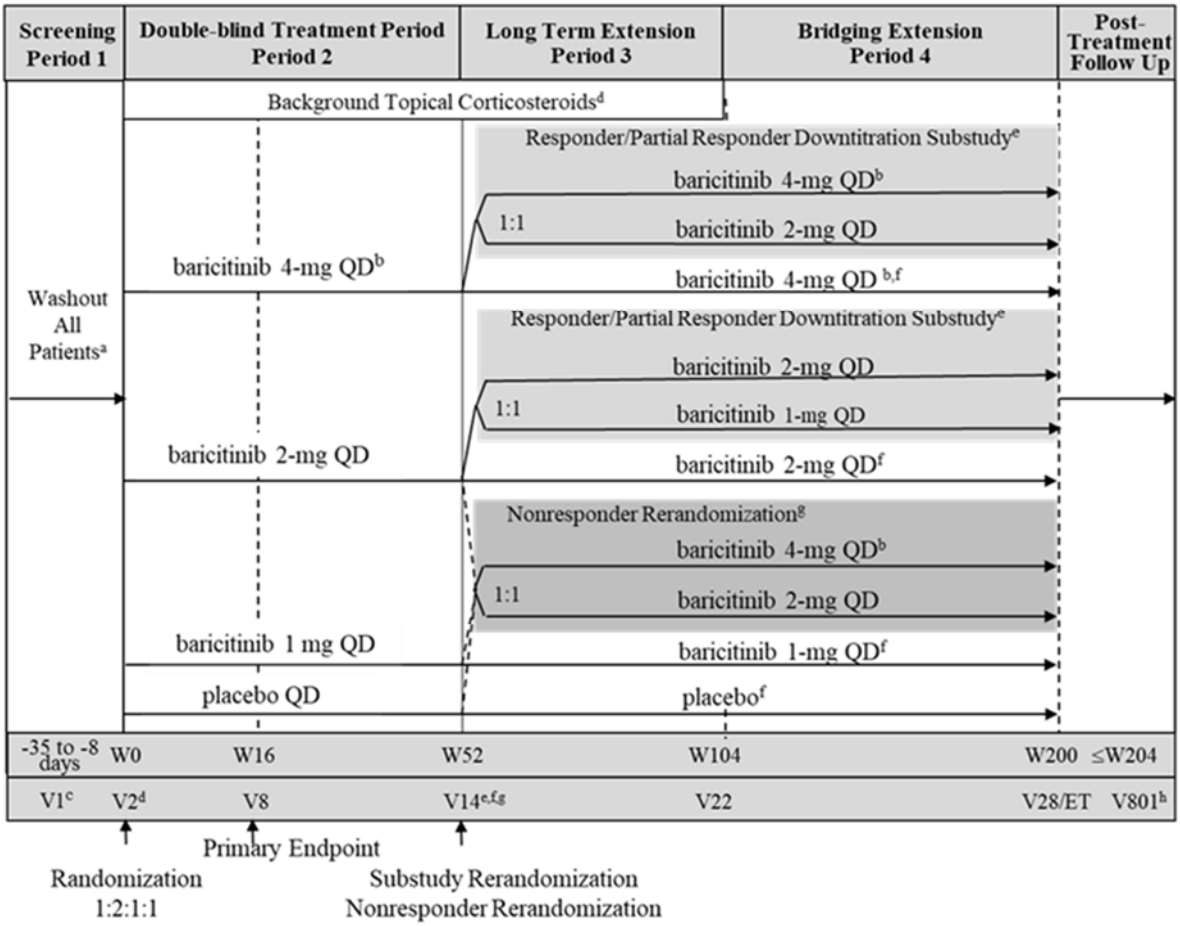
In JAIN and JAIY, patients used background TCS therapy (either triamcinolone 0.1% cream and/or hydrocortisone 2.5% ointment, or equivalent potency TCS/TCIs approved for AD in the country of trial) on active lesions.

In all trials, the placebo that was administered was not specified, but the company sourced three different placebo tablets that looked similar to each dose of baricitinib administered in a trial. Patients receiving an active treatment were given the placebo tablet for the doses that they were not randomised for.

3.2.1.1 BREEZE-AD4 (JAIN): Baricitinib + TCS

This trial was conducted in 103 sites in 14 different countries across Europe, South America and Asia. Six of these sites were located in the UK. The trial consisted of a 52-week treatment period, followed by a 52-week long-term extension period. Patients were randomised in a 1:1:2:1 ratio to receive placebo, 1 mg baricitinib, 2 mg baricitinib or 4 mg baricitinib, respectively. The study design is shown in Figure 3. The primary outcome was measured as the proportion of patients who achieved EASI75 at week 16 with either 2 mg or 4 mg baricitinib. However, only 4 mg baricitinib was used in the ITC and safety analyses. Safety outcomes were assessed at 16 and 24 weeks.

Figure 3 Study design for the BREEZE-AD4 (JAIN) trial (from CS, Figure 5)



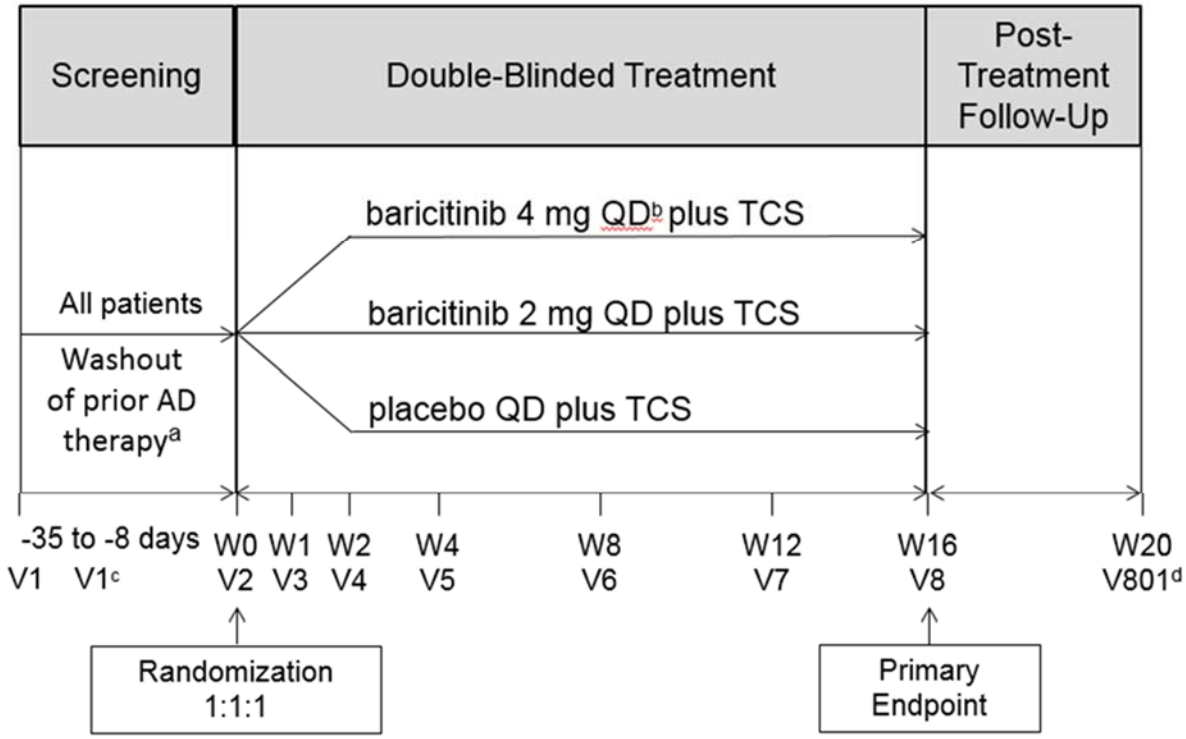
[a] Applicable to patients taking topical treatments (excluding emollients) or systemic treatments for AD at the time of screening. [b] Maximum dose of baricitinib for patients with renal impairment (defined as eGFR <60 mL/min/1.73 m^2) was 2 mg QD. [c] Patients for whom PPD skin test for the evaluation of TB infection was performed at V1 had to return and PPD test was read 48–72 hours after V1 (post-PPD). [d] At Visit 2 (Week 0), patients were supplied with mild- and moderate-potency TCS to be applied throughout the trial. [e] At Week 52, responders (IGA 0 or 1) and partial responders (IGA 2) who were assigned to baricitinib 4 mg or 2 mg at randomisation were enrolled into the down-titration study only if they did not have interrupted study drug at the time and had not used high- or ultra-high-potency TCS in the previous 14 days. If a patient in the sub-study had an IGA ≥3 at any time, they were retreated with their pre-sub-study baricitinib dose for the remainder of the study. [f] At Week 52, those who were in the baricitinib 1 mg or placebo groups and responders (IGA 0 or 1) and partial responders (IGA 2) in the baricitinib 4 mg or baricitinib 2 mg treatment groups who were not eligible for the randomised down-titration sub-study remained on their current dose of IP. If a patient had an IGA ≥3 at any time, except for patients in the baricitinib 4 mg group, they were rerandomised automatically at a 1:1 ratio to baricitinib 2 mg QD or baricitinib 4 mg QD. Re-randomisation occurred only once. Patients in the baricitinib 4 mg group remained on 4 mg. [g] Beginning at Visit 14 (Week 52), non-responders (IGA ≥3) in the placebo, baricitinib 1 mg or baricitinib 2 mg treatment groups were rerandomised at a 1:1 ratio to baricitinib 4 mg QD or baricitinib 2 mg QD. Non-responders randomised to baricitinib 4 mg at baseline remained on 4 mg. After re-randomisation, patients remained on the same dose of baricitinib for the remainder of the study. [h] Occurred approximately 28 days after the last dose of IP.
AD: atopic dermatitis; eGFR: estimated glomerular filtration rate; ET: early termination; IGA: Investigator's Global IP: investigational product; PPD: purified protein derivative; QD: once daily; TB: tuberculosis; TCS: topical corticosteroids; V: visit; W: week.
Source: BREEZE-AD4 (JAIN) Clinical Study Report.

3.2.1.2 BREEZE-AD7 (JAIY): Baricitinib + TCS

This trial was conducted in 68 sites in 10 countries in Europe, South America, Asia, and Australia. No patients were enrolled from the UK. Patients were randomised in a 1:1:1 ratio to treatment with placebo, 2 mg baricitinib or 4 mg baricitinib for a 16-week treatment period. However, only the 4 mg arm was used in the ITC and safety analyses. The study design is presented in Figure 4 and included a 4-week post-treatment follow-up period. The primary outcome was measured as the proportion of

patients who achieved IGA ≤ 1 with a ≥ 2-point improvement at week 16. Safety outcomes were assessed at 16 weeks.

Figure 4 Study design for the BREEZE-AD7 (JAIY) trial (from CS, Figure 6)



[a] Applicable to patients taking topical treatments (excluding emollients) or systemic treatments for AD at the time of screening. [b] For patients randomised to the 4 mg QD dose who had renal impairment (defined as eGFR <60 mL/min/1.73 m^2), the baricitinib dose was 2 mg QD. [c] Patients for whom PPD skin test for the evaluation of tuberculosis infection was performed at V1 had to return and PPD test was read 48–72 hours after V1 (post-PPD). [d] Occurred approximately 28 days after the last dose of the study treatment (was not required for those patients entering the long-term extension Study JAHN).
AD: atopic dermatitis; eGFR: estimated glomerular filtration rate; PPD: purified protein derivative; QD: once daily; V: visit; W: week.
Sources: BREEZE-AD7 (JAIY) Clinical Study Report.[28]

3.2.1.3 BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM): Baricitinib monotherapy

The two monotherapy trials recruited from sites internationally, however no patients were enrolled from the UK. Patients were randomised in a 2:1:1:1 ratio receiving placebo, baricitinib 1 mg, baricitinib 2 mg and baricitinib 4 mg, respectively. Only the 4 mg arms were used in the ITC and safety analyses. The treatment period was 16 weeks long with a 4-week post-treatment follow-up period. The primary outcome was measured as the proportion of patients who achieved IGA ≤ 1 with a ≥ 2-point improvement at week 16. Safety outcomes were assessed at 16 weeks.

3.2.1.4 Phase-II study JAHG: Baricitinib monotherapy

In response to the PFCs, the company provided details for the JAHG trial.[31] This was a Phase II, double blind, randomised, placebo-controlled study that recruited patients from 10 sites in the US and 3 sites in Japan. Patients were randomised in a 1:1:1 ratio to treatment with placebo, 2 mg and 4 mg

baricitinib for a 16-week treatment period followed by a 4-week post-treatment follow-up. This study was only used for assessing safety data.

3.2.2 Trial populations

The population of interest in the CS is adult patients who are candidates for systemic therapy who had a history of inadequate response to topical therapy and who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. Clinical advice to the ERG is that baricitinib may be given to patients at the same point in the treatment pathway as immunosuppressants, not just to patients who have a history of intolerance to, contraindication to, or inadequate response to ciclosporin. The National Eczema Society also states that patients with moderate to severe eczema for whom topical treatments are insufficiently effective and who must progress to second-line treatments would benefit from the introduction of a new second-line treatment option. Therefore, the ERG considers the base-case trial population to be restrictive and not fully representative of the population that could be given baricitinib, in practice.

Disease severity

The patient population was stated to be patients with moderate to severe AD, however the inclusion criteria may exclude patients on the lower end of the moderate scale as discussed in Table 4. Patients were required to have a baseline EASI score ≥ 16, IGA score ≥ 3 and involvement of ≥ 10% of the body surface area. A study by Chopra et al.[1] found that moderate disease was associated with EASI scores as low as 6 and a broad range of values (6.0–22.9). Therefore, the EASI inclusion criteria of a score ≥ 16 excludes patients on the lower end of the moderate scale and may bias the trial populations towards more severe disease. Furthermore, the mean EASI scores in all four trials ranged from [redacted] while published EASI strata for severe AD ranged from 21.1 to 50.[1, 2] Therefore, the mean EASI scores in the trials indicate that the patients included are more likely to have severe AD. The CS did not present subgroup analyses based on disease severity, although baseline disease severity by IGA was a pre-planned subgroup in all four trials. This is discussed further in Section 3.2.4.3.

Prior therapy

The vast majority of patients in the four baricitinib trials had prior TCS therapy (range: [redacted] - [redacted]) (Table 6). Clinical advice is that this is representative of TCS use in the NHS population. However, the JAIN trial had the lowest proportion of patients who received TCS therapy ([redacted] in the placebo arm and [redacted] in the baricitinib 4 mg arm). The majority of patients in the JAIN trial had prior ciclosporin use ([redacted] in the placebo arm and [redacted] in the baricitinib 4 mg arm) as this was part of the inclusion criteria. Whereas, this was lower in the JAIY, JAHL and JAHM trials (range: [redacted] to [redacted]). Prior phototherapy use was only reported for the JAIN trial, with nearly half of patients receiving phototherapy ([redacted] in the placebo arm and [redacted] in the baricitinib 4 mg

arm). Prior biologic use ranged from [redacted] to [redacted] in the four trials, with most patients receiving dupilumab. Clinical advice is that other biologics are typically only given experimentally.

ERG comment

The ERG considers that the population in the trials may not be fully representative of the NHS population eligible for systemic therapy who have had an inadequate response or intolerance to existing topical treatments.

Exclusion criteria for all four trials included patients currently experiencing, or who have a history of, other concomitant skin conditions, which would interfere with evaluation of the effect of the study medication on AD, or which require frequent hospitalisation and/or intravenous treatment for skin infections. Patients with eczema herpeticum within 12 months prior to screening or more than twice in the past, or with any serious concomitant illness anticipated to require the use of systemic corticosteroids or require active frequent monitoring, were also excluded. Additionally, the JAIN and JAIY trials excluded patients who have an important side-effect to TCS, which prevents further use. The key inclusion and exclusion criteria for JAIN are listed in Table 5 of the CS and for JAIY, JAHL and JAHM in Table 6 of the CS. Full inclusion and exclusion criteria were not presented in the CS but were accessible from the clinical study reports (CSR).

The baseline characteristics of the JAIN, JAIY, JAHL and JAHM trials are reported in Tables 8 to 11 of the CS, respectively. Patients included in the four trials were mostly comparable. Patients had a mean age of [redacted] years old and an average duration since diagnosis of approximately [redacted] years. Patients had a weight of [redacted] kg and a BMI of [redacted]. The proportion of the trial population who were female was slightly different between trials (range: [redacted]%). Overall, the baseline characteristics of the intention to treat (ITT) populations for the placebo and 4 mg baricitinib arms do not show any concerning imbalances across the treatment groups. In the JAIN trial, there were more female patients in the placebo group compared to the baricitinib 4 mg group. The JAIY trial had a higher proportion of patients who had prior systemic corticosteroid therapy in the placebo group than the baricitinib 4 mg group. The JAHM trial also had a higher percentage of patients who had prior systemic therapy in the placebo group than the baricitinib 4 mg group, particularly prior systemic corticosteroid therapy (Table 6).

The majority of patients in the JAIN, JAHL and JAHM trials were Caucasian. However, in the JAIY trial, [redacted]% of the placebo group, and [redacted]% of the baricitinib 4 mg group, were Asian (Table 6). AD presents differently in Asian patients, affecting both severity classification and response to treatment.[32] There were no data on the proportions of patients who were Black, and there were no subgroup analyses presented on race or skin colour, although in all four trials race was a pre-planned subgroup. This is discussed further in Section 3.2.4.3.

Patients included in the long-term extension JAHN trial were comparable to the other BREEZE-AD trials. Patients across the trial were largely Caucasian, representing [redacted]% of patients in all treatment arms in the main treatment phase and [redacted]% in the open-label addendum. There were no data on prior therapy use in the JAHN study. An overview of the baseline characteristics of patients included in the JAHN trial is presented in the CS (Table 51 of Appendix M).

Table 6 Baseline characteristics of the JAIN, JAIY, JAHL and JAHM trials (adapted from Tables 6 and 7 of the CS)

Characteristics	JAIN		JAIY		JAHL		JAHM	
	PBO + TCS (N=93)	4 mg + TCS (N=92)	PBO + TCS (N=109)	4 mg +TCS (N=111)	PBO (N=249)	4 mg (N=125)	PBO (N=[redacted])	4 mg (N=[redacted])
Age (years), mean (SD)	39 (14)	39 (13)	34 (13)	34 (11)	35 (12.6)	37 (12.9)	35 (13.0)	34 (14.1)
Female, %	47	38	35	32	101 (40.6)	42 (33.6)	90 (36.9)	41 (33.3)
Caucasian, %	80	77	42	49	147 (59.5) [a]	70 (56.5)	169 (69.3)	82 (66.7)
Asian, %	[redacted]	[redacted]	[redacted]	[redacted]	73 (29.6) [a]	41 (33.1)	72 (29.5)	38 (30.9)
Other, %	[redacted]	[redacted]	[redacted]	[redacted]	27 (10.9) [a]	14 (11.2)	3 (1.2)	2 (2.4)
Duration since AD diagnosis (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	26 (15.5)	25 (14.9)	25 (14)	23 (15)
Weight (kg), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	73 (15.7)	74 (17.2)	72 (16)	73 (15)
Body mass index (kg/m^2), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	25 (4.5)	25 (4.3)	25 (4.3)	25 (4.2)
Europe, %	[redacted]	[redacted]	[redacted]	[redacted]	135 (54.2)	68 (54.5)	111 (45.5)	56 (45.5)
Japan, %	[redacted]	[redacted]	[redacted]	[redacted]	45 (18.1)	22 (17.6)	45 (18.4)	23 (18.7)
Rest of world, %	[redacted]	[redacted]	[redacted]	[redacted]	69 (27.7)	35 (28.0)	88 (36.1)	44 (35.8)
IGA of 4 at screening Visit 1, %	54	51	44.4	45.0	[redacted]	[redacted]	[redacted]	[redacted]
IGA of 4 Visit 2, %	31 (11.6)	33 (13.7)	29 (12.3)	31 (12.6)	105 (42.2)	51 (40.8)	121 (49.6)	63 (51.2)
EASI, mean (SD)	69 (13.0)	69 (13.4)	67 (13.8)	68 (13.2)	32 (13.0)	32 (12.7)	33 (12.8)	33 (12.7)
SCORAD, mean (SD)	48 (21.3)	54 (23.8)	48 (24.4)	52 (23.3)	68 (14.0)	68 (13.0)	68 (12.7)	68 (13.6)
BSA affected, mean (SD)	21 (5.7)	21 (6.0)	21 (6.7)	21 (6.0)	53 (23.1)	52 (21.8)	52 (21.7)	54 (21.5)
POEM, mean (SD)	1.6 (1.6)	2.1 (1.8)	1.8 (2.0)	1.8 (2.3)	21 (5.6)	21 (5.6)	21 (6.3)	20 (6.3)
ADSS Item 2, mean (SD)	14.5 (6.9)	14.0 (8.1)	15 (7.9)	15 (7.9)	3.4 (5.2)	3.3 (5.2)	1.8 (2.1)	1.9 (2.5)
DLQI, mean (SD)	7.1 (1.9)	6.7 (2.3)	7.4 (1.7)	7.0 (2.0)	14 (7.4)	14 (7.1)	15 (8.1)	14 (8.4)
Itch NRS, mean (SD)	6.5 (2.3)	6.1 (2.6)	6.8 (2.3)	6.0 (2.5)	7 (2.0)	6 (2.0)	6.8 (2.2)	6.6 (2.2)

Skin Pain NRS, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	6 (2.5)	6 (2.4)	6.2 (2.5)	6.0 (2.6)
Prior TCS therapy, n (%) [a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior topical calcineurin inhibitor use, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior systemic therapy, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Systemic corticosteroid use	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Systemic immunosuppressant use	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ciclosporin use	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Biologic use, n (%) b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Only TCS use in the 12 months preceding screening was recorded. [b] Biologics use included 10 patients on dupilumab, 1 patient on lebrikizumab, 4 patients on nemolizumab, 1 patient on omalizumab, and 7 patients on tralokinumab.
Abbreviations: AD: atopic dermatitis; ADSS: Atopic Dermatitis Sleep Scale; BSA: body surface area; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; EQ-5D-5L: 5-level EuroQol 5 Dimensions; HADS: Hospital Anxiety Depression Scale; HIS: health index score; IGA: Investigator's Global Assessment; NRS: Numeric Rating Scale; PBO: placebo; PGI-S-AD: Patient Global Impression of Severity–Atopic Dermatitis; POEM: Patient-Oriented Eczema Measure; SCORAD: SCORing Atopic Dermatitis.; SD: standard deviation; TCS: topical corticosteroids; VAS: visual analogue score.

3.2.3 Quality Assessment

A summary of the quality assessment of the JAIN, JAIY, JAHL and JAHM trials is presented in Table 7. Full justifications for the risk of bias decisions were not provided. All four trials were RCTs with placebo arms. Randomisation appears to be appropriate, patients in all four trials were randomised by an interactive web response system. Re-randomisation at 52 weeks was also conducted using the interactive web response system. The concealment of treatment allocation for all three trials appears adequate.

Participants and investigators were blinded during the treatment phase and long-term extension phase of all four trials. Identical placebo tablets were used, minimising the risk of performance bias. There were few imbalances between treatment groups in the trials, which are described in Section 3.2.2. The number of patients who discontinued was similar among all arms in the JAIY, JAHL and JAHM arms. However, discontinuation in the JAIN trial was [redacted], ([redacted] vs [redacted], respectively).

The outcomes listed in the protocol match the ones reported in the trial clinical study reports (CSR), therefore the risk of selective outcome reporting is low. Intention-to-treat analysis, with non-responder imputation (NRI) or mixed model for repeated measures (MMRM) for missing data, was used for all analyses. The CS did not report the total proportion of missing values imputed, however

the number of patients rescued and therefore imputed as missing was similar across arms in the JAIN and JAIY trials, although this was higher in the placebo arms for the JAHL ([redacted]% in the placebo arm vs [redacted]% in the 4 mg baricitinib arm) and JAHM ([redacted]% in the placebo arm vs [redacted]% in the 4 mg baricitinib arm) trials. Overall, the ERG considers that the four baricitinib trials are of good quality with a low risk of bias.

Table 7 Quality assessment results for the baricitinib trials (from Table 15 of the CS)

	BREEZE-AD1	BREEZE-AD2	BREEZE-AD4	BREEZE-AD7
Was randomisation carried out appropriately?	Y	Y	Y	Y
Was the concealment of treatment allocation adequate?	Y	Y	Y	Y
Were the groups similar at the outset of the study in terms of prognostic factors?	Y	Y	Y	Y
Were the care providers, participants and outcome assessors blind to treatment allocation?	Y	Y	Y	Y
Were there any unexpected imbalances in dropouts between groups?	N	N	N	N
Is there any evidence to suggest that the authors measured more outcomes than they reported?	N	N	N	N
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Y	Y	Y	Y

N: no; NA: not applicable; RCT: randomised controlled trial; Y: yes.

3.2.4 Summary of the results of the included trials

3.2.4.1 Censoring

Efficacy endpoints for the BREEZE-AD trials were analysed using the following censoring rules:

i) ***Primary censoring rule:*** Continuous data were censored as missing and a MMRM used for analysis and dichotomous data had non-responder imputation after either

- permanent discontinuation of the study drug

  or

- the initiation of rescue therapy with TCS (any potency in the monotherapy trials or high or ultra-high potency in the combination trials) or systemic therapy.

This censoring rule is equivalent to using all the data up to rescue. Results for all outcomes for all trials using the primary censoring rule are reported in Section B.2.6 of the CS.

ii) ***Secondary censoring rule:*** Continuous data were censored as missing, with MMRM used for analysis and dichotomous data had non-responder imputation after

- permanent discontinuation of the study drug
  or
- the initiation of systemic rescues therapies.

Using the secondary censoring rule, data were not censored if patients were rescued using TCS alone. Clinical advice to the ERG is that this censoring rule is more likely to reflect clinical practice as it is expected that rescue medication will be used concomitantly with baricitinib. For patients on baricitinib a flare would not be considered an indicator of loss of response or grounds to discontinue treatment. A flare is considered a short event which can be managed with rescue therapy, with the patient either continuing on with treatment for AD or resuming treatment after stopping the rescue therapy. Clinical advice to the ERG was that in the case of a flare, control can be re-established on the same medication following rescue. Thus, data should not be censored after the initiation of rescue therapy.

3.2.4.2 Efficacy Outcomes

Results for primary and secondary outcomes assessing efficacy and quality of life for the JAIN, JAIY, JAHL and JAHM trials are presented in Section B.2.6 of the CS, including IGA, EASI, SCORAD, Itch NRS, Skin pain NRS, Item 2 of ADSS, DLQI and EQ-5D-5L. This section of the report focuses on the main outcomes that were included in the ITC and economic models. The ERG also considers skin pain NRS an important outcome, as skin pain has been highlighted by the ERG's clinical expert and the National Eczema Society (NES) as an important measure from the patients' perspective. However, it has not been included in the ITC or the health economic model. The main outcomes included were:

- **EASI score:** measures disease extent at four body regions: head and neck, trunk, upper limbs, and lower limbs. A higher score represents higher disease burden. EASI50, EASI75 and EASI90 represent an improvement of 50%, 75% and 90% in EASI score from baseline, respectively and are a dichotomous measure of the proportion of patients who have achieved a 50/75/90% improvement from the baseline score.
- **Itch NRS:** assesses the overall severity of itch experienced by patients within the last 24 hours. Higher scores represent a worse itch.
- **Skin pain NRS:** assesses the overall severity of skin pain experienced by patients within the last 24 hours. Higher scores represent worse pain.

- **DLQI:** assesses quality of life across six domains, where the higher the score the greater the impairment of life. The company assessed three DLQI outcomes – the mean change from baseline (MCFB) in DLQI, the proportion of patients who achieved a DLQI score of 0 or 1, and the proportion of patients who achieved a ≥ 4-point improvement in DLQI scores.

Results for IGA, SCORAD, Item 2 of ADSS and EQ-5D-5L are described in the CS. Categorical variables in the four trials were analysed using logistic regression, whereas continuous variables were analysed using MMRM. The statistical methods are detailed in Appendix L of the CS.

Combination Therapy Trial: BREEZE-AD4 (JAIN)

EASI

EASI75 at 16 weeks was the primary efficacy endpoint for the JAIN trial (Table 8). Patients treated with 4 mg baricitinib were more likely to achieve EASI75 compared to placebo at 16 weeks (OR: [redacted]) using the primary censoring rule. The results using the secondary censoring were consistent (OR: [redacted]). However, the difference between the 4 mg baricitinib and placebo groups was not statistically significant using either censoring rule for EASI75 at 24 weeks. Additionally, the ERG notes that peak response for EASI 75 was reached at week 8, as shown in Figure 8 of the CS.

The results for EASI50 were similar to those observed for EASI75. Patients in the 4 mg baricitinib group were more likely to achieve EASI50 at week 16 using the primary and secondary censoring rules. The difference between the two treatment groups was not statistically significant at 24 weeks using the primary censoring rule. No results were available for EASI50 at 24 weeks using secondary censoring (Table 8).

The difference between the proportion of patients who achieved EASI90 in the placebo and 4 mg baricitinib groups was not statistically significant at 16 or 24 weeks for either censoring rule (Table 8).

Table 8 Proportion of patients achieving EASI50, EASI75 and EASI90 for the JAIN trial (adapted from Tables 16, 17, 19, 20 and 21 of the CS)

		Week 16		Week 24	
		Placebo + TCS (N= 93)	Baricitinib 4 mg + TCS (N=92)	Placebo +TCS (N=93)	Baricitinib 4mg + TCS (N=92)
Primary Censoring					
EASI50	OR vs placebo (95% CI)		[redacted]		[redacted]

	p-value vs placebo	[redacted]	[redacted]
EASI75	OR vs placebo (95% CI)	[redacted]	[redacted]
	p-value vs placebo	[redacted]	[redacted]
EASI90	OR vs placebo (95% CI)	[redacted]	[redacted]
	p-value vs placebo	[redacted]	[redacted]
Secondary Censoring			
EASI50†	OR vs placebo (95% CI)	[redacted]	NA
	p-value vs placebo	[redacted]	NA
EASI75	OR vs placebo (95% CI)	[redacted]	[redacted]
	p-value vs placebo	[redacted]	[redacted]
EASI90	OR vs placebo (95% CI)	[redacted]	[redacted]
	p-value vs placebo	[redacted]	[redacted]

† Secondary censoring results for EASI50 at any time-point were not included in the relevant CSR. The secondary censoring results for 16 weeks were only stated in the CS.
OR: odds ratio; CI: confidence interval; NA: not available; TCS: topical corticosteroids

Itch NRS ≥ 4-point improvement at Weeks 1, 2, 4, 16 and 24

In the 4 mg baricitinib group, patients were more likely to achieve a ≥ 4-point improvement in NRS Itch scores compared to patients treated with placebo at week 2 (OR: [redacted]), week 4 (OR: [redacted]), week 16 (OR: [redacted]) and week 24 (OR: [redacted]) using primary censoring. Results using secondary censoring were consistent with those for primary censoring (Table 9). The ERG notes that the greatest proportion of patients achieved ≥ 4-point improvement in Itch NRS at week 5 in the baricitinib 4 mg arm, as shown in Figure 12 of the CS.

Table 9 Proportion of patients in the JAIN trial with ≥ 4 itch NRS at baseline achieving a ≥ 4-point Itch NRS improvement at Weeks 1, 2, 4 and 16 (adapted from Table 24 of the CS and Table 139 of the CSR).

		Primary Censoring		Secondary Censoring‡	
		Placebo + TCS (N=85)	Baricitinib 4 mg + TCS (N=[redacted])	Placebo + TCS (N=85)	Baricitinib 4 mg + TCS (N=[redacted])
Week 1	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 2	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 4	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 16	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 24	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]

†In Table 24 of the CS, N=[redacted], however, in the CSR, N=[redacted] for the baricitinib 4 mg group.
‡ Secondary censoring results extracted from the CSR for JAIN [27]
OR: odds ratio; CI: confidence interval; TCS: topical corticosteroids

Skin pain NRS mean change from baseline

The average mean change from baseline (MCFB) in skin pain NRS at weeks 16 and 24, using primary and secondary censoring is summarised in Table 10. Treatment with 4 mg baricitinib was associated with a significant improvement in the MCFB for skin pain NRS at 16 and 24 weeks using both censoring methods. Additionally, the ERG notes that the greatest improvement in the MCFB for skin pain NRS was seen at week 9 for the baricitinib 4 mg arm, as shown in Figure 13 of the CS.

Table 10 Mean change from baseline in Skin Pain NRS at week 16 and 24 in JAIN (adapted from Table 25 of CS and Table 147 of CSR)

	Week 16		Week 24	
	Placebo +TCS (N=93)	Baricitinib 4 mg + TCS (N=92)	Placebo +TCS (N=93)	Baricitinib 4 mg + TCS (N=92)
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]
Primary Censoring				
MCFB LSM	-1.56†	-3.02†	[redacted]	[redacted]
Mean Diff (95% CI)	[redacted]		[redacted]	
p	[redacted]		[redacted]	
Secondary Censoring				
MCFB LSM	[redacted]	[redacted]	[redacted]	[redacted]
Mean Diff (95% CI)	[redacted]		[redacted]	
p	[redacted]		[redacted]	

† The results presented in this table were reported by the company in the CS and in the CSR [27]in a table reporting skin NRS MCFB over a period of 0-16 weeks. Alternative results were presented in a table reporting over 0-24 weeks. While these results are slightly different (MCFB LSM for placebo =[redacted], MCFB LSM for Baricitinib 4 mg=-[redacted] and Mean Diff= [redacted]), the conclusions remain unchanged. The ERG believes that these discrepancies were possibly due to differences in imputation.

CI: confidence interval; MCFB: mean change from baseline; LSM: least-squares mean; Mean Diff: mean difference; TCS: topical corticosteroids

DLQI

The company assessed the mean change from baseline (MCFB) in DLQI, the proportion of patients who achieved a DLQI score of 0 or 1, and the proportion of patients who achieved a ≥ 4-point improvement in DLQI scores at week 16 and 24 using primary censoring. Results using secondary censoring were only presented for patients who achieved a ≥ 4-point improvement, and this was the outcome assessed in the ITC.

The number of patients being assessed differed across DLQI outcomes (Table 11). Treatment with 4 mg baricitinib was associated with a statistically significant improvement in MCFB DLQI and a higher proportion of patients achieved a DLQI score of 0 or 1 compared to placebo (Table 11).

Using primary censoring, patients in the 4 mg baricitinib arm were more likely to achieve a ≥ 4-point improvement in DLQI compared to those in placebo at week 16 (OR: [redacted]). The difference between the two arms at week 24 is not statistically significant. The results at week 16 using secondary censoring were consistent with the

results obtained using primary censoring, however, the magnitude of the effect was [redacted] (OR: [redacted]), due to the difference in the populations for the two censoring methods (Table 11). The ERG notes that the greatest improvement in MCFB and in the proportion of patients achieving ≥ 4-point improvement in DLQI score was seen at week 4 in the baricitinib 4mg arm, as shown in Figure 15 and 17 of the CS.

Table 11 DLQI outcomes at week 16 and 24 for patients in the BREEZE-AD4 (JAIN) trial (adapted from Table 27 of the CS)

	Week 16		Week 24	
	Placebo + TCS	Baricitinib 4 mg +TCS	Placebo + TCS	Baricitinib 4 mg +TCS
Primary Censoring				
N	93	92	93	92
Baseline Mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB				
N	93	92	93	92
MCFB LSM	-4.95	-7.95	[redacted]	[redacted]
Mean Diff (95% CI)		-3.00 [redacted]		[redacted] [redacted]
p		[redacted]		[redacted]
Score of 0 or 1				
N	93	92	93	92
OR (95% CI)		[redacted] [redacted]		[redacted] [redacted]
p		[redacted]		[redacted]
≥ 4- point improvement				
N	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted] [redacted]		[redacted] [redacted]
p		[redacted]		[redacted]
Secondary Censoring				
≥ 4- point improvement				
N	[redacted]	[redacted]	[redacted]	[redacted]
OR† (95% CI)		[redacted] [redacted]		[redacted]
p		[redacted]		[redacted]

†Secondary censoring results for ≥ 4-point improvement at any time-point were not reported in the relevant CSR. The secondary censoring results for 16 weeks were only stated in the CS.

OR: odds ratio; CI: confidence interval; MCFB: mean change from baseline; LSM: least-squares mean; Mean Diff: mean difference; NA: not available; TCS: topical corticosteroids

Combination Therapy Trial: BREEZE-AD7 (JAIY)

EASI

Patients treated with 4 mg baricitinib were more likely to achieve EASI50 (OR: [redacted]), EASI75 (OR: [redacted]) and EASI90 (OR: [redacted]) at 16 weeks. Results obtained using secondary censoring were consistent with those using primary censoring (Table 12). The ERG notes that peak response in EASI75 was seen at week 12 for the baricitinib 4mg arm, as shown in Figure 19 of the CS.

Table 12 Proportion of patients achieving EASI50, EASI75 and EASI90 for the BREEZE-AD7 (JAIY) trial at week 16 (adapted from Table 31 of the CS)

		Placebo + TCS (N= [redacted])	Baricitinib 4 mg + TCS (N= [redacted])
Primary Censoring			
EASI50	OR vs placebo (95% CI)		[redacted]
	p-value vs placebo		[redacted]
EASI75	OR vs placebo (95% CI)		[redacted]
	p-value vs placebo		[redacted]
EASI90	OR vs placebo (95% CI)		[redacted]
	p-value vs placebo		[redacted]
Secondary Censoring			
EASI50	OR vs placebo (95% CI)		[redacted]
	p-value vs placebo		[redacted]
EASI75	OR vs placebo (95% CI)		[redacted]
	p-value vs placebo		[redacted]
EASI90	OR vs placebo (95% CI)		[redacted]
	p-value vs placebo		[redacted]

OR: odds ratio; CI: confidence interval; TCS: topical corticosteroids

Itch NRS ≥ 4-point improvement at Day 2 and Weeks 1, 2, 4 and 16

The difference in patients achieving ≥ 4-point improvement in Itch NRS was not statistically significant at two days or at week 1 (Table 13). However, results at week 2 (OR: [redacted]), week 4 (OR: [redacted]), and week 16 (OR: [redacted]) achieved statistical significance.

Table 13 Proportion of patients in the BREEZE-AD7 (JAIY) trial with ≥ 4 itch NRS at baseline achieving a ≥ 4-point Itch NRS improvement at Weeks 1,2,4 and 16 (from Table 34 of the CS and Table 138 of the CSR)

		Primary Censoring		Secondary Censoring	
		Placebo + TCS (N=[redacted])	Baricitinib 4 mg + TCS (N=[redacted])	Placebo + TCS (N=[redacted])	Baricitinib 4 mg + TCS (N=[redacted])
Day 2	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 1	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 2	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 4	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]
Week 16	OR vs placebo (95% CI)		[redacted]		[redacted]
	p-value vs placebo		[redacted]		[redacted]

OR: odds ratio; CI: confidence interval; TCS: topical corticosteroid

Skin pain NRS mean change from baseline

The average MCFB in skin pain NRS at 16 weeks is presented in Table 14. Treatment with 4 mg baricitinib was associated with a statistically significant reduction in the skin pain NRS MCFB. The results were consistent using secondary censoring (Table 14).

Table 14 Mean change from baseline in skin pain NRS at week 16 in JAIY (from Table 35 of the CS and Table 153 of the CSR)

	Placebo +TCS ([redacted])	Baricitinib 4 mg + TCS ([redacted])
Baseline mean	[redacted]	[redacted]
Primary Censoring		
MCFB LSM	[redacted]	[redacted]
Mean Diff (95% CI)	[redacted]	
p	[redacted]	
Secondary Censoring		
MCFB LSM	[redacted]	[redacted]
Mean Diff (95% CI)	[redacted]	
P	[redacted]	

CI: confidence interval; MCFB: mean change from baseline; LSM: least-squares mean; Mean Diff: mean difference; TCS: topical corticosteroids

DLQI

The populations assessed for all outcomes and censoring methods were the same: there were [redacted] patients in the placebo arm and [redacted] patients in the 4 mg baricitinib group (Table 15). Therefore, the proportion of patients who experienced ≥ 4-point improvement in the DLQI score was higher than in the placebo arm, which was consistent between primary censoring (OR: [redacted]) and secondary censoring (OR: [redacted]).

Treatment with 4 mg baricitinib was associated with a statistically significant improvement in MCFB DLQI and a higher proportion of patients achieved a DLQI score of 0 or 1 compared to placebo (Table 15). The ERG notes that the greatest improvement in the MCFB in DLQI score and in the proportion of patients achieving ≥ 4-point improvement was seen at week 4 and 2, respectively (Figures 24 and 25 of the CS).

Table 15 DLQI outcomes at 16 weeks in the BREEZE-AD7 (JAIY) trial (adapted from Table 37 of the CS)

	Week 16	
	Placebo + TCS	Baricitinib 4 mg +TCS
N	[redacted]	[redacted]
Baseline Mean	[redacted]	[redacted]
Primary Censoring		
MCFB		
N	[redacted]	[redacted]
MCFB LSM	[redacted]	[redacted]
Mean Diff (95% CI)	[redacted]	
p	[redacted]	
Score of 0 or 1		
N	[redacted]	[redacted]
OR (95% CI)	[redacted]	
p	[redacted]	
≥ 4- point improvement		
N	[redacted]	[redacted]
OR (95% CI)	[redacted]	
p	[redacted]	
Secondary Censoring		
≥ 4- point improvement		
N	[redacted]	[redacted]
OR (95% CI)	[redacted]	
p	[redacted]	

OR: odds ratio; MCFB: mean change from baseline; Mean Diff: mean difference; CI: confidence interval; TCS: topical corticosteroid

Monotherapy Trials: BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM)

In their submission, the company described efficacy outcomes and safety data for the JAHL and JAHM- trials where baricitinib was administered as a monotherapy. However, it is unlikely that baricitinib would be used as a monotherapy in NHS practice, as AD is a complex condition that is not typically treated with just a single therapy. It is expected that TCS would be given alongside baricitinib in practice.

EASI

In the JAHL and JAHM trials, patients in the 4 mg baricitinib group were more likely to achieve EASI50, EASI75 and EASI90 than patients in the placebo group (Table 16). Results obtained using secondary censoring were consistent with those obtained using primary censoring, however for all three EASI outcomes in both studies, the magnitude of the estimates using primary censored data is larger.

Table 16 Proportion of patients achieving EASI50, EASI75, and EASI90 for the JAHL and JAHM trials at week 16 (adapted from Table 40 of the CS)

		BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)	
		Placebo (N= 249)	**Baricitinib 4 mg (N=125)**	**Placebo (N=244)**	**Baricitinib 4mg (N=123)**
Primary Censoring					
EASI50	OR vs placebo (95% CI)	[redacted]		[redacted]	
	p-value vs placebo	[redacted]		[redacted]	
EASI75	OR vs placebo (95% CI)	3.72 (2.01, 6.89)		4.41 (2.22, 8.76)	
	p-value vs placebo	<0.001		<0.001	
EASI90	OR vs placebo (95% CI)	4.13 (1.91, 8.91)		6.20 (2.42, 15.91)	
	p-value vs placebo	<0.001		<0.001	
Secondary Censoring					
EASI50	OR vs placebo (95% CI)	[redacted]		[redacted]	
	p-value vs placebo	[redacted]		[redacted]	
EASI75	OR vs placebo (95% CI)	[redacted]		[redacted]	
	p-value vs placebo	[redacted]		[redacted]	
EASI90	OR vs placebo (95% CI)	[redacted]		[redacted]	
	p-value vs placebo	[redacted]		[redacted]	

Itch NRS ≥ 4-point improvement at Weeks 1, 2, 4 and 16

In both trials, the proportion of patients who achieved ≥ 4-point improvement in the Itch NRS score in the 4 mg baricitinib group was statistically significantly higher compared to those in the placebo group for weeks 1, 2, 4 and 16 (Table 17).

Table 17 Proportion of patients in the BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) trials with ≥ 4 itch NRS at baseline achieving a ≥ 4-point Itch NRS improvement at Weeks 1,2,4 and 16 (adapted from Table 43 of the CS, Table 114 of the JAHL CSR and Table 114 of the JAHM CSR)

		Primary Censoring				Secondary Censoring			
		BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)		BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)	
		Placebo (N=222)	Baricitinib 4 mg (N=107)	Placebo (N=213)	Baricitinib 4 mg (N=107)	Placebo (N=222)	Baricitinib 4 mg (N=107)	Placebo (N=213)	Baricitinib 4 mg (N=107)
Week 1	OR vs placebo (95% CI)	31.93 (2.29, >99.99) †		6.65 (1.17, 37.99)		[redacted]		[redacted]	
	p-value vs placebo	0.010		0.033		[redacted]		[redacted]	
Week 2	OR vs placebo (95% CI)	88.26 (5.67, >99.99) †		11.03 (2.83, 42.90)		[redacted]		[redacted]	
	p-value vs placebo	0.001		<0.001		[redacted]		[redacted]	
Week 4	OR vs placebo (95% CI)	10.00 (4.07, 24.56)		9.93 (3.74, 26.37)		[redacted]		[redacted]	
	p-value vs placebo	<0.001		<0.001		[redacted]		[redacted]	
Week 16	OR vs placebo (95% CI)	4.80 (2.47, 9.32)		4.91 (2.22, 10.86)		[redacted]		[redacted]	
	p-value vs placebo	<0.001		<0.001		[redacted]		[redacted]	

† The confidence intervals are extremely wide due to no patients in the placebo arm of JAHL achieving a ≥ 4-point increase in the Itch NRS in the first two weeks
OR: odds ratio; CI: confidence interval

Skin pain NRS mean change from baseline

The average MCFB in skin pain NRS in the JAHL and JAHM trials is summarised in Table 18. Treatment with 4 mg baricitinib in both trials was associated with a statistically significant improvement in skin pain NRS MCFB compared to placebo using both censoring methods.

Table 18 Mean change from baseline in skin pain NRS at week 16 in BRREZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) (adapted from Table 44 of the CS, Table 134 of the JAHL CSR and Table 133 of the JAHM CSR)

	BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)	
	Placebo +TCS (N=249)	Baricitinib 4 mg + TCS (N=125)	Placebo +TCS (N=244)	Baricitinib 4 mg + TCS (N=123)
Baseline mean	6.07	5.74	6.21	5.95
Primary Censoring				
MCFB LSM	-0.84	-1.93	-0.86	-2.49
Mean Diff (95% CI)		-1.09 (-1.79, -0.39)		-1.63 (-2.37, -0.87)
p		0.002		<0.001
Secondary Censoring				
MCFB LSM	[redacted]	[redacted]	[redacted]	[redacted]
Mean Diff (95% CI)		[redacted]		[redacted]
p		[redacted]		[redacted]

CI: confidence interval; MCFB: mean change from baseline; LSM: least-squares mean; Mean Diff: mean difference; TCS: topical corticosteroids

DLQI

The number of patients assessed were the same for all outcomes. The proportion of patients who achieved ≥ 4-point improvement in the DLQI score with baricitinib was significantly higher than those with placebo in both trials. Treatment with 4 mg baricitinib was also associated with a statistically significant improvement in MCFB DLQI and a higher proportion of patients achieved a DLQI score of 0 or 1 compared to placebo (Table 19).

Table 19 DLQI outcomes at Week 16 in the BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) trials (adapted from Table 46 and 47 of the CS)

	BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)	
	Placebo	Baricitinib 4 mg	Placebo	Baricitinib 4 mg
Primary Censoring				
N	[redacted]	[redacted]	[redacted]	[redacted]
Baseline Mean	[redacted]	[redacted]	[redacted]	[redacted]
MCFB				
N	[redacted]	[redacted]	[redacted]	[redacted]
MCFB LSM	[redacted]	[redacted]	[redacted]	[redacted]
Mean Diff (95% CI)		[redacted]		[redacted]

p		[redacted]		[redacted]
Score of 0 or 1				
N	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]		[redacted]
p		[redacted]		[redacted]
≥ 4- point improvement				
N	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]		[redacted]
p		[redacted]		[redacted]

OR: odds ratio; MCFB: mean change from baseline; Mean Diff: mean difference; CI: confidence interval

3.2.4.3 Subgroup Analysis

The trials included a range of pre-specified subgroup analyses, listed in Tables 5 and 6 of the CS. Subgroup analyses were performed on the pooled monotherapy population (JAHL and JAHM) and on the combination therapy JAIY trial. The proportion of patients achieving IGA ≤ 1, EASI75 or a ≥ 4-point improvement in Itch NRS at week 16 for subgroups with significant interactions ($p < 0.1$) is presented in Table 49 of the CS. The CS did not present subgroup data for the JAIN trial as the company stated they were not available at the time of submission. However, in response to the PFC, the company provided the proportion of patients achieving IGA ≤ 1, EASI75 or a ≥ 4-point improvement in Itch NRS at Week 16 for subgroups with significant interactions ($p < 0.1$) for the JAIN trial and for the JAIN + JAIN-like JAIY population.

Region & Skin colour

One of the subgroups presented by the company is region, specifically Europe, Japan and rest of world (ROW). In the JAIY trial, and the JAIN + JAIN-like JAIY population, significant interactions were observed for the region and specific region subgroups for EASI75. For the region subgroup of the JAIY trial, European patients had a greater response with baricitinib compared with placebo (risk ratio [RR]: [redacted]) than Japanese patients (RR: [redacted]) or ROW (RR: [redacted]). Similarly, for the specific region subgroup non-Japanese patients responded better with baricitinib relative to placebo (RR: [redacted]) than Japanese patients (RR: [redacted]). This effect was also seen in East Asian patients (RR: [redacted]) versus all other patients (RR: [redacted]). There were significant interactions observed for Itch NRS ≥ 4-point improvement for the JAIY trial, for East Asian patients vs all other patients (Table 20). In the JAIN + JAIN-like JAIY population, a similar significant interaction was seen for EASI75 and European patients (RR: [redacted]) vs Japanese patients (RR: [redacted]) vs ROW (RR: [redacted]).

In response to the points for clarification, the company provided observations other than geographical factors that may impact on treatment efficacy which may explain why there is a greater response in European patients. Clinical practice, particularly the use of rescue TCS, which has a high potency, is different in Europe compared with Japan. In Europe clinical practice broadly limits the use of high potency TCS, whereas Japan favours it. The rescue rates observed in Europe ([redacted]%) were much lower than in Japan ([redacted]%) and the rest of the world ([redacted]%). Therefore, response rates, as assessed by the primary censoring rule, which censors following the use of rescue therapy or permanent study drug discontinuation, was found to be higher in the European population. The company also suggested that previous insufficient use or potency of TCS in non-European countries may be why there is a higher response rate in the placebo + TCS arm. However, the ERG is unsure whether this explains the higher response rate in the placebo + TCS arm in non-European patients. The company noted that some phenotypes, specifically the Th17 axis are more prevalent in Asian patients compared with European American patients and provided exploratory analyses, which showed a covariate effect for Japanese patients vs non-Japanese patients in the JAHL and JAHM monotherapy trials for IGA 0 or 1. Differences in baseline EASI score and BSA were noted, indicating that Japanese patients have more severe disease. The company also provided baseline characteristics for the JAIN + JAIN-like JAIY population, stratified by region. It reported that Japanese patients had a higher EASI score ([redacted] in the placebo arm and [redacted] in the baricitinib 4 mg arm) than European patients ([redacted] in the placebo arm and [redacted] in the baricitinib 4 mg arm). Additionally, Japanese patients had a higher SCORAD score and BSA than European patients. The observations suggest that European patients and Japanese patients have different baseline severity and different treatment practices which can lead to differences in treatment efficacy. This is a source of uncertainty, indicating that the trial populations are not fully generalisable to the NHS population, which should be considered when interpreting the results.

Notably, the company provided results of subgroup analyses for the proportion of patients achieving IGA 0 or 1 at Week 16 by region for the monotherapy and JAIY trials. They did not report a statistically significant difference in response using baricitinib 4 mg in East Asian countries not including Japan compared with Europe or Global regions, suggesting there is not a specific effect of East Asian ethnicity. The evidence provided in the company's response, plausibly outlines alternative explanations for the observed differences in Japanese patients. The evidence from non-Japanese East Asian patients, appears to suggest that the differences are not driven primarily by ethnicity, but rather other characteristics of the recruited Japanese patients.

Skin colour was a subgroup specified by NICE in the final scope. However, the CS stated that data were not available to conduct subgroup analyses for skin colour. [redacted] were enrolled in any of the pivotal trials. In response to the points for clarification, the company stated that the trial

program was not designed to investigate baricitinib efficacy in Black patients compared with other patient populations. Patients were not recruited from the US; if they had been, it would be expected that a higher proportion of Black patients would have been included. The company note that there is some evidence that the pathology of AD could be distinct in Black patients: mutations in the FLG gene leading to a deficiency in filaggrin have been associated with AD that is more severe and persistent than its wild-type counterpart. These mutations are detected in up to 30% of individuals, but they are rarely identified in African-American populations with AD.[32] The differences in the cytokine pathways involved in atopic dermatitis across ethnic groups were also noted in the dupilumab appraisal (TA534),[20] but the Appraisal Committee considered that there was insufficient evidence to determine the extent to which different cytokine pathways modify treatment effect. For this reason, the company did not consider this further. However, the ERG notes that this is a potential equalities issue as the lack of data on Black patients means that it is not possible to establish baricitinib efficacy in this population (Table 4)

Severity of AD

The CS reported a significant interaction between EASI75 and baseline IGA score in the pooled JAHL and JAHM population. Patients with an IGA score of 3 responded better on baricitinib relative to placebo (RR: [redacted]) than patients with an IGA score of 4 (RR: [redacted]) (Table 20). Patients with moderate AD and severe AD was a specified subgroup in the NICE scope. However, no clinical evidence was presented for the subgroups. The company submission states that this is because of the lack of a widely accepted classification system. However, there are published strata that allow classification by EASI score. In response to the points for clarification, the company stated that EASI does not reflect all aspects of moderate to severe disease, as AD is a flaring disease, and thus EASI score alone would not provide consistent classification of disease severity. The company reported that feedback from UK clinicians, experienced in the treatment of AD, confirmed that strata based on EASI score would not be used in isolation to inform treatment decisions for patients with moderate and severe disease in UK clinical practice. Furthermore, no subgroup analyses for moderate versus severe AD are available for dupilumab, and thus it was not feasible to conduct any efficacy comparisons with the key comparator in these populations. For these reasons, the company considered it inappropriate and infeasible to conduct subgroup analyses based on AD severity. However, clinical advice to the ERG is that EASI is widely accepted and considers all relevant aspects of the clinical signs of AD. Therefore, it would have been plausible and beneficial to conduct subgroup analyses based on AD severity.

Ciclosporin failure

The CS reported an interaction between EASI75 and ciclosporin failure for the JAIY population. A higher proportion of patients with no previous ciclosporin failure responded to baricitinib treatment

relative to placebo (RR: [redacted]), compared with patients with previous ciclosporin failure (RR: [redacted]) (Table 20).

Gender

There were significant interactions between IGA ≤ 1 and gender in the JAIY population and in the JAIN+JAIN-like JAIY population. Female patients had a greater response to baricitinib relative to placebo than male patients (Table 20). There was a significant interaction between EASI75 and gender in the JAIY population, with a higher proportion of male patients achieving EASI75 (RR: [redacted]) with baricitinib than female patients (RR: [redacted]). In the JAIN population, a higher proportion of male patients achieved a ≥ 4-point Itch NRS improvement with baricitinib compared with placebo (RR: [redacted]) than female patients (RR: [redacted]).

Table 20 Subgroup analyses with significant interactions at week 16 (adapted from Table 49 of the CS)

Outcome	Subgroup	Category	PBO	4 mg BARI	RR vs PBO	p-value
Combination therapy: JAIN						
Itch NRS improvement ≥ 4-point	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	
Combination therapy: JAIY						
IGA ≤1	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]
EASI75	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]
	Region	Europe (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		Japan (N=[redacted])	[redacted]	[redacted]	[redacted]	
		ROW (N=[redacted])	[redacted]	[redacted]	[redacted]	
	Specific region	Japan (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		Not Japan (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]
	Specific region	East Asia (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		All other (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]
Itch NRS improvement ≥ 4-point	Specific region	East Asia (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		All other (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]

Pooled monotherapy: JAHL and JAHM						
EASI75	Baseline IGA score	IGA 3 (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		IGA 4 (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Relative risk	[redacted]	[redacted]	[redacted]	[redacted]
Pooled combination therapy: JAIN + JAIN-like JAIY patients						
IGA ≤ 1	Gender	Male (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		Female (N=[redacted])	[redacted]	[redacted]	[redacted]	
EASI75	Region	Europe (N=[redacted])	[redacted]	[redacted]	[redacted]	
		Japan (N=[redacted])	[redacted]	[redacted]	[redacted]	[redacted]
		ROW (N=[redacted])	[redacted]	[redacted]	[redacted]	

3.2.5 Adverse Events (AEs)

The company investigated the safety of 4 mg baricitinib with or without the concurrent use of TCS in comparison to placebo for up to 16 weeks. The safety analysis looked at two separate datasets: safety data for the combination therapy JAIN trial (N=93 in placebo and N=92 in 4 mg baricitinib), and an 'integrated safety analysis' dataset. The integrated safety analysis comprised of patients from the two monotherapy trials (JAHL and JAHM), a combination therapy Phase II trial (JAHG) and the combination therapy trial JAIY. A breakdown of the patients included in the integrated safety analysis is presented in Table 21.

Table 21 Overview of patients that contribute to the integrated safety analysis (from Table 71 of the CS

Study	PBO ± TCS	BARI 4 mg ± TCS
JAHG	[redacted]	[redacted]
JAHL	249	125
JAHM	244	123
JAIY	[redacted]	[redacted]
Total	[redacted]	[redacted]

PBO: Placebo; BARI: Baricitinib; TCS: Topical corticosteroid

In their safety assessment, the company reported the following AEs:

(i) **Treatment-emergent adverse events (TEAEs):** These were defined by the company as untoward medical events that emerged or worsened during the treatment period and were not causally related to the treatment.

(ii) **Serious adverse events (SAEs):** SAEs were defined as any AE which resulted in death, a life-threatening experience, persistent or significant disability or incapacity, a congenital

abnormality or birth defect or any important medical event which jeopardises the patient or requires intervention to prevent any of these outcomes.

(iii) Adverse events leading to permanent discontinuation from study treatment: Patients were permanently discontinued from the baricitinib treatment arm if treatment had to be ceased due to medical, safety, regulatory or for reasons consistent with applicable laws, regulations and good clinical practice, or if patients required treatment with any systemic therapeutic agent that is not allowed as part of a rescue therapy. The criteria for permanent discontinuation are presented in detail in Appendix L.

(iv) Adverse events of special interest (AESI): The company defined AESIs as infections, malignancies, hepatic events (as defined by abnormal clinical liver tests), major cardiovascular events (including myocardial infarction or stroke), and thrombotic events (including deep vein thrombosis)

3.2.5.1 Treatment-emergent adverse events (TEAEs)

In both the integrated analysis and JAIN, although differences were not clinically meaningful, patients in the 4 mg baricitinib group experienced a higher proportion of TEAEs than in the placebo group.

In JAIN, at 16 weeks, 75.0% (n=69) of the patients in the 4 mg baricitinib group experienced at least one TEAE compared to 53.8% (n=50) of patients in the placebo group. In the 1 and 2 mg baricitinib groups, [redacted] and [redacted] experienced at least one TEAE, respectively. Between 16 weeks and 24 weeks, an additional [redacted] TEAEs were experienced in the 4 mg baricitinib group, so that [redacted] of the patients had experienced a TEAE. In the 1 mg baricitinib group [redacted] and the 2 mg baricitinib group [redacted] of patients experienced at least one TEAE.

At 16 weeks, the most commonly observed TEAEs in all four treatment arms were nasopharyngitis (Placebo: 12.9%, 1 mg baricitinib: [redacted]2 mg baricitinib: [redacted], 4 mg baricitinib: 26.1%). Other TEAEs that were observed in > 3% of the patients in placebo, 1 mg baricitinib, 2 mg baricitinib and 4 mg baricitinib at 16 weeks are detailed in Table 22. A summary of commonly observed TEAEs at 24 weeks is presented in the Appendix (Table 66).

Table 22 Commonly observed treatment-emergent adverse events (TEAEs) in JAIN at 16 weeks (adapted from Table 73 of the CS and Table 186 from the CSR)

TEAEs affecting > 3% of patients, n (%)	PBO + TCS (N=93)	BARI 1 mg + TCS ([redacted])	BARI 2 mg +TCS ([redacted])	BARI 4 mg +TCS (N=92)
TEAEs at 16 Weeks				
≥ 1 TEAE	50 (53.8)	[redacted]	[redacted]	69 (75.0)
Nasopharyngitis	12 (12.9)	[redacted]	[redacted]	24 (26.1)
Headache	6 (6.5)	[redacted]	[redacted]	7 (7.6)
Influenza	2 (2.2)	[redacted]	[redacted]	6 (6.5)
Upper abdominal pain	2 (2.2)	[redacted]	[redacted]	5 (5.4)
Diarrhoea	3 (3.2)	[redacted]	[redacted]	5 (5.4)
Oral herpes	3 (3.2)	[redacted]	[redacted]	5 (5.4)
Folliculitis	[redacted]	[redacted]	[redacted]	[redacted]
Herpes simplex	[redacted]	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Skin infection	[redacted]	[redacted]	[redacted]	[redacted]
Peripheral oedema	0	[redacted]	[redacted]	4 (4.3)
Abdominal pain	3 (3.2)	[redacted]	[redacted]	3 (3.3)
Back pain	3 (3.2)	[redacted]	[redacted]	3 (3.3)
Asthma	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Dry eye	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]
Blood creatine phosphokinase increased	[redacted]	[redacted]	[redacted]	[redacted]
Oropharyngeal pain	[redacted]	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]	[redacted]

TEAE: treatment-emergent adverse event; PBO: placebo; BARI: baricitinib; TCS: topical corticosteroid

In the integrated analysis, [redacted] of the patients in the 4 mg baricitinib group experienced at least 1 TEAE, compared to [redacted] in the placebo group. The most commonly observed TEAEs in the placebo and 4 mg baricitinib groups are summarised in Table 23.

Table 23 Commonly observed TEAEs in the integrated safety analysis dataset (adapted from Table 73 of the CS)

TEAEs affecting > 3% of patients, n (%)	PBO ± TCS ([redacted])	BARI 4 mg ±TCS ([redacted])
≥ 1 TEAE	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Blood creatine phosphokinase increased	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]

TEAE: treatment-emergent adverse event; PBO: placebo; BARI: baricitinib; TCS: topical corticosteroid

3.2.5.2 Serious Adverse Events (SAEs)

A higher proportion of patients in the 4 mg baricitinib group (6.5%) in JAIN experienced at least 1 SAE compared to placebo (2.2%). However, in the integrated analysis, [redacted] of the patients in the 4 mg baricitinib group experienced at least one SAE compared to [redacted] of the patients in the placebo group. These differences were not clinically meaningful. The number of patients who experienced at least one SAE in JAIN in all treatment arms at 16 and 24 weeks and a summary of the SAEs experienced are presented in Table 24. A summary of the SAEs experienced in the placebo and 4 mg baricitinib treatment groups of the integrated safety analysis is provided in Table 25.

Table 24 Serious Adverse Events (SAEs) observed in the JAIN trial (adapted from Table 74 of the CS and Table 210 of the CSR)

		PBO + TCS (N=93)	BARI 1 mg + TCS ([redacted])	BARI 2 mg +TCS ([redacted])	BARI 4 mg +TCS (N=92)
16 Weeks	Patients with ≥ 1 SAE, n (%)	2 (2.2)	[redacted]	[redacted]	6 (6.5)
	SAEs	Erysipelas (n=1) Atopic dermatitis (n=1) Bowen's Disease (n=1)	[redacted]	[redacted]	Pyelitis (n=1) Staphylococcal infection (n=1) Atopic dermatitis (n=2) Allergic conjunctivitis (n=1) Soft tissue inflammation (n=1) Ligament rupture (n=1)
24 Weeks	Patients with > 1 SAE, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
	SAEs*		[redacted]	[redacted]	

* SAEs reported here are new events observed between 16 and 24 weeks
PBO: placebo; BARI: baricitinib; TCS: topical corticosteroids; SAE: serious adverse events

Table 25 Serious adverse events (SAEs) observed in the integrated safety analysis (adapted from Table 74 of the CS)

	PBO ± TCS ([redacted])	BARI 4 mg ±TCS ([redacted])
Patients with ≥ 1 SAE, n (%)	[redacted]	[redacted]
Adverse Events	[redacted]	[redacted]

PBO: placebo; BARI: baricitinib; TCS: topical corticosteroids

3.2.5.3 Adverse events leading to permanent discontinuation

The company did not consider any AE to be cause for concern as AEs varied between the treatment groups in the JAIN trial and the integrated analysis and the occurrence of adverse events was relatively balanced between all treatment arms.

AEs that resulted in permanent discontinuation at 16 weeks in JAIN are summarised in Table 26. At week 24, an additional [redacted] patients had discontinued due to AEs. [redacted] in the baricitinib 1 mg arm discontinued due to [redacted]; and [redacted] in the baricitinib 4 mg arm discontinued due to [redacted]. [redacted] in the baricitinib 2 mg arm discontinued due [redacted].

Table 26 Adverse events (AEs) leading to permanent discontinuation in the JAIN trial at 16 weeks (adapted from Table 75 in the CS and Table 214 from the CSR)

	PBO + TCS N=93	BARI 1-mg + TCS [redacted]	BARI 2-mg + TCS [redacted]	BARI 4-mg + TCS N=92
AEs leading to permanent DC from study treatment, n (%)	1 (1.1)	[redacted]	[redacted]	1 (1.1)
Blood alkaline phosphatase increased	1 (1.1)	[redacted]	[redacted]	0
Skin infection	0	[redacted]	[redacted]	1 (1.1)
Abdominal pain	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Atopic dermatitis	[redacted]	[redacted]	[redacted]	[redacted]

AEs: adverse events; DC: discontinuation; PBO: placebo; BARI: baricitinib; TCS: topical corticosteroid

A summary of the AEs that resulted in permanent discontinuation in the integrated safety analysis is summarised in Table 75 of the CS. In the 4 mg baricitinib group, the most common AEs that lead to permanent discontinuation were toxic skin eruptions (n = [redacted] and a decrease in white blood cell count (n = [redacted] whereas lymphopenia (n = [redacted] and dizziness (n = [redacted] lead to discontinuation in the placebo arm.

3.2.5.4 Adverse events of special interest (AESI)

AESIs that were observed in JAIN are summarised in Table 27. The most common AESIs were treatment emergent (TE) infections. A higher proportion of patients experienced TE infections in the 4 mg baricitinib arm ([redacted] compared to placebo ([redacted], 1 mg baricitinib ([redacted] and 2 mg baricitinib [redacted]). In particular, a higher proportion of patients in the 4 mg baricitinib group experienced herpes simplex ([redacted] compared to placebo ([redacted], 1 mg baricitinib [redacted] and 2 mg baricitinib [redacted].

Table 27 Adverse events of special interest (AESI) at 16 weeks in the JAIN trial (adapted from Table 76 of the CS and Table 90 of the CSR)

	PBO + TCS N=93	BARI 1-mg + TCS [redacted]	BARI 2-mg + TCS [redacted]	BARI 4-mg + TCS N=92
Any TE infection	[redacted]	[redacted]	[redacted]	[redacted]
Serious infections	[redacted]	[redacted]	[redacted]	[redacted]
Opportunistic infection	[redacted]	[redacted]	[redacted]	[redacted]
Herpes zoster	[redacted]	[redacted]	[redacted]	[redacted]
Herpes simplex	[redacted]	[redacted]	[redacted]	[redacted]

TE: treatment-emergent; PBO: placebo; BARI: baricitinib; TCS: topical corticosteroid

AESIs for the integrated safety analysis are presented in Table 76 of the CS. The most common AESIs were TE infections, experienced by [redacted] of the patients in the placebo arm and [redacted] of the 4 mg baricitinib arm. Similar to the JAIN trial, the most commonly observed infection was herpes simplex (observed by [redacted] of the patients in the placebo group and [redacted] of the patients in the 4 mg baricitinib group).

3.2.5.5 Overview of adverse events

An overview of the AEs observed is presented in Table 28. No deaths were observed in any of the trials for the 16-week duration, or up to 24 weeks for JAIN.

Table 28 Overview of AEs observed in JAIN, the integrated safety analysis and trials contributing to the integrated safety analysis (adapted from Table 72 of the CS, Table 81 of the JAIY CSR, Table 63 of the JAHL CSR and Table 63 of the JAHM CSR)

	BREEZE-AD4 (JAIN)		BREEZE-AD7 (JAIY)		BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)		Integrated* Analysis	
	PBO+ TCS (N=93)	BARI 4mg + TCS (N=92)	PBO + TCS ([redacted])	BARI 4mg + TCS ([redacted])	PBO (N=249)	BARI 4 mg (N=125)	PBO (N=244)	BARI 4mg (N=123)	PBO ± TCS ([redacted])	BARI 4 mg ± TCS ([redacted]) †
Patients with ≥ 1 TEAE (%)	50 (53.8)	69 (75.0)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SAEs (%)	2 (2.2)	6 (6.5)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs leading to permanent discontinuation (%)	1 (1.1)	1 (1.1)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AESIs (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

***** Integrated analysis consists of patients from the phase II JAHG trial, monotherapy trials JAHL and JAHM and combination therapy study JAIY

† The ERG corrected this from Table 72 of the CS, where N was reported to be 93, using the N reported in other AE tables

TCS : topical corticosteroids; PBO: placebo; BARI: baricitinib; AE: adverse events; AESI: adverse events of special interest; SAE: serious adverse events; TEAE: treatment-emergent adverse events

ERG Comments

It was unclear to the ERG why the company chose to look at the AEs observed in monotherapy trials and combination therapy trials collectively for the integrated safety analysis and further why the Phase II trial was included in the safety analysis.

In their submission, the company only assessed safety data for patients who received either placebo or 4 mg baricitinib. The company did not present safety outcomes for patients who received 1 mg or 2 mg baricitinib, in the CS. As patients in these groups do not receive a higher dose of baricitinib than is

recommended, the ERG considers these data useful for safety analysis, particularly as the company proposes a 2 mg dose for patients aged over 75 years. The Company subsequently provided these data.

Similar to the baricitinib trials, in the CAFÉ and CHRONOS trials the most common adverse events with dupilumab treatment were also infections and infestations (45.8% and 57%, respectively), particularly nasopharyngitis (20.6% and 23%, respectively). However, eye disorders (19.6% and 31%, respectively), particularly conjunctivitis (11.2% and 14%, respectively) were observed as adverse effects with dupilumab and not baricitinib.

3.2.5.6 Patients rescued and rescue therapies

The number of patients who were rescued during the trial are summarised in Table 77 of the CS, and the rescue medicines administered are reported in Table 78.

At week 16, fewer patients in the baricitinib 4 mg treatment arm needed to be rescued compared to placebo in the JAIY, JAHL and JAHM trials. However, in the JAIN trial, the same number of people were rescued in both the placebo and 4 mg baricitinib arms (n = [redacted] and n=[redacted], respectively). At 24 weeks, more patients were rescued in the baricitinib 4 mg arm (n = [redacted] compared to the placebo arm (n = [redacted]. Table 29 presents the cumulative number of patients who were rescued at 16 and 24 weeks in all four baricitinib trials. An integral part of management of AD is the control of flares, as AD is episodic in nature. Flares are typically treated using high potency TCS and sometimes systemic agents. Clinical advice to the ERG and the NES both state that patients are reluctant to use topical corticosteroids on a routine basis to control their symptoms because of concerns about adverse effects. Reducing flares and TCS use is a priority to patients and clinicians. Although flare is not an outcome presented in the submission, receipt of rescue can be considered a proxy for flare. The higher proportion of patients rescued in the JAIN trial in the baricitinib arm relative to the placebo arm suggests that baricitinib treatment is not effective at reducing flares.

Table 29 Cumulative number of patients rescued during the BREEZE-AD trials at 16 and 24 weeks (adapted from Table 77 of the CS)

Cumulative Number of patients rescued, n (%)								
	BREEZE-AD4 (JAIN)		BREEZE-AD7 (JAIY)		BREEZE-AD1 (JAHL)		BREEZE-AD2 (JAHM)	
Time-point	PBO +TCS (N=[redacted])	BARI 4mg +TCS (N=[redacted])	PBO +TCS (N=[redacted])	BARI 4mg +TCS (N=[redacted])	PBO (N=[redacted])	BARI 4mg (N=[redacted])	PBO (N=[redacted])	BARI 4mg (N=[redacted])
16	[redacted]	[redacted]	[redacted]	[redacted]	166 (66.7)	51 (40.8)	187 (76.6)	72 (58.5)
24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

PBO: placebo; BARI: baricitinib; TCS: topical corticosteroids; NA: not applicable

Most patients who had to be rescued were treated with topical corticosteroids. Table 30 summarises the rescue medications used in the JAIN and JAIY trials. In the JAIN trial, [redacted] of patients in the baricitinib 4 mg arm were rescued with topical corticosteroids compared to [redacted] in the placebo arm. In JAIY, [redacted] of patients in 4 mg baricitinib were rescued with topical corticosteroids compared to [redacted] in the placebo arm. The proportion of patients who had to be rescued with topical corticosteroids was higher in the JAHL and JAHM trials. In JAHL, [redacted] of patients in 4 mg baricitinib were rescued with topical corticosteroids ([redacted] in placebo) and [redacted] of 4 mg baricitinib patients in JAHM ([redacted] in placebo).

Table 30 Summary of rescue medications used in the JAIN and JAIY trials (adapted from Table 78 of the CS)

	BREEZE- AD4 (JAIN)		BREEZE-AD7 (JAIY)	
Rescue Medications, n (%)	**PBO + TCS ([redacted])**	**BARI 4 mg +TCS ([redacted])**	**PBO + TCS ([redacted])**	**BARI 4 mg +TCS ([redacted])**
Any rescue	[redacted]	[redacted]	[redacted]	[redacted]
Rescue TCS	[redacted]	[redacted]	[redacted]	[redacted]
Phototherapy	[redacted]	[redacted]	[redacted]	[redacted]
Systemic Medication	[redacted]	[redacted]	[redacted]	[redacted]
Corticosteroids	[redacted]	[redacted]	[redacted]	[redacted]
Ciclosporin	[redacted]	[redacted]	[redacted]	[redacted]
Biologics	[redacted]	[redacted]	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]

The percentages reported for JAIN were adjusted to present the proportion of patients in the treatment arm who received the rescue therapy.

TCS: topical corticosteroid, PBO: placebo; BARI: baricitinib; NA: not applicable

3.2.5.7 Long-term study JAHN (BREEZE-AD3)

The company did not report safety data for the on-going long-term JAHN study, although it provides safety results for the longest-term data. Patients in this study were originally participants in JAHL and JAHM and were assessed for up to an additional 52 weeks after the end of their original trial.

Safety results for JAHN are summarised in Table 31. The proportion of patients who experienced at least 1 TEAE was [redacted] in patients receiving 4 mg baricitinib ([redacted] compared to 2 mg baricitinib [redacted]), 1 mg baricitinib ([redacted] and placebo ([redacted]). While [redacted] patients in the placebo and 1 mg baricitinib arms experienced AEs resulting in permanent discontinuation, [redacted] of the patients in the 2 mg baricitinib arm and [redacted] of patients in the 4 mg baricitinib arm discontinued due to AEs.

[redacted]

Table 31 Overview of Adverse Events for JAHN (BREEZE-AD3) (adapted from Table 159 of CSR)

	PBO (N=[redacted])	BARI-1 mg (N=[redacted])	BARI- 2 mg (N=[redacted])	BARI-4 mg (N=[redacted])
Patients with ≥ 1 TEAE (%)	[redacted]	[redacted]	[redacted]	[redacted]
Deaths (%)	[redacted]	[redacted]	[redacted]	[redacted]
SAEs (%)	[redacted]	[redacted]	[redacted]	[redacted]
AEs leading to permanent discontinuation (%)	[redacted]	[redacted]	[redacted]	[redacted]

TEAE: treatment-emergent adverse events; SAEs: serious adverse events; AEs: adverse events; PBO: placebo; BARI: baricitinib

ERG comments

The SmPC for baricitinib (Olumiant®)[33] in rheumatoid arthritis (where it is administered as monotherapy or in combination with disease-modifying antirheumatic drugs) outlines additional adverse events not reported in the company's safety analysis as they do not appear to occur frequently over the duration of the BREEZE trials. However, due to the lack of long-term studies of baricitinib for the treatment of AD it is worth noting that these adverse events could potentially also occur in patients with AD, including:

(i) An increase in blood lipid parameters which could potentially have an impact on cardiovascular morbidity and mortality.

(ii) Deep vein thrombosis (DVT) and pulmonary embolism (PE).

[redacted]

(iii) An increase in alanine transaminase (ALT) and aspartate transaminase (AST).

[redacted]

(iv) The risk of malignancies including lymphoma was increased in patients with rheumatoid arthritis, although there is insufficient evidence to assess the incidence of malignancies after baricitinib exposure.

[redacted]

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

The company only identified two relevant comparators to baricitinib: dupilumab and BSC. Therefore, the 40 unique studies identified in the SLR were screened to identify studies investigating the use of baricitinib or dupilumab. Of the 40 studies, there were 12 published studies of dupilumab and 3 published studies of baricitinib, listed in Table 50 of the CS. However, no head to head clinical trials comparing baricitinib and dupilumab were identified. Therefore, these studies were reviewed for the purposes of conducting an ITC to assess the clinical effectiveness of baricitinib versus dupilumab.

The studies were screened against the inclusion criteria described in Table 51 of the CS, which appear appropriate. The inclusion criteria were the same as those for the SLR. However, only phase III RCTs in patients who had experienced failure with, or were intolerant to or had contraindication to ciclosporin, including at least one trial arm of baricitinib (4 mg QD) or dupilumab (300 mg Q2W) were included. The methods of data extraction and quality assessment are specified in section 3.1.3 and 3.1.4. The quality assessment suggests that generally, the risk of bias for the studies included was low.

The ERG did not undertake independent searches to check that all relevant studies were included in the ITC, due to time constraints. However, a comparison of studies included in this STA, with the earlier STA of dupilumab, and recent published network meta-analyses of atopic dermatitis, was undertaken. No relevant trials appear to have been excluded from the ITC.

Table 52 of the CS lists all the trials included and excluded from the ITC. The ITC included a total of 8 studies. Four studies of baricitinib: JAHL, JAHM, JAIN, JAIY; and four studies of dupilumab: LIBERTY AD CAFÉ (CAFÉ), LIBERTY AD CHRONOS (CHRONOS), SOLO1 and SOLO2. Individual patient data (IPD) were available for the baricitinib trials, and therefore data relevant to the eligible population of patients who had experienced failure with, were intolerant to or had a contraindication to ciclosporin (JAIN-like subgroup) were extracted from the JAHL, JAHM and JAIY trials. To maximise sample sizes, data were pooled for the baricitinib monotherapy studies JAHL and JAHM by pooling JAIN-like patients from the JAHL study with patients in the JAHM study's JAIN-like subgroup. The data for the baricitinib plus TCS studies JAIN and JAIY were also pooled by combining the JAIN and JAIY JAIN-like subgroups.

Full trial data were not available for the dupilumab trials. However, post hoc pooled analyses from the dupilumab appraisal (TA534),[20] which included the eligible population of patients who have experienced failure with, are intolerant to or have a contraindication to ciclosporin (CAFÉ-like subgroup), were used. The data for the CAFÉ-like subgroups of the dupilumab monotherapy studies SOLO1 and SOLO2 were pooled. The data for the dupilumab plus TCS studies CAFÉ and CHRONOS were also pooled by combining the CAFÉ and CHRONOS CAFÉ-like subgroup data.

All trials reported efficacy endpoints at week 16. A composite outcome of EASI50 response and ≥ 4-point improvement in DLQI was used as the base case, which was available for all trials and pooled groups except for the CAFÉ trial alone. EASI50 and EASI75 outcomes were available from all populations, whereas EASI90 was only available from the CAFÉ trial. All analyses and outcomes considered in the ITC are summarised in Table 32.

Table 32 Summary of analyses in the ITC (from CS, Table 55)

Comparison	Populations		Outcomes (Week 16)
	Baricitinib	Dupilumab	
Baricitinib + TCS versus dupilumab + TCS	**JAIN** All trial data	**CAFÉ** All trial data	• EASI50 • EASI75 • EASI90 • Itch NRS ≥4-point Improvement
	JAIN + JAIN-like JAIY JAIN trial data combined with post hoc data from the subgroup of patients with ciclosporin failure, intolerance or contradiction from the JAIY study	**CAFÉ + CAFÉ-like CHRONOS** CAFÉ trial data combined with post hoc data from the subgroup of patients with ciclosporin failure, intolerance or contradiction from the CHRONOS study	• EASI50 + DLQI ≥4-point Improvement • EASI50 • EASI75
Baricitinib monotherapy versus dupilumab monotherapy	**JAIN-like JAHL + JAHM** Pooled post hoc data from the subgroup of patients with ciclosporin failure, intolerance or contradiction from the JAHL and JAHM studies	**CAFÉ-like SOLO1 and SOLO2** Pooled data for the sub-population of patients with ciclosporin failure, intolerance or contradiction from the SOLO1 and SOLO2 studies	• EASI50 + DLQI ≥4-point Improvement • EASI50 • EASI75

The CS presents the baseline characteristics of the studies included in the ITC in Table 56. In response to the PFC, the company provided baseline characteristics for the pooled data broken down into its component trials: JAIY JAIN-like group, JAHL JAIN-like group and JAHM JAIN-like group. The company also provided baseline characteristics for the pooled populations JAIN + JAIY JAIN-like group and the JAHL/JAHM JAIN-like group (Table 33). However, details of patients' previous use of therapies, including systemic therapy, phototherapy, TCS, TCI and biologic therapy were not provided. The company also did not provide separate baseline characteristics for the CHRONOS CAFÉ-like, SOLO1 CAFÉ-like or SOLO2 CAFÉ-like populations, since they are not publicly available. The baseline characteristics reported were similar across arms in each of the pooled populations.

However, there were some notable differences in patient characteristics across trials, which are discussed on page 113 of the CS. There was a substantial difference in the proportion of Asian patients between the JAIN ([redacted]%) and CAFÉ (2%) studies. In the JAIN + JAIN-like JAIY population, [redacted]% of patients were Japanese and [redacted]% of patients were non-European. The proportion of Japanese or non-European patients in the CAFÉ + CHRONOS CAFÉ-like population was not

reported, although there were some non-European sites in the CHRONOS trial. As discussed in Section 3.2.4.3, significant interactions were observed in the JAIY trial for specific region (Japan vs all others, and East Asia vs all others) for the EASI75 outcome. Therefore, geographic region may be an effect modifier and therefore this is a source of inconsistency between these trials. There was also a difference in the baseline severity of the patients included in the trials. The eligibility criteria for the baricitinib trials included was an EASI score of ≥ 16, whereas for the CAFÉ and CHRONOS trials it was an EASI score of ≥ 20. This indicates that patients in the CAFÉ and CHRONOS trials are likely to be more severe than those included in the baricitinib trials. This is also reflected in the baseline EASI scores, which are slightly higher for the CAFÉ + CHRONOS CAFÉ-like pooled group (34.8 in the placebo arm and 33.6 in the dupilumab arm), compared with the JAIN + JAIY JAIN-like pooled group ([redacted] in the placebo + TCS arm and [redacted] in the baricitinib + TCS arm).

There were some additional differences in trial design and analysis between the baricitinib and dupilumab trials. The baricitinib trials had a washout period of 5 half-lives for biologic AD treatments, 4 weeks for systemic treatments and 2 weeks for topical treatments (including TCS), excluding emollients. Whereas, in the dupilumab trials there was a 2-week TCS standardisation period before randomisation at baseline, during which patients applied medium-potency TCS once daily to active lesion areas or low-potency TCS on areas of thin skin. The differences in length of washout period and in TCS use before randomisation may indicate that patients in the dupilumab trials are less likely to experience a flare and more likely to have a better response than patients in the baricitinib trials.

Both the baricitinib trials and the dupilumab trials applied different censoring rules. Primary censoring rules were the same in both trials. However, secondary censoring in the dupilumab trials was different to the baricitinib trials, as all observed data regardless of rescue treatment was used, including data collected after withdrawal (referred to as 'all patients' analysis in TA534[20]). In the baricitinib trials, data were still considered as missing or had non-responder imputation after permanent study drug discontinuation or after initiation of systemic rescue therapies, but not considered missing after rescue with TCS. Therefore, in the dupilumab trials there may be a higher response rate using secondary censoring as data from additional rescued patients are included. The differences in relative treatment effects using both primary and secondary censoring are discussed in Section 3.4.2.

The differences described increase the risk of between study, across-comparison heterogeneity, which reduces the reliability of the ITC results. The ERG recognises that there are no more than 2 studies within each comparison so statistical assessment of heterogeneity was not possible.

Table 33 Baseline characteristics for the pooled populations considered in the ITC (from Table 7 of the company clarification response)

Intervention	JAIN + JAIY JAIN-like pooled		JAHL/JAHM JAIN-like pooled	
	PBO +TCS	BARI 4 mg QD +TCS	PBO	BARI 4 mg QD
N	[redacted]	[redacted]	[redacted]	[redacted]
Male, %	[redacted]	[redacted]	[redacted]	[redacted]
Race, (%)				
White	[redacted]	[redacted]	[redacted]	[redacted]
Asian	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]
Age (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Baseline scores, mean (SD)				
EASI	[redacted]	[redacted]	[redacted]	[redacted]
SCORAD	[redacted]	[redacted]	[redacted]	[redacted]
IGA	[redacted]	[redacted]	[redacted]	[redacted]
DLQI	[redacted]	[redacted]	[redacted]	[redacted]
Itch NRS	[redacted]	[redacted]	[redacted]	[redacted]
BSA affected	[redacted]	[redacted]	[redacted]	[redacted]
POEM	[redacted]	[redacted]	[redacted]	[redacted]
HADS[a]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D VAS	[redacted]	[redacted]	[redacted]	[redacted]

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

3.4.1 Critique of the indirect comparison methods

The ITC results presented a composite outcome of EASI50 + ≥ 4-point improvement in DLQI as the base-case, which is consistent with the previous NICE STA submission for dupilumab (TA534).[20] However, this outcome was not available for the CAFÉ trial and so could only be reported for the pooled comparisons. The company stated that the composite endpoint was used as the base-case due to what clinicians considered to be clinically meaningful changes in outcomes, while the CAFÉ and CHRONOS trial endpoints were dictated by the requirements of regulatory agencies. The clinical experts to the Committee in TA534[20] explained that EASI75 and IGA 0/1, the endpoints of the clinical trials, are difficult to achieve in practice, and that the composite endpoint was more sensitive to changes in treatment outcomes and more clinically relevant. However, the British Association of Dermatologists state that a clinically significant improvement is defined as a reduction in EASI score of 75% (i.e. EASI75), or a fall in IGA of 2 points.

The ITC also reported EASI response rates (EASI50, EASI75 and EASI90). However, Itch NRS ≥ 4-point improvement was only reported for the JAIN vs CAFÉ comparison not the pooled comparisons. Indirect treatment comparisons were carried out using the Bucher method[34] to compare baricitinib with dupilumab, via the placebo common comparator. This is a frequentist method, which takes the relative effect estimated for one treatment vs placebo and subtracts it from the treatment effect of the other treatment vs placebo, to obtain an indirectly estimated relative effect (and variance) of the two active treatments. Relative effects and variances of each treatment vs placebo, used in the equations of the Bucher method, came from single studies or, when more than one study was available, were obtained through meta-analysis. For binary outcomes, the Mantel-Haenszel method was used. There was no ITC of continuous outcomes. Binary outcomes were assessed as odds ratio (OR), relative risk (RR) and risk difference (RD). The ERG will comment mainly on the OR results for conclusions on clinical effectiveness. Separate ITC were carried out for each of the outcomes EASI50, EASI75 and EASI90. No joint analysis of EASI cut points was considered as the company stated that from a medical point of view, EASI75 is considered to be the most important EASI outcome (since it is the most sensitive in clinical practice). There were also few cases of EASI90, which may compromise the results if combined with other EASI measures. Therefore, only standalone results were presented. In addition, the company stated that as the mean change from baseline in absolute EASI score is available from the BREEZE-AD trials, there is no additional value in conducting an analysis where categorical EASI measures are combined.

Assumptions of indirect comparison methods are that all the studies included both in the within-comparison (i.e. pairwise meta-analysis) and between-comparison syntheses (Bucher indirect comparison) are sufficiently homogeneous to allow meaningful pooling. Lack of homogeneity of the relative effects within comparisons is usually termed heterogeneity, and lack of homogeneity across comparisons is termed inconsistency. The company planned to assess statistical heterogeneity between the studies on each direct treatment comparison by identifying and quantifying Tau^2 (the DerSimonian-Laird approach) and I^2 and to use a fixed effect (FE) model to obtain the pooled estimator of the corresponding treatment effect if no between-study heterogeneity was identified. However, the hypothesis of within-comparison homogeneity cannot be reliably tested when there are less than 3 studies in each meta-analysis as there is insufficient information to estimate the between-study heterogeneity. The assumption of consistency cannot be tested unless direct evidence on the relative efficacy of the two active treatments was available. In the ITC presented in the CS, the maximum number of studies in each meta-analysis was two (usually just one) and there was no direct evidence comparing baricitinib to dupilumab. Therefore, none of these assumptions can be statistically evaluated.

However, the expectation of statistical homogeneity can be validated clinically by comparison of study conditions, patient characteristics, and outcome measures. The trials included in the ITC comparisons vary by ethnicity, baseline severity and other features, as discussed in Section 3.3. These variations contribute to differences in placebo response rates which can be an indicator of potential differences in the relative efficacy of the interventions compared with placebo. In response to the PFC, the company stated that a fixed-effect model was chosen for all comparisons as there were too few studies included to produce reliable between-study variations for random-effects models. Therefore, the ERG considers the analyses based on FE meta-analysis models acceptable but they need to be interpreted with caution due to the possibility of imbalance in effect modifying covariates across studies.

Analyses were conducted in R with meta-analyses conducted using the package *meta*[35] and the Bucher ITC using Lilly's own Cheetah-tool. The company provided raw data tables and the R code to run the ITC. The ERG has checked and validated the code and results obtained.

3.4.2 Indirect comparison results

Results presented in the ITC are based on data where the primary censoring rule was applied. However, results using secondary censoring were also presented for the base-case JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients. The ERG notes that results from secondary censoring are more likely to reflect clinical practice, as noted in Section 3.2.4.1.

EASI50 and DLQI ≥ 4-point improvement

The relative treatment effects for EASI50 and DLQI ≥ 4-point improvement in the base-case population: JAIN and JAIN-like JAIY patients and CAFÉ and CAFÉ-like CHRONOS patients are presented in Table 34. The results show that dupilumab is more effective than baricitinib in achieving EASI50 + ≥4-point improvement in DLQI using both primary censoring (OR: [redacted]) and secondary censoring (OR: [redacted]) at week 16.

Table 34 Relative treatment effects for EASI50 + DLQI ≥4-point improvement at week 16 for JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients (adapted from Table 57 of the CS and the company clarification response)

Source	n/N (%) Placebo	n/N (%) Active Treatment	OR†	95% CI
Primary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
Pooled: JAIN+ JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	NA/108 (NA)	NA/107 (NA)	NA	
CAFÉ-like CHRONOS	NA/61 (NA)	NA/23 (NA)	NA	
Pooled: CAFÉ + CAFÉ-like CHRONOS	35/169	89/130	[redacted]	[redacted]
ITC: Pooled JAIN+ JAIN-like JAIY vs. Pooled CAFÉ + CAFÉ-like CHRONOS (fixed-effects model)			[redacted]	[redacted]
Secondary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
Pooled: JAIN+ JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	NA/108 (NA)	NA/107 (NA)	NA	NA
CAFÉ-like CHRONOS	NA/61 (NA)	NA/23 (NA)	NA	NA
Pooled: CAFÉ + CAFÉ-like CHRONOS	47/169	95/130	[redacted]	[redacted]
ITC: Pooled JAIN+ JAIN-like JAIY vs. Pooled CAFÉ + CAFÉ-like CHRONOS (fixed-effects model)			[redacted]	[redacted]

† ORs presented for fixed effects meta-analyses

OR: odds ratio; CI: confidence interval; BARI: baricitinib; DUPI: dupilumab; PBO: placebo; TCS: topical corticosteroids; NA: not available

Other outcomes

There was a statistically significant higher odds of achieving EASI50 with dupilumab compared to baricitinib for the analyses using the JAIN + JAIN-like JAIY and CAFÉ + CAFÉ-like CHRONOS patients (OR: [redacted]) and for the analyses using only the JAIN and CAFÉ populations (OR: [redacted]) using primary censoring. Secondary censoring results were similar for the analyses using the JAIN + JAIN-like JAIY and CAFÉ + CAFÉ-like CHRONOS patients but not presented for the JAIN and CAFÉ populations (Table 67, Appendix 2).

Dupilumab also showed higher odds of achieving EASI75 than baricitinib but there was no statistically significant difference for both the JAIN + JAIN-like JAIY and CAFÉ + CAFÉ-like CHRONOS populations (OR: [redacted]) or the JAIN vs CAFÉ populations (OR: [redacted]) using primary censoring. The results using secondary censoring were similar for

the analyses using the JAIN + JAIN-like JAIY and CAFÉ + CAFÉ-like CHRONOS patients but not presented for the JAIN and CAFÉ populations (Table 68, Appendix 2).

The EASI90 results for the JAIN vs CAFÉ population using primary censoring also favoured dupilumab, however they were not statistically significant (OR: [redacted]) (Table 69, Appendix 2). Results for secondary censoring were not presented.

Baricitinib showed a higher probability of achieving Itch NRS ≥ 4-point improvement compared to dupilumab for the JAIN vs CAFÉ population but the difference was not statistically significant and results are very uncertain (OR: [redacted]) (Table 70, Appendix 2).

In response to the PFC, the company stated that the results for EASI50 and EASI75 for the JAIN vs CAFÉ comparison were comparable to the results obtained for the analyses using JAIN + JAIN-like JAIY and CAFÉ + CAFÉ-like CHRONOS. The reason for conducting both comparisons was mainly due to the availability of outcomes: the composite outcome of EASI50 with DLQI ≥ 4-point improvement was not available for the CAFÉ trial alone, and EASI90 data were not available for the pooled CAFÉ + CAFÉ-like CHRONOS population.

Sensitivity analysis

A sensitivity analysis was conducted comparing JAIN and CAFÉ, where only European patients from JAIN were included. The CS stated that the rationale for this analysis was due to the significant interactions observed for specific region (East Asia vs all others) in the JAIY trial for the Itch NRS outcome, indicating that region may be a treatment effect modifier. In response to the PFC, the company provided baseline characteristics of European patients only from the JAIN trial. There were slight differences between the placebo arm, and the 4 mg baricitinib arm, in the proportion that were female ([redacted]% vs [redacted]%, respectively), the duration since AD diagnosis ([redacted] years vs [redacted] years, respectively) and BSA ([redacted] vs [redacted], respectively). The baseline characteristics were mostly similar to the patients in the CAFÉ trial.

There were no significant differences observed for this European-only population between dupilumab and baricitinib for EASI50 (Table 67, Appendix 2), EASI75 (Table 68, Appendix 2), EASI90 (Table 69, Appendix 2) or for Itch NRS (Table 70, Appendix 2). However, the comparisons between baricitinib and placebo for the JAIN European-only patients showed a better response than the full JAIN population for each outcome. This indicates that European patients may have a better response with baricitinib than non-European patients. As discussed earlier, this could be due to differences in clinical practice relating to rescue treatment. However, the true reasons are unclear and this remains an area of uncertainty.

Limitations of the ITC

For EASI50 and ≥ 4-point improvement in DLQI, EASI50 alone and EASI75, baricitinib had a greater relative effect than dupilumab relative to placebo using secondary censoring rather than primary censoring. In the JAIN trial, a higher proportion of patients in the baricitinib arm used rescue TCS than in the placebo arm. Whereas, in the CAFÉ trial, a higher proportion of patients used rescue TCS in the placebo arm than the dupilumab arm. Secondary censoring did not consider data missing after rescue with TCS for the baricitinib trials, which may explain why baricitinib has a better odds of response than dupilumab using secondary censoring compared with primary censoring. As noted above, the ERG considers secondary censoring to better reflect what would happen in clinical practice.

Patient reported outcomes, such as skin pain NRS and ADSS that were presented in the JAIN and JAIY trials, were not included in the ITC due to them not being available from the CAFÉ and CHRONOS trials. Itch NRS ≥ 4-point improvement was not available for the pooled populations. Clinical advice to the ERG is that these patient-reported outcomes are very important due to the effect they have on the quality of life of patients, particularly itch and scratching as it is correlated with flares, reduced performance or ability to concentrate on a task and lack of sleep in patients with AD. Additionally, the NES stated that itchiness, skin pain and sleep disturbance are the most debilitating symptoms of AD, with constant itchiness being one of the most challenging aspects of eczema.

A comparison based on IGA could not be conducted because the dupilumab and baricitinib trials used different IGA scales. In response to the PFC, the company stated that mean IGA scores are not available for dupilumab from CAFÉ or the CAFÉ + CHRONOS CAFÉ-like and SOLO1/2 CAFÉ-like pooled populations; only the proportion of patients achieving IGA of 0 or 1 is reported. As such, it was not possible to conduct an indirect comparison using standardised mean differences in IGA. The ERG also notes that in the CHRONOS trial, dupilumab reported greater flare suppression when compared with placebo (16% vs 52% respectively), which significantly reduced the need for rescue therapy in patients treated with dupilumab. Whereas, as discussed in Section 3.2.5.6, baricitinib had a higher rate of rescue therapy at week 24 in the JAIN trial compared with placebo. This indicates that relative to dupilumab, baricitinib may not be as effective at controlling flares and reducing the use of high potency TCS.

The ITC was only available for outcomes at 16 weeks, therefore, the long-term efficacy of baricitinib compared with dupilumab is uncertain. There was no ITC for adverse events carried out, however this could have been done for adverse events that were reported by the CAFÉ trial, including ≥ 1 TEAE, ≥ 1 TE SAE, death, diarrhoea, abdominal pain, back pain and asthma.

3.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG carried out checks of all data and code used for the ITC. Additional analyses were carried out to obtain absolute probabilities of response to be used in the ERG economic analyses. These are described in Section 6.2

3.6 Conclusions of the clinical effectiveness section

The clinical evidence presented in the submission is based on four multicentre RCTs (JAIN, JAIY, JAHL and JAHM). JAIN and JAIY are the most relevant trials as they compare baricitinib in combination with TCS to placebo and baricitinib is most likely to be given alongside TCS in practice. Whereas JAHL and JAHM compare baricitinib monotherapy with placebo. An ITC was conducted to compare baricitinib with dupilumab, as there was no head to head evidence directly comparing both treatments.

All four baricitinib trials were reasonably good quality and the results are likely to be reliable. The population considered in the submission was adult patients who are candidates for systemic therapy and who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This is different to the NICE scope, which states the population is adults with moderate-to-severe atopic dermatitis (AD) who are candidates for systemic therapy that had an inadequate response or intolerance to existing topical treatments. Clinical advice to the ERG is that baricitinib is likely to be used at the same point in the treatment pathway as other immunosuppressants, prior to dupilumab. Therefore, the ERG considers the base-case trial population to be restrictive and not fully representative of the population for this indication. Additionally, clinical advice is that currently dupilumab is favoured above a "second" systemic and thus started early in the pathway due to its low side effect profile and less frequent monitoring requirements. Consequently, the comparators to baricitinib should not be restricted to dupilumab but should also include systemic immunosuppressants (Table 4). Trial inclusion criteria appear to have been appropriate. However, the mean age of patients in all four trials was [redacted] years old, which is higher than would be expected in the NHS population. The ERG also notes that the inclusion criteria for the baricitinib trials presented in the submission specified an EASI score ≥ 16, IGA score ≥ 3 and BSA involvement ≥ 10%. Several published strata on EASI score, state that moderate disease is associated with EASI scores as low as 6 ranging to 22.9[1]. This includes patients with EASI scores far below the trial inclusion criteria of 16, which may exclude patients on the lower end of the moderate scale and bias the trial populations towards more severe disease. Furthermore, the published strata on EASI score indicate that the mean EASI scores ([redacted]) in the trial populations represent severe disease (see Section 3.2.2). Therefore, in terms of age and disease severity, the ERG considers the population in the clinical evidence presented may not represent all moderate to severe patients in the NHS population.

In the JAIN trial, patients treated with 4 mg baricitinib were more likely to achieve EASI50 (OR: [redacted]) and EASI75 (OR: [redacted]) compared to placebo at 16 weeks using the primary censoring rule. However, there was no statistically significant difference at week 24. In the JAIY trial, patients treated with 4 mg baricitinib were more likely to achieve EASI50 (OR: [redacted]), EASI75 (OR: [redacted]) and EASI90 (OR: [redacted]) at 16 weeks. In JAIN, the 4 mg baricitinib group, patients were more likely to achieve a ≥ 4-point improvement in NRS Itch scores compared to patients treated with placebo at week 16 ([redacted]) and week 24 (OR: [redacted]) using primary censoring. In JAIY, the difference in patients achieving ≥ 4-point improvement in Itch NRS was statistically significant at week 16 (OR: [redacted]). Using primary censoring, patients in the 4 mg baricitinib arm were more likely to achieve a ≥ 4-point improvement in DLQI compared to those in placebo at week 16 in JAIN (OR: [redacted]) and JAIY (OR: [redacted]). The results using secondary censoring were consistent with the results reported here (see Section 3.2.4.2).

The company did not present subgroup analyses on skin colour or on severity, although these were pre-planned subgroups and specified in the NICE scope. The company notes differences in disease pathology and efficacy in AD across ethnic groups and the ERG notes that AD can be more severe and persistent in Black patients. [redacted] Black patients were recruited in the baricitinib trials, therefore the ERG notes that this is a potential equalities issue as the lack of data on Black patients means that it is not possible to establish baricitinib efficacy in this population (Table 4). The company did not provide subgroup analyses by severity as it stated that EASI does not reflect all aspects of moderate or severe disease and it does not provide consistent classification of disease severity. However, clinical advice to the ERG is that EASI is widely accepted and considers all relevant aspects of AD clinical signs. The ERG considers that although there are limitations to using one severity classification, it would have been plausible and beneficial to present separate subgroups of moderate and severe AD based on published EASI strata (Table 4).

In the JAIY trial, and the JAIN + JAIN-like JAIY population, a higher proportion of European patients achieved EASI75 with baricitinib than placebo compared with Japanese patients. In the JAIY trial, a higher proportion of European patients had ≥ 4-point improvement in Itch NRS than Non-European patients. The greater response seen in European patients may be due to differences in clinical practice, particularly the use of rescue TCS and the baseline severity between European and Japanese patients (see Section 3.2.4.3). Notably, the results of regional subgroup analyses do not support a lower response for baricitinib 4 mg in East Asian countries not including Japan, suggesting there is not a specific effect of East Asian ethnicity but rather other characteristics of the recruited Japanese patients. However, this is a source of uncertainty, indicating that the trial populations are not fully generalisable to the NHS population, which should be considered when interpreting the results.

The ITC appears to have included all relevant trials of baricitinib and dupilumab. Studies were assessed for quality, which suggested generally, the risk of bias for most studies was low. Inevitably the trials included in the ITC vary by design and patient characteristics. There was a substantial difference in the proportion of Asian patients between the JAIN ([redacted]%) and CAFÉ (2%) studies. As discussed, geographic region may be an effect modifier, therefore this is a source of inconsistency between these trials. There were also differences in the length of the washout periods and in TCS use before randomisation between the trials. Additionally, differences in the secondary censoring rule indicate that the dupilumab trials may have a higher response rate using secondary censoring as data from additional rescued patients are included. These differences increase the risk of between study, across-comparison heterogeneity, which reduces the reliability of the ITC results.

The ITC results using primary censoring report that dupilumab is more effective than baricitinib in achieving EASI50 + ≥4-point improvement in DLQI (OR: [redacted]) and EASI50 (OR: [redacted]) at week 16 using primary censoring, which were similar using secondary censoring. For the full JAIN vs CAFÉ populations, there was a [redacted] of achieving EASI50 with dupilumab than baricitinib (OR: [redacted]). However, there is considerable uncertainty in most ITC results due to wide confidence intervals. Patient reported outcomes, such as skin pain NRS and ADSS were not included in the ITC due to them not being available from the CAFÉ and CHRONOS trials. Itch NRS ≥ 4-point improvement was not available for the base case population. Clinical advice to the ERG is that these patient-reported outcomes are very important due to the effect they have on the quality of life of patients, particularly itch as it is correlated with flares and lack of sleep in patients with AD. Additionally, the ITC was only available for outcomes at 16 weeks, therefore, the long-term efficacy of baricitinib compared with dupilumab is uncertain. There was also no ITC for adverse events carried out, however this could have been done for adverse events that were reported by the CAFÉ trial, including ≥ 1 TEAE, ≥ 1 TE SAE, death, diarrhoea, abdominal pain, back pain and asthma.

In JAIN, at 16 weeks, a higher proportion of patients in the 4 mg baricitinib group (75.0%) experienced at least one TEAE compared to the placebo group (53.8%). Between 16 weeks and 24 weeks, an additional [redacted] TEAEs were experienced in the 4 mg baricitinib group. In the integrated analysis, [redacted] experienced at least 1 TEAE in both groups. In both JAIN and the integrated analysis, the most common AESIs were treatment emergent (TE) infections ([redacted]% and [redacted]%, respectively), and in particular herpes simplex ([redacted]% and [redacted]%, respectively). In the CAFÉ and CHRONOS trials the most common adverse events with dupilumab treatment were also infections and infestations (45.8% and 57%, respectively), particularly nasopharyngitis (20.6% and 23%, respectively). However, eye disorders (19.6% and 31%, respectively) were also observed as

adverse effects with dupilumab. No deaths were observed in any of the trials for the 16-week duration, or up to 24 weeks for JAIN.

An integral part of management of AD is the control of flares, as AD is episodic in nature. Flares are typically treated using high potency TCS. Reducing flares and TCS use is a priority to patients and clinicians due to the adverse effects of using TCS. Although flare is not an outcome presented in the submission, receipt of rescue can be considered a proxy for flare. In the JAIN trial, a similar number of people were rescued in both the placebo and 4 mg baricitinib arms (n = [redacted] and [redacted], respectively) and more patients were rescued in the baricitinib 4 mg arm (n = [redacted]) compared to the placebo arm (n = [redacted]) at week 24. The ERG also notes that in the CHRONOS trial, dupilumab reported greater flare suppression when compared with placebo (16% vs 52% respectively), which significantly reduced the need for rescue therapy in patients treated with dupilumab. This indicates that baricitinib treatment is not as effective at controlling flares and reducing the need for high potency TCS.

4 COST EFFECTIVENESS

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

The company performed three systematic literature reviews (SLR) in order to identify relevant economic evaluations, as well as information on resource use, costs and quality of life estimates for adults with moderate-to-severe AD. The details of the SLRs are provided in Appendices G-I of the company submission.

Search strategy

The searches for the three SLRs were undertaken in the following electronic databases: MEDLINE, MEDLINE In Progress, EMBASE, Econlit as well as the Centre for Reviews and Dissemination (CRD) databases including the Database of Abstracts and Reviews of Effects (DARE), NHS Economic Evaluations Database (NHS EED) and the HTA database. Conference abstracts from the American Academy of Dermatology, CEA Registry, ISPOR and the European Academy of Dermatology and Venerology were also searched. The websites of the HTA agencies from the UK, US, Canada, Australia, Sweden, Norway, Germany, France, The Netherlands, Italy, Belgium and Spain were also searched for relevant appraisals of therapies for adults with AD.

The searches were first conducted in March 2018, with an updated search taking place in February 2020. The updated search on the websites for HTA agencies was only carried out for English speaking countries (UK, US, Canada and Australia).

Inclusion/exclusion criteria

The inclusion criteria for the SLRs on cost-effectiveness models, health related quality of life (HRQL) estimates, and resource use and costs are presented in Appendix G.1.2, Appendix H.1.2 and Appendix I.1.2 of the CS respectively. In brief, the cost-effectiveness review included studies if they assessed the cost-effectiveness of any treatments for AD. A broad range of studies were considered for inclusion. These included cost-effectiveness, cost-utility, cost-minimisation, cost studies and utility studies. The quality of life and resource use reviews adopted similar criteria with a focus on outcomes relevant to each of these reviews.

Interpretation of the review: Cost-effectiveness review

The SLR identified seventeen studies that met the eligibility criteria (summarised in the CS Appendix G, Table 19). The company describes the structure of their economic model presented as being based on the approach described in two HTA reports: TA534[36] which appraised dupilumab for moderate-to-severe AD, and a US Institute for Clinical and Economic Review report,[37] which considered the clinical and cost-effectiveness of dupilumab for moderate-to-severe AD. The ERG considers both of these evaluations to be highly relevant to the decision problem and useful sources of information. The

ERG, however, notes some important differences between the company's model and that presented in the Institute for Clinical and Economic Review report. Namely that the Institute for Clinical and Economic Review used several additional health states to delineate different categories of response and the use of separate models to consider patients with moderate and severe AD. The ERG considers that it may have been appropriate to adopt features of this model in the company's analysis and it was not clear from the CS why they were rejected, see Section 4.2.1 for further discussion.

An important feature of AD is that patients will experience periods of relative disease control followed by exacerbations in which symptoms flare. Several models identified in the cost-effectiveness review sought to capture the relapsing-remitting nature of AD (including Ellis *et al.*[38, 39] and Pitt *et al.*[40]), but were not considered relevant by the company. The ERG considers this a potentially important omission, see Section 4.2.1 for further details.

Interpretation of the review: Health Related Quality of Life Studies

A SLR was conducted to identify studies with relevant data on health-related quality of life (HRQL). Twenty-three relevant studies containing information on the quality of life of individuals with moderate-to-severe AD were identified (summarised in the CS Appendix H, Table 24). No studies reported HRQL data for individuals using baricitinib for the treatment of moderate-to-severe AD. As HRQL data was collected using the EQ-5D-5L tool in the BREEZE-AD trials, the company instead used this to inform health state utilities in the model.

Interpretation of the review: Cost and Healthcare Resource Identification, Measurement, and Validation

A SLR was conducted to identify studies with relevant data on costs associated with treatment for moderate-to-severe AD. Four studies that included appropriate information on resource use and costs were identified (summarised in the CS Appendix I, Table 29). One study[41] identified in the SLR presents evidence that severity may impact resource use, particularly regarding disease management during flares and response maintenance periods. Although this study[41] provides limited coverage of resource use and is somewhat outdated, the notion that severity may impact on resource use is particularly relevant to the decision problem. Another study, Ameen *et al.*,[42] which was used to inform the resource utilisation and costs in the dupilumab submission (TA534[36]), was also used by the company. Finally, data on resource use and costs accepted by the committee in TA534[36] were used to populate their model.

Conclusions of the economic reviews

The company's cost-effectiveness review did not identify any relevant economic assessments of baricitinib. It did, however, identify several economic evaluations of other therapies for AD, including recent HTAs[20, 36, 37] considering moderate-to-severe AD. The critical appraisal of these studies largely

focused on describing model features rather than a thorough analysis of the various modelling approaches, key assumptions, and data sources. The review, however, provided useful contextual information and allowed the company to identify and justify several assumptions and data sources used in the company model.

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

Table 35 summarises the economic submission and the ERG's assessment of whether the *de novo* evaluation meets NICE's reference case and other methodological recommendations.

Table 35 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 model considered QALY benefits to treated individuals.
Perspective on costs	NHS and PSS	NHS and PSS costs were considered.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Fully incremental cost–utility analysis.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The economic model used a time horizon of 62 years – sufficient to capture important differences.
Synthesis of evidence on health effects	Based on systematic review	Systematic review was conducted for evidence of health effects. Indirect treatment comparison was conducted to combine relevant clinical trial data.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Health effects were presented in QALYs. Measured directly from patients in the trials using EQ-5D-5L and mapped to EQ-5D-3L. Utility gained from response to treatment estimated using a GLM regression model.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Utilities were populated using quality of life data collected from JAIN and JAIY JAIN like patients.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	UK population valuation set used within mapping, described in Dolan *et al.*[43]
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No special weighting undertaken.
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	Costs considered were NHS and PSS. Resource use was primarily taken from TA534[36] using prices relevant to the NHS and PSS.

Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with reference case
Sensitivity analysis	Probabilistic sensitivity analysis	The presented probabilistic analysis was not properly implemented as estimates of uncertainty were arbitrary and did not reflect parameter uncertainty. The revised model received at PFCs addressed this issue.

NHS, national health service; PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome: PFC, points for clarification.

4.2.1 Model structure

The company developed a *de novo* cost-effectiveness model to evaluate the cost-effectiveness of baricitinib versus dupilumab and best supportive care (BSC). The model was built in Microsoft Excel, and allows for both pairwise comparisons and fully incremental analysis. The analysis uses a 62-year (lifetime) time horizon, and was chosen to reflect the chronic nature of AD. In line with the NICE reference case, costs and QALYs are discounted at a rate of 3.5%.

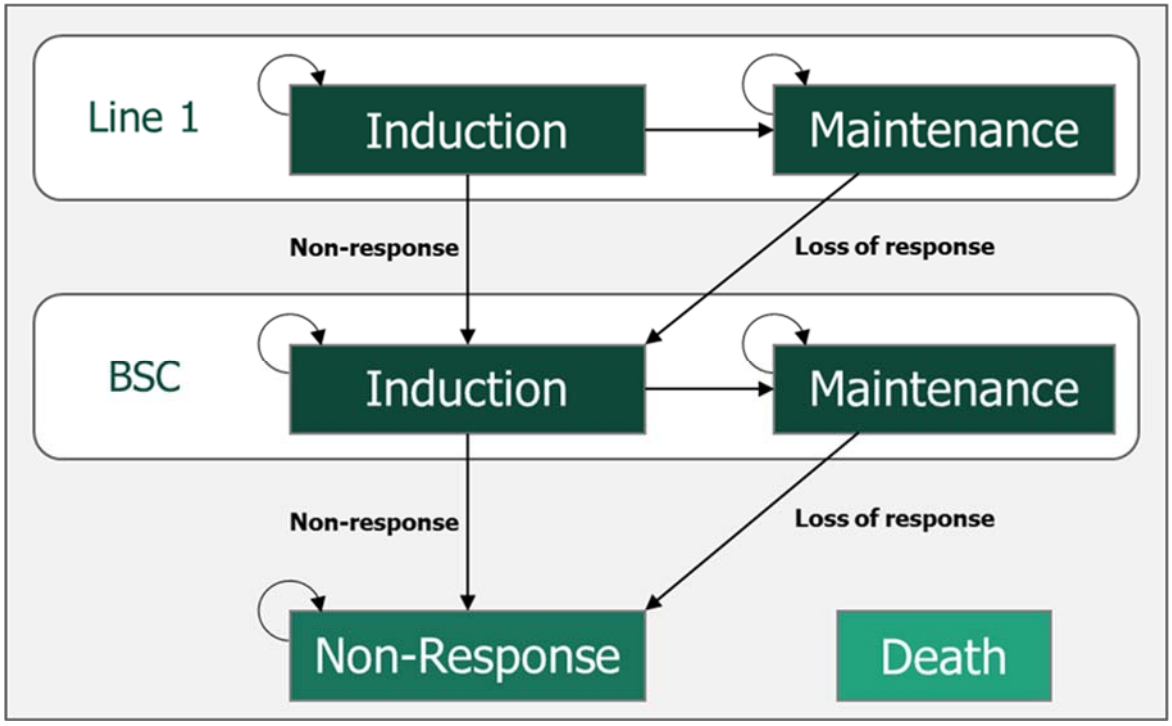
The population included in the model are adult patients with moderate-to-severe AD, who have failed at least one systemic immunosuppressant. The company also submitted scenario analyses whereby only patients recruited to European sites in the JAIN and CAFÉ (dupilumab) trials were analysed.

The Markov model captures the treatment of atopic dermatitis in four distinct health states, 'Induction' (representing a set of tunnel states), 'Maintenance', 'Non-Response' and 'Death'. Each cycle is four weeks in duration. As the cycle length is only four weeks, no half-cycle correction was applied. All model inputs were scaled to the cycle length and for the most part, were linked to response status rather than treatment received. Differences in total costs and utilities are therefore primarily driven by differences in the proportion of patients achieving and maintaining response to treatment.

The structure of the company's model is depicted in Figure 5. Patients enter the model and are allocated to baricitinib or the comparator (either BSC or dupilumab) and enter into the induction stage for that treatment. Patients remain in the induction stage for 16 weeks. During this period, patients cannot discontinue treatment, and can only transition into the 'Death' health state.

After the 16-week induction period, the patient's response to treatment is assessed. In the base case, patient response is defined by a composite of EASI50 and an improvement of four points or more in DLQI score ($\Delta DLQI \geq 4$). This definition of response is in line with the committee's preferred definition of response for TA534.[36] If patients respond, they transition to the 'Maintenance' (responder) health state, while non-responders transition to BSC treatment. The company also present scenarios in which response is defined as EASI50, EASI75, or Itch NRS ≥ 4.

Figure 5. Model structure in Company Submission. (Source: CS, Figure 40).



Patients who transition to the 'Maintenance' health state are modelled to receive continuous treatment until they lose response (Weeks 16-52), or discontinue for other reasons such as adverse events (known as all cause discontinuation) from Year 2 onwards. The annual probability of all-cause discontinuation was obtained from the dupilumab CHRONOS trial, and was assumed to be constant rate of 3.7% per annum (Source: CS, Table 87).

In the dupilumab and baricitinib treatment arms of the model, patients who discontinue treatment, owing to non-response at the end of the induction period, loss of response from Weeks 16-52, or all-cause discontinuation from the beginning of Year 2 onwards, transition to first-line BSC. Upon their transition to first-line BSC, patients enter a second set of induction tunnel states. Patients remain in the BSC induction state for 16 weeks, during which they cannot discontinue and can only transition to the 'death' health state. After the 16-week induction period, patients who respond to the first-line BSC (defined as EASI50 and $\Delta DLQI \geq 4$) transition to the first-line BSC 'maintenance' health state; those who do not respond to first-line BSC transition to 'second-line BSC' ('non-responder BSC'). These two lines of BSC are characterised by different costs (patients responding to first-line BSC have lower costs) and different utilities (patients cannot respond to second-line BSC) and represent distinct treatment options in the model. For clarity, the ERG henceforth refers to these as first- and second-line BSC.

Patients who enter the non-responder heath state and who receive second-line BSC remain there until death or the end of simulation. Death is an absorbing health state; which patients can transition to

from any other health state. It is assumed that there is no treatment effect on mortality, so the transitional probability to death is equal across health states, and is based on the normal UK mortality rates.

ERG Comment

Comparison with other atopic dermatitis cost-effectiveness models

The company reference two HTA reports[36, 37] that were used to inform the inputs of the economic model. Both HTA reports assessed the cost-effectiveness of dupilumab for the treatment of moderate-to-severe AD. Key assumptions used each of these models are outlined in Table 36.

Table 36 Comparison of Model Structures used for Model Inputs in the Company Submission

	ICER Evidence Report: Dupilumab (2017)	NICE TA534: Dupilumab (2018)	NICE ID1622 Baricitinib (2020)
Summary of Model	Markov Model with 3 health states: Usual Care (baseline/no response) Responder (split by EASI response) Death	One-year decision tree. Response and Non-Response feed into Markov Model. Markov model made of three health states: Maintenance treatment BSC Death	Markov Model with 4 health states: Induction Maintenance Non-Response Death Induction/Maintenance for two lines of therapy: first-line treatment and BSC.
Induction Period to Treatment	Patients are assumed to show no response in the first cycle of the model. They can transition to the 'Responder' health state in the following cycle.	The short-term decision tree models the induction period. If patients respond at the 16-week assessment, they continue on the allocated treatment. Patients who respond to dupilumab at 16 weeks remain on the treatment until 1 year.	Induction period is modelled by a set of four tunnel health states, lasting 16 weeks. Patients are unable to discontinue treatment during this period. A 16-week induction period is modelled for all treatments including best-supportive care.
Transition within Model	Patients enter the model in the non-responder health state. Then transition to responder states after first cycle (four months). Could transition back to 'non-response' state.	Two time points in the one-year decision tree where patients can move between response and non-response.	Patient enter the model and enter the induction period of the allocated treatment. After this, responders transition to 'maintenance' health state. Non-responders transition to BSC induction. Non-responders to BSC transition to 'non-response' health state (second-line BSC).
Cycle Length	4 months	1 year	4 weeks
Comparator	Usual Care: emollients	Best supportive care: emollients, low-to-mid potency topical corticosteroids and rescue therapy.	Best supportive care. emollients, low-to-mid potency topical corticosteroids, phototherapy, psychological support and rescue therapy.
Patient Population	Population with moderate-to-severe atopic dermatitis who have failed topical therapy.	Population with moderate-to-severe AD who are contraindicated to, intolerant of,	Population with moderate-to-severe AD who have failed at least one current systemic

	Moderate and severe patients modelled separately (Proportion of AD: 53% severe, 47% moderate).	and had an inadequate response to a systemic immunosuppressant.	immunosuppressant due to intolerance, contraindication or inadequate disease control.
HRQL	Utility values were based on EQ-5D-3L results collected in the LIBERTY dupilumab trial.	Utility values were based on EQ-5D-3L results collected in the LIBERTY dupilumab trial.	Utility values were based on EQ-5D-5L results collected in the BREEZE-AD studies and were converted to EQ-5D-3L utilities.
Resource and Costs	Costs were obtained from the Truven Health Marketscan® Commercial Claims and Encounters database (2013).	Costs were obtained from the BNF (2017), PSSRU and the National Reference Costs (2015) and the National Schedule of Reference Costs (2015-2016), and NHS Reference Costs (2014). Resource use for AEs were based on dupilumab clinical trials	Costs were obtained from the BNF (2019), PSSRU and National Reference Costs (2019) and the National Schedule of NHS Costs (2018-2019). Resource use was based on TA534.

The ERG considers particular aspects of the previous HTA models to better represent AD and its management in clinical practice.

In the Institute for Clinical and Economic Review[37] model for dupilumab, the model's structure allows for a more nuanced approach to assessing treatment-effectiveness, where patients who transition to the 'responder' health state have varying utilities associated with their level of response (EASI50, EASI75 or EASI90). This approach is likely to better reflect the clinical reality, as patients may achieve differing levels of response which are likely to correspond to different improvements in quality of life compared to BSC. Importantly, such an approach would improve model precision and better reflect the benefits of more efficacious treatments Regarding the use of the simpler two state approach in TA534[20], the ERG notes that this was combined with treatment specific utilities, which would mitigate the limitations of this approach.

The model described in the US Institute for Clinical and Economic Review[37] report also stratifies the treatment effect, costs, and utilities by baseline severity of AD, i.e. moderate and severe. Costs and quality of life estimates were calculated based on the proportion of moderate and severe AD patients in each health state at one time. The ERG considers this to be more accurate compared to the model presented in the CS, where there is little stratification by severity of AD. For example, it is likely that patients with severe AD who achieve a response will experience greater improvements in quality of life compared to someone with moderate AD (this can be seen in the Institute for Clinical and Economic Review model[37]). Depending on the proportion of moderate/severe patients who achieve response, the HRQL for responders used in the company's economic model may result in a

under/overestimate of the overall benefit from the treatment, which may lead to an over/underestimate of the ICER, respectively.

Best supportive care modelled as a line of therapy

The company models BSC as a distinct line of therapy where, in principle, response can be achieved and maintained. This is followed by a further line of BSC for those who do not respond to initial BSC treatment (called 'non-response' in the company's model) and referred to by the ERG as 2nd line BSC.

This approach to modelling BSC as a distinct line of therapy, however, mischaracterises BSC, while also ignoring important features of AD; specifically, the waxing and waning nature of symptoms. In the ERG's view, BSC is not a treatment aimed at achieving disease control in the same way as either dupilumab or baricitinib, but rather the treatment of choice to manage disease-related symptoms when disease control cannot be achieved with existing treatment.

The modelling of BSC should therefore reflect the fact that these patients are likely to have lower levels of disease control than patients receiving therapies such as baricitinib and dupilumab. Importantly, the modelling of BSC should also reflect the fact that patients receiving BSC will have periods of both good and poor disease control, rather than a stable but very poor QoL. The company's economic model, however, does not permit this. Instead, owing to the high discontinuation rate assumed in first-line BSC (57.0% annually from the second year), patients move rapidly and permanently to second-line BSC. The company model therefore implies patients receiving 2nd line BSC will remain in a health state associated with poorer quality of life and increased costs for the majority of the time horizon.

Crucially, the implied assumption that patients failing treatment remain perpetually in state of poor disease control is inconsistent with the longer-term evidence and expert opinion. This is illustrated in Table 37 which reports data from the LIBERTY-AD-CHRONOS trial[44]. There are several important features of these data. Firstly, a substantial proportion of patients achieve a clinically important reduction in symptom severity between Week 0 and Week 16. Secondly, a substantial proportion of these patients lose those improvements by Week 52 (this is illustrated by the conditional response rates reported in TA534[20]), suggesting that such improvements are temporary, and fluctuate over longer periods. Thirdly, the rates of patients achieving improvements in symptoms are near constant across Week 16 to 52, suggesting that for every patient that loses their disease control, another sees an improvement in their symptoms.

Table 37. Number and percentage of patients in the placebo arm of the LIBERTY AD CHRONOS trial who achieved EASI50 across the study period.

Week	1	2	4	6	8	12	16	20	24	28	32	36	40	44	48	52
EASI50 (%)	39 (15%)	71 (27%)	104 (39%)	119 (45%)	114 (43%)	107 (41%)	97 (37%)	93 (35%)	93 (35%)	98 (37%)	90 (34%)	85 (32%)	82 (31%)	77 (29%)	82 (31%)	79 (30%)
EASI75 (%)	6 (2%)	22 (8%)	45 (17%)	61 (23%)	68 (26%)	65 (25%)	62 (23%)	63 (24%)	69 (26%)	64 (24%)	64 (24%)	66 (25%)	67 (25%)	62 (23%)	62 (23%)	57 (22%)
EASI90 (%)	2 (1%)	9 (3%)	16 (6%)	22 (8%)	29 (11%)	29 (11%)	33 (13%)	38 (14%)	41 (16%)	42 (16%)	38 (14%)	47 (18%)	47 (18%)	44 (17%)	47 (18%)	41 (16%)

This pattern is exactly what we would expect from a BSC group experiencing waxing and waning of disease, and suggests that there will always be a non-negligible proportion of patients receiving BSC who have good disease control. Further, the relative stability of the proportion of patients achieving EASI50[44] indicates that the observed rates of placebo 'response' are therefore likely to be driven primarily by regression to the mean, owing to strict trial inclusion criteria, where patients are in an uncontrolled disease state. As such, the observed rate of placebo 'response' would indicate the proportion of patients who achieve disease control at any one point in time.

This failure to acknowledge that a proportion of BSC patients will have good disease control is an important omission and results in the effectiveness of BSC being underestimated, with corresponding consequences for the resulting ICER. Rather than modelling BSC as a line of therapy, where patients may experience loss of response and hence transition to another health state, the ERG's preference would be to model BSC as a health state in which the costs and QALYs are an average of responders and non-responders. The ERG considers this to better reflect the waxing and waning nature of AD, and better reflects how BSC is used in clinical practice.

Induction period of baricitinib

In the economic model, the company assume that patients enter into an 'induction' period of their allocated treatment. For all three treatment arms, this is modelled as a set of four, temporary tunnel health states, which can only be visited once in a fixed sequence. As each cycle length is four weeks, this induction period lasts for 16 weeks, at the end of which, patients are assessed for response. During the induction period, patients are unable to discontinue their treatment, and can only transition to the 'Death' health state.

With regards to baricitinib, the ERG is concerned that a 16-week induction period may not reflect how baricitinib is used in practice, nor the patterns of response observed in existing trial data. While, the ERG acknowledges that the bulk of the trial evidence assumes a 16-week induction period, the draft SmPC for baricitinib provided by the company indicates that discontinuation of baricitinib should be considered if the patient shows no response by 12 weeks. The SmPC suggests that partial

responders may continue treatment, which is perhaps reflected in the use of the company's composite outcome with a lower threshold for response than the IGA0-1/EASI75 definition of response used in the trials. This discrepancy between what is modelled in the CS and that of draft SmPC is likely to have an impact on the ICER, as patients who do not respond to baricitinib in this period will be modelled with a higher quality of life.

At the clarification stage, the ERG asked the company to comment on their expectations of how baricitinib is likely to be used in practice given the recommendation in the draft SmPC. The ERG also requested that the company provided a scenario in which a 12-week induction period is modelled for baricitinib. The company declined to provide such a scenario analysis. The company stated that they believed an early clinical assessment of efficacy would risk discontinuing treatment in patients who would go on to respond. However, the ERG does not consider this to be supported empirically, and highlights Figure 8a in the company's clarification response (reproduced as Figure 6 and Figure 7 below), which illustrates that the proportion of patients responding (EASI75) peaks in Week 8, and drops to what might optimistically be described as a plateau beyond this time point.

Figure 6 Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI75 over trial period (reproduced from PFC Response Figure 8a)

While equivalent data are not available for the composite EASI & DLQI outcome, Figure 6 demonstrates that the majority of patients achieved a ≥ 4-point improvement in DLQI within one week of initiating baricitinib, after which point the proportion of patients responding peaks and plateaus between Week 2-4, before dropping gradually for the remainder of the trial period. These figures demonstrate the rapid response of symptoms to baricitinib treatment, and a similar pattern can

be seen in pruritus symptoms (See CS Figure 12). This is not supportive of the company's argument that the response period stipulated in the SmPC should be ignored to allow for more patients to respond.

Figure 7 Proportion of patients in BREEZE-AD4 (JAIN) achieving a > 4-point improvement in DLQI score over the trial period (reproduced from CS PFC response Figure 17a)

The ERG, therefore, does not agree that the evidence presented in the CS (Section 2.6.1) supports the company's suggestion that an earlier clinical assessment risks patient discontinuation before they show response.

Given the above and guidance for use in the SmPC, the ERG disagrees with the company's decision to model the first clinical assessment at 16 weeks. The ERG is, however, unable to present scenario analysis using a 12 week endpoint due lack of appropriate data to populate it.

4.2.2 Population

The company's base-case analysis considers patients with moderate-to-severe AD who have previously failed one or more systemic therapies. As discussed in Table 4 and Section 3.2.2, this population is narrower than that covered by the anticipated market authorisation for baricitinib, which covers all adult patients with moderate-to-severe AD who are eligible for systemic therapy. The CS explains that the narrower modelled population *"is consistent with the anticipated eligible patient population for baricitinib in UK clinical practice"* and reflects the company's positioning of baricitinib as an alternative to dupilumab. With respect to this latter point, the ERG notes that the modelled population also reflects the recommendations made in TA534.[20]

In line with the narrower focus of the base-case analysis, the modelled population is based upon the JAIN + JAIN-like JAIY trial populations, which includes only patients with a history of intolerance to, contraindication to, or had an inadequate response to ciclosporin. The baseline characteristics of the modelled population are presented in Table 38 and include age, sex, EASI score, and EQ-5D. Of these listed characteristics, age and sex are directly parametrised in the model and are used to drive per cycle mortality. Age also drives age related utility adjustments, as well as the model time horizon which is set so that patients are modelled up to 100 years of age.

Table 38 Baseline patient characteristics (adapted from Table 82 CS and executable model)

	Age (years)	Gender (% Male)	Mean EASI score	Mean EQ-5D HIS
Base-case population (n = 293)	[redacted]	[redacted]	[redacted]	[redacted]
JAIN Europe subgroup	[redacted]	[redacted]	[redacted]	[redacted]
JAIN only	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: EASI: Eczema Area and Severity Index; EQ-5D: EuroQol 5 Dimensions; HIS: Health index Score.

In addition to the base-case analysis, the company also presents scenario analyses considering several alternative populations: JAIN Europe, JAIN only, and JAIN-like JAHL and JAHM. No scenario analysis is, however, presented for other potentially relevant subgroups including subgroups defined with respect to disease severity and skin colour, both of which are listed in the NICE scope. The omission of a disease severity subgroup analysis was justified on the grounds that the clinical classification of severity is applied inconsistently in practice, meaning that it is not possible to meaningfully define separate subgroups of moderate and severe AD. The omission of a skin colour subgroup analysis was justified on the grounds that there is insufficient data available to populate such an analysis. The ERG notes in relation to this latter subgroup that the JAIN Europe subgroup analysis may act as reasonable proxy for a white only subgroup as there are few non-white patients (< 5%) in this analysis. No counterpart analysis was presented for other skin types.

ERG comment

Disease severity and eligibility for treatment

The ERG has some concerns regarding the generalisability of the trial data to the modelled population. The inclusion criteria applied in all the pivotal trial evidence presented by the company require patients to have IGA ≥ 3 and EASI score > 16. This minimum requirement is, however, potentially overly restrictive and may mean that the recruited population has more severe disease than patients treated in practice. As described in Section 3.2.4.3, there is currently no widely accepted definition of what constitutes moderate disease. Examination of the literature, however, suggests that many clinicians consider eligibility for systemic treatment an indicator of moderate disease as

evidenced by several large cohorts.[45] While this represents a somewhat circular definition, it potentially means that the eligible patient population may be wider than that considered in the trials. This is further supported by several published strata based on EASI score, as described in Section 3.2.2, which indicates that the mean EASI score (30.1) in the modelled population represents severe disease.

The implications of excluding more moderate patients are difficult to fully distil. Subgroup analysis reported for the monotherapy population (Table 49 of CS) suggest baricitinib is more effective in a moderate population, which in turn suggests that results based on a moderate population would favour baricitinib. This subgroup analysis, however, defines moderate disease with respect to the recruited population (moderate disease defined as an IGA score of 3) and as such it is difficult to be certain that similar results would be seen if the population was extended to include patients will less severe disease. It is also unclear whether similar results would be observed in a combination therapy population (data were requested by the ERG at the clarification stage, but the company declined to provide this). Further, in terms of the economic analysis, it is possible that the higher response rates may be ameliorated by reductions in the quality of life gains associated with response; moderate patients have potentially less to gain from treatment and therefore the utility gains associated with response may also be lower, meaning treatment success is relatively less valuable in QALY terms. The ERG attempted to quantify the quality of life gains in moderate patients by requesting that the company provide utility values separately for moderate and severe patients. The company, however, chose not to provide these in their response.

Ethnicity and skin colour: East Asian patients

Ethnicity and skin colour may represent important treatment effect modifiers, as there is a body of evidence suggesting that immune phenotypes, the primary drivers of disease in AD, differ across ethnic groups.[46-49] Because of this, baricitinib may be more effective in some groups than others. Evidence suggesting such a differential effect is observed in the subgroup analysis presented in Table 49 of the CS for the JAIY population, which show that baricitinib is substantially more effective in patients recruited to European centres compared with Japanese centres.

As discussed in Section 3.2.2 the ERG requested that the company comment on these results and any biological reason why East Asian patients may not benefit from treatment with baricitinib. The company's response noted the potential for biological differences which may explain this result, but provided evidence to suggest that this is more likely to be as a result of differences in other geographical factors that may impact on treatment efficacy such as use of concomitant medications, adherence to treatment, and differences in the investigator assessment of the disease. Specifically, the company highlighted results of regional subgroup analyses, which do not support a lower response for

baricitinib 4 mg in East Asian countries not including Japan, suggesting there is not a specific effect of East Asian ethnicity.

While the ERG notes that the anticipated market authorisation for baricitinib does not preclude the use of baricitinib in any particular ethnicity,[33] the noted differences in disease pathogenesis and evidence of differential effects in Japanese centres do represent a source of uncertainty and suggest that the data from this cohort may be less relevant to the decision problem than patients recruited from other centres.

Ethnicity and skin colour: Black patients

Related to the above concerns regarding the effectiveness of baricitinib in East Asian patients, the ERG also finds concerning that there were [redacted] Black patients included in the JAIN and JAIN-like, JAIY, JAHL, and JAHM populations comprising the company's ITC.[47, 50, 51]

As is described in Section 3.2.4.3, the ERG requested that the company to comment on the potential for differences in response for black people, with the company response outlining that the trial program was not designed to investigate baricitinib efficacy in black patients compared with other patient populations. The company further stated that ethnic differences in the cytokine pathways involved in atopic dermatitis were noted in the dupilumab appraisal (TA534[20]), but not explored due to lack of evidence.

The reasons for [redacted] black patients from the trial are unclear, though the ERG notes that several centres were located in countries where black people represent a non-negligible proportion of the population. [redacted] black patients from the AD-BREEZE trials, however, means that it is not possible to establish that baricitinib is effective in this population. Further, given the noted differences in disease pathology, and differences in efficacy demonstrated across ethnic groups in other inflammatory disorders, it is questionable whether it is appropriate to assume the efficacy results observed in white patients are transferrable to other ethnicities.

Positioning as a comparator to dupilumab

The modelled population of patients who have failed one or more systemic treatments reflects the company's positioning of baricitinib as an alternative to dupilumab, and aligns with the modelled population in TA534[20]. The ERG considers this reasonable in principle but notes that this position is likely only relevant to an incident population who are currently naïve to dupilumab rather the prevalent population. This sub-population of dupilumab experienced patients is likely to represent a significant number of patients, and is not well represented by the JAIN and JAIY JAIN-like populations (only [redacted] of JAIN patients have previous dupilumab experience).

Further, the validity of the modelled population is somewhat conditional on dupilumab representing the most appropriate comparator to baricitinib. As outlined in Table 4 and further discussed in Section 4.2.3, there are several reasons to consider that immunosuppressive agents such as ciclosporin, methotrexate, and azathioprine represent a more natural comparator to baricitinib than dupilumab due to the similarities in their mode of action. In which case, the broader population covered by the marketing authorisation may be the most appropriate population in which to consider baricitinib.

4.2.3 Interventions and comparators

The economic model compared baricitinib combination therapy with dupilumab combination therapy, and BSC in a fully incremental analysis i.e. an analysis where all three alternatives were considered simultaneously. A summary of the modelled interventions and comparators is included in Table 39 and outlined below.

Table 39: Summary of modelled interventions and comparators

Baricitinib combination therapy	**Dupilumab combination therapy**	**1st line BSC**	**2nd line BSC**
Daily 4mg baricitinib	Loading dose of 600mg dupilumab followed by 300mg every two weeks	Blended comparator (66%, Mometasone (0.1%; TCS); 25% Tacrolimus (0.1%; TCI), Prednisolone 30mg)	Blended comparator (66.7%, Mometasone (0.1%; TCS); 22.3% Tacrolimus (0.1%; TCI), Prednisolone 30mg)
Concomitant treatment			
Bathing products: (33% Aqueous cream; 25% Dermol 200 shower emollient; 17% Aveeno bath oil; 15% Dermol 600 bath emollient; 10% Oilatum bath formulation)	Bathing products: (33% Aqueous cream; 25% Dermol 200 shower emollient; 17% Aveeno bath oil; 15% Dermol 600 bath emollient; 10% Oilatum bath formulation)	Bathing products: (33% Aqueous cream; 25% Dermol 200 shower emollient; 17% Aveeno bath oil; 15% Dermol 600 bath emollient; 10% Oilatum bath formulation)	Bathing products: (33% Aqueous cream; 25% Dermol 200 shower emollient; 17% Aveeno bath oil; 15% Dermol 600 bath emollient; 10% Oilatum bath formulation)*
Emollients: (Aveeno cream, Cetraben ointment, Dermol cream, Diprobase ointment, Epaderm ointment, Hydromol ointment, White soft paraffin 50%/ liquid paraffin 50% ointment, Oilatum cream)	Emollients: (Aveeno cream, Cetraben ointment, Dermol cream, Diprobase ointment, Epaderm ointment, Hydromol ointment, White soft paraffin 50%/ liquid paraffin 50% ointment, Oilatum cream)	Emollients: (Aveeno cream, Cetraben ointment, Dermol cream, Diprobase ointment, Epaderm ointment, Hydromol ointment, White soft paraffin 50%/ liquid paraffin 50% ointment, Oilatum cream)	Emollients: (Aveeno cream, Cetraben ointment, Dermol cream, Diprobase ointment, Epaderm ointment, Hydromol ointment, White soft paraffin 50%/ liquid paraffin 50% ointment, Oilatum cream)*
TCS: Mometasone 0.1%	TCS: Mometasone 0.1%	TCS: Mometasone 0.1%	TCS: Mometasone 0.1%*
			TCI: Tacrolimus 0.1%

Abbreviations: BSC: best supportive care; TCI: topical calcineurin inhibitor TCS: topical corticosteroids.
*Double intensity

Baricitinib combination therapy was modelled as consisting of orally administered baricitinib and concomitant supportive care. Dosing for all patients receiving baricitinib was 4 mg once daily in line

with the SmPC for patients aged under 75 years of age. Concomitant supportive care was assumed to consist of bathing, emollient products, and background TCS (0.1% Mometasone).

Dupilumab combination therapy was assumed to consist of dupilumab delivered via self-administered sub-cutaneous injection and concomitant supportive care. Dosing of dupilumab was a 600 mg loading dose followed by 300 mg every other week; this aligns with marketing authorisation for dupilumab and the relevant supporting evidence from the CAFÉ and CHRONOS trials.[52, 53] Concomitant supportive care was identical to that received by patients receiving baricitinib.

Best supportive care, which the ERG denotes as 1st line BSC, was modelled as a blended comparator consisting of topical mometasone (TCS), topical tacrolimus (TCI), and oral Prednisolone (a corticosteroid). Patients were also assumed to receive concomitant supportive care identical to that received by baricitinib and dupilumab patients. Patients could receive 1st line BSC either as an initial treatment or following failure of either baricitinib or dupilumab.

Patients failing to respond to 1st line BSC following induction or after loss of subsequent response were also assumed to receive BSC. The composition of BSC received by non-responders, however, differed to that outlined above and in effect represents another line of treatment distinct from 1st line BSC. This 2nd line BSC consisted of a blended comparator of topical mometasone (TCS), topical tacrolimus (TCI) and oral Prednisolone (a corticosteroid identical to that modelled for 1st line BSC). Second line BSC also included concomitant supportive care which consisted of bathing and emollient products, background TCS (0.1% Mometasone) similar to that received as part of 1st line treatments. The intensity of these supportive treatments was, however, assumed to be double that used in 1st line BSC. Additionally, 2nd line BSC was also assumed to include a TCI (0.1% Tacrolimus).

In addition to the systematic treatments described above, patients were modelled to receive a range of rescue therapies in response to disease flares. Rescue therapies included betamethasone valerate cream (TCS), cutivate (0.05%) cream (a TCS), eumovate (0.05%) ointment (a TCS), dermovate (0.05%) cream (a TCS), prednisolone (5mg) (a systemic steroid) and tacrolimus (0.1%) ointment (a TCI). The rate of rescue therapies used was dependent upon treatment received, see Section 4.2.7.2 for details.

Several comparators listed in the NICE scope were not included in the economic analysis: Phototherapy, alitretinoin (approved for the treatment of AD affecting the hands) and systemic immunosuppressants (azathioprine, ciclosporin, methotrexate and mycophenolate mofetil). The company justified the exclusion of phototherapy and immunotherapies on the grounds that phototherapy is typically given in early lines of therapy. The ERG notes that this is broadly consistent with TA534[20] and that phototherapy is also considered as part of health state costs. The omission of

alitretinoin was justified by the company on the grounds that it is licenced only for the treatment of hand eczema, and therefore not relevant to the current population, which is also consistent with TA534[20]. The exclusion of immunosuppressants was justified in relation to the model population which focuses on patients who have failed one of more systemic therapies.

ERG comment

Sequential therapy

The company's economic model does not consider the potential for sequential treatment with active therapies i.e. the possibility that baricitinib and dupilumab may be used in sequence, either as dupilumab followed by baricitinib or baricitinib followed by dupilumab. The ERG notes that the functionality to model treatments as part of sequence has been added to the executable model but is not used in the company's base-case analysis or any presented scenario analysis.

At the clarification stage the ERG enquired as to the company's justification for this approach. The company's response stated that the anticipated positioning for baricitinib in UK clinical practice is as fifth line alternative to dupilumab or BSC for those patients for whom dupilumab is not recommended or contraindicated. The company further highlighted that baricitinib would not be considered after dupilumab in the treatment pathway, due to the very limited evidence from the BREEZE clinical trials for the efficacy and safety of baricitinib in patients who have received prior dupilumab.

The ERG does not agree with the company's line of reasoning and notes that nowhere in the CS is it stated that the company wishes NICE to consider baricitinib for use only in patients naïve to dupilumab. Further, by the company's logic it would be necessary for any NICE recommendation for baricitinib to explicitly prohibit the use of dupilumab as subsequent therapy, as there is currently nothing in the NICE recommendations or marketing authorisation for dupilumab that precludes its use following another biologic agent. The ERG questions the acceptability of any such recommendations, given the superior effectiveness of dupilumab compared with baricitinib and the limited treatment options available in this population.

The ERG further questions the clinical rationale for excluding treatment sequences, and notes that clinical advice provided to the ERG suggests that these treatments are very likely to be used in a sequence by clinicians. The ERG also notes precedent for the evaluation of treatment sequences rather than simple comparisons of active treatments in many of the recent appraisals of biologics for the treatment of psoriasis.[54-56] In these appraisals it is typically assumed that patients will cycle through 3 or more active treatments, and as such, treatments are often modelled as part of a sequence.

The ERG therefore considers the modelling of baricitinib as part of a sequence to be highly relevant to this appraisal and to be in line with how baricitinib will be used in practice. The modelling of

treatment sequences also allows for the positioning of a drug within the pathway to be optimised. For example, it may be more cost-effective to use baricitinib as a 5th line treatment before dupilumab. The ERG explored the cost-effectiveness of treatment sequencing in a scenario analysis (Section 6.1)

Omission of comparators listed in the scope

The ERG is in general agreement with the company's reasoning regarding the exclusion of phototherapy and alitretinoin as comparators within the economic analysis. The ERG is, however, concerned about the exclusion of immunosuppressants as a comparator. The ERG notes the focus of the company's economic model on a population who have failed one or more immunosuppressants, but considers that there is scope for further immunosuppressant use in many of these patients given both the availability of several different immunosuppressants and the potential for response to be re-induced. The ERG also notes that comparisons to ciclosporin were made in TA534[20] for the committee to consider (though these were not considered suitably robust for decision making).

Given these concerns, the ERG requested the company provide further justification for excluding immunosuppressants as a comparator and to provide a relevant scenario analysis; however, the company declined to attempt such an analysis.

As outlined above, the ERG does not consider this line of reasoning plausible, given the availability of numerous alternative immunosuppressants. Failure of one therefore does not preclude the use of another such agent in subsequent lines of therapy. Further, the mode of action of baricitinib potentially places it as a more a natural comparator to the immunosuppressants such as ciclosporin and methotrexate, than dupilumab. This is because as a JAK1/JAK2 inhibitor, baricitinib has a more targeted mode of action as compared with other systemics, but is more broadly immunosuppressive than dupilumab which has a more focused mode of action, distinct from both baricitinib and immunosuppressants. Consequently, clinicians are only likely to consider baricitinib as an alternative to dupilumab upon exhaustion of systemic immunosuppressants. The comparisons to dupilumab and BSC presented in the CS are therefore potentially most plausible when baricitinib is positioned further down the treatment pathway, following failure of multiple immunosuppressants, or following failure of dupilumab.

Dosing in over 75s

The SmPC for baricitinib states that patients over the age of 75 or those otherwise susceptible to infection should be treated at lower dose of 2 mg. In the model approximately [redacted] of patients remain on treatment by the age of 75. The lower dose of baricitinib is however not applied in these patients, nor are any patients assumed to require the lower dose due to increased infection risk. Evidence from JAIN, JAIY and JAHL shows that the 2 mg dose is less effective than the 4 mg. As a result, the model may overestimate the effectiveness of baricitinib in these patients. Further, because the acquisition

costs of the 2 mg dose are the same as the 4 mg dose, this will lead to the model underestimating the ICER. Clinical advice received by the ERG suggests that the proportion of eligible patients over 75 is likely to be small and that due to co-morbidities present in this population, phototherapy and topical agents are often the preferred treatment options. This may suggest that the time horizon adopted is too long, as realistically patients will not continue or initiate treatment when over the age of 75.

4.2.4 Perspective, time horizon and discounting

The analyses assumed the perspective of the NHS and Personal Social Services (PSS), and future costs and benefits were discounted at 3.5% per annum. This is in line with the NICE reference case.

The time horizon of the base case analyses was approximately 62 years and was based on modelling patients up to 100 years of age. The company justified the choice of time horizon noting that it is consistent with the NICE reference case and that it ensures all costs and QALY gains associated with the interventions are fully captured.

The ERG considers the choice of a time horizon reasonable in the context of AD and the expectation that it is a life-long condition. However, the ERG notes that this choice of such a long time horizon does mean that the comparatively short-term effectiveness evidence is projected over a very long period, increasing uncertainty in the model results. Furthermore, the long time horizon also means that many patients spend much of the modelled time horizon on BSC. This means that rates of discontinuation associated with baricitinib and dupilumab have a significant impact on the outcomes of the model. It also means that the costs and utilities associated with BSC influence not only the comparison between baricitinib and BSC, but also the comparison between baricitinib and dupilumab.

Further, as outlined in Section 4.2.3, the adoption of an extended time horizon means that some patients are assumed to remain on treatment well into old age. This may not be realistic given the burden of comorbidities in this population and it may have therefore been more reasonable to have modelled a shorter time horizon, where the maximum age is lower.

4.2.5 Treatment effectiveness and extrapolation

4.2.5.1 Assessment of treatment response

Clinical response is defined in the company's base-case analysis as a relative improvement in patients' EASI score of 50% relative to baseline, i.e. EASI50 after 16 weeks of treatment. In addition, patients required an absolute reduction in their DLQI score of 4 points or more relative to baseline. Patients who discontinued treatment prior to Week 16 were classed as non-responders for the purposes of the economic analyses.

The company's submission explored the use of a number of alternative definitions of clinical response and their effect upon the cost-effectiveness of baricitinib. The company presented scenarios which used EASI50 only, EASI75, and Itch NRS ≥ 4 at 16 weeks as the definition of response (see Section 5.1).

The rates of clinical response applied in the base-case for baricitinib are derived from the company's ITC, which included data from BREEZE-AD4 (JAIN) and data on patients from BREEZE-AD7 (JAIY) who had previously failed on, or were intolerant or contraindicated to ciclosporin, referred to as JAIN-like JAIY. Equivalent response probabilities for dupilumab were based on a population comprising CAFÉ and CAFÉ-like CHRONOS patients. Indirect comparisons were made using placebo + TCS (as a proxy for BSC) as a common comparator in the company's ITC analysis described in Section 3.4 Response probabilities were generated from the ITC by simply adding the risk difference (RD) for dupilumab vs placebo from the CAFÉ plus CHRONOS CAFÉ-like population [redacted] to the placebo response rate in the JAIN plus JAIN-like JAIY population (31.25%). The company assumed that the standard error associated with the adjusted response probability would be the same as in the dupilumab trial.

The base-case response probabilities for baricitinib, dupilumab, and BSC are presented in Table 40, along with probabilities of response according to selected alternative response definitions. In the base-case analysis, patients gained no further health benefit by achieving a higher level of response, i.e. EASI75-89. When response was defined as achievement of EASI50 only, patients who achieved EASI75 had a higher HRQL than those whose response was between EASI50-74 (see Section 4.2.6 for more detail).

Table 40 Summary of response probabilities for alternative response definitions at 16 weeks (adapted from CS Tables 83 and 84)

	Response probabilities from the ITC (calculated by company) (%, SE%)		
	EASI50 + DLQI≥4 (base-case)	**EASI50**	**EASI75**
Baricitinib	48.99 (4.09)	[redacted]	[redacted]
Dupilumab	79.25 (3.00)	[redacted]	[redacted]
BSC	31.25 (3.86)	[redacted]	[redacted]

ERG Comment

Calculation of response rates from the ITC

The response rates applied in the model were derived using absolute measures of the treatment effect rather than relative effects. The use of absolute treatment effects can result in bias where the response rates in the common comparator (placebo +TCS) differ across studies and are therefore less robust to

differences between the contributing trials. A comparison of placebo response rates for contributing JAIN+JAIY JAIN like and CAFÉ plus CHRONOS CAFÉ-like does show a degree of difference, with consistently lower placebo response rates reported for the CAFÉ plus CHRONOS CAFÉ-like population (31% vs 26% for the base case analysis). The company's approach will therefore tend to favour dupilumab overestimating the treatment effect.

The company's approach also departs from the analysis suggested in NICE DSU TSD5[3], which recommends that absolute response rates on a control and the relative effects on the active treatments used in the ITC are pooled on the log-odds scale. Using this approach would also have allowed the company to use the correct SE derived from the ITC in the probabilistic sensitivity analysis, rather than using an unadjusted value based on the dupilumab trial. The ERG's updated base-case analyses in Section 6.3 use response rates estimated using the relative treatment effects used in the ITC, following the recommendations in TSD5[3] (see Section 6.1).

Defining a clinically meaningful response to treatment

The use of the *post hoc* composite of EASI50 and ΔDLQI ≥ 4 secondary outcomes is consistent with that accepted by the committee in TA534[20]; however, the ERG notes a number of weaknesses with this definition of response and that there are alternatives which may be more valid, and may provide a different picture of the relative effectiveness of baricitinib.

The ERG considers it uncertain whether the definition of response used by the company would be recognised and treated as clinically meaningful in practice. The submission from the British Association of Dermatologists for the present appraisal states that a clinically significant improvement is defined as a reduction in EASI score of 75% (i.e. EASI75), or a fall in IGA of 2 points. The company also stated in their clarification response that *"disease severity scales, including EASI (...) are not practical for routine use in clinical practice"*.

Further, previous studies have found poor correlation between EASI score and HRQL in AD.[20] This was supported by clinical advice received by the ERG, which indicated that morphological response (i.e. EASI-based) is less important to many patients than improvement in pruritus (itch) and resulting sleeplessness. While it is possible that this combination of outcomes may be better correlated with a higher HRQL, there did not appear to be any correlation observed in the regression analysis based on [redacted] patients presented by the company (see Section 4.2.6.) As discussed in the ERG's critique of the model structure (see Section 4.2.1), the model does not appear to be fit for the purpose of demonstrating any potential clinical benefit associated with baricitinib.

The ERG notes that both JAIN and JAIY studies assessed response using a ≥ 4-point improvement on the Itch Numeric Rating Scale at Weeks 4 and 16, which is widely considered to represent a clinically

meaningful improvement in itch symptoms in AD and psoriasis.[57, 58] The results of the company's ITC indicated that baricitinib offers similar effectiveness to dupilumab in terms of itch (see Table 4), and thus could prove to be a valuable treatment option where pruritus is an important factor in a patient's disease.

Primary vs secondary censoring

As discussed in Section 3.2.4.1, two alternative censoring methods were employed to analyse the supporting trial data:

- **Primary censoring** where patients were censored and classified as non-responders on initiation of rescue therapy with TCS;
- **Secondary censoring** where patients were censored and classified as non-responders only on the initiation of systemic rescue therapies.

As described previously, use of rescue therapy is not a good indication of patients losing response and does not necessarily indicate treatment failure. Clinical advice received by the ERG suggests that rescue medication may be used concomitantly with systemic treatments including baricitinib and dupilumab and will be used to overcome the often short-term symptom flares that many patients will periodically experience. In such circumstances, clinicians would expect that, disease control can be re-established on the same medication following rescue, and therefore use of rescue medication would not necessarily be grounds to discontinue treatment. The secondary censoring rule may therefore be considered a better reflection of clinical practice, permitting some use of rescue therapy. The ERG notes that such an approach would also align with TA534[20] where similar conclusions were drawn regarding the use of rescue medications.

4.2.5.2 Treatment discontinuation between Week 16 and 52

Following the 16-week treatment induction period, patients who have achieved a clinically significant response enter the Markov component of the model, representing the maintenance phase, in which patients continue treatment with a stable response. Between Weeks 16 and 52, the magnitude of response remains constant, but patients are at risk of discontinuing treatment as a consequence of losing their response to treatment.

The company were unable to produce data on continued treatment response between 16 and 52 weeks from the BREEZE trials, and concluded that there are no reliable estimates for treatment discontinuation and durability of response for baricitinib. The company therefore chose to assume equivalence to dupilumab in their base-case analysis. This assumption forms much of the basis of the cost-effectiveness estimates produced for baricitinib; however, the company provided no biological or clinical rationale in support of this assumption.

The probability of continued response at Week 52 amongst Week 16 responders in the CHRONOS trial was used to calculate a per-cycle discontinuation rate, which was then applied to baricitinib. The modelled response probabilities at Week 52 conditional on response at Week 16 are presented for the three selected response definitions Table 41. Per cycle discontinuation probabilities were calculated as a linear rate simply by rescaling the probability of discontinuing over weeks 16 to 52 into the four-week cycle length, yielding a per-cycle discontinuation rate for dupilumab and baricitinib of 0.697%.

At the clarification stage the ERG noted that some of the discontinuation rates applied in the model did not align with the values reported in Table 41, with the company erroneously using data for EASI75 when applying the EASI50 and EASI50 + DLQ I≥ 4 response criteria. These were corrected by the company and a revised model supplied to the ERG. This update had only a minor impact on the resulting ICER (corrected ICER reported in Company PFC response, Table 12).

Table 41 Summary of response probabilities for alternative response definitions at Week 52, conditional upon Week 16 response (adapted from CS Tables 85 and 86, Company's executable model)

Response definition	Week 52 conditional response probability (%, SE%)					
	Baricitinib		Dupilumab		BSC	
	52-weeks	Per cycle	52-weeks	Per cycle	52-weeks	Per cycle
EASI50 + DLQI≥4 (base-case)	93.9 (2.8)*	0.697	93.9 (2.8)	0.697	76.7 (4.8) †	2.90
EASI50	94.5 (2.5)*	0.627	94.5 (2.5)	0.627	81.3 (3.5) †	2.27
EASI75	82.1 (5.3)*	2.17	82.1 (5.3)	2.17	70.6 (6.4) †	3.79

* Assumed to be the same as dupilumab † Based on corresponding values from TA534

The company provided conditional response probabilities (EASI50 + ΔDLQI ≥ 4) at treatment weeks 16 and 52 for patients enrolled in the JAHN extension study. Of [redacted] patients who received 4 mg baricitinib as a monotherapy in both the originating and extension studies, [redacted] achieved the composite response outcome at Week 16. Of these patients, [redacted] were still responding at Week 52. Data on a smaller population of placebo patients were also available; [redacted] placebo patients [redacted] responding at Week 16 remained in response at Week 52. However, this only included those patients who also achieved an IGA response at Week 16, as the remaining patients were re-allocated to receive baricitinib, or discontinued the study.

ERG Comment

Validity of assuming continued response

The majority of health benefit generated by baricitinib in the company's model is based on an assumption of equivalence with dupilumab in terms of long-term efficacy and discontinuation. The

ERG has a number of substantive issues with this assumption, and highlights available data that suggest these benefits are unlikely to be realised in practice.

The company's analysis assumes an enduring and diverging difference in long-term response rate between baricitinib and BSC, based on long-term data on dupilumab. However, this is not supported by available trial data.

Firstly, by Week 24 in the JAIN study, baricitinib was no longer associated with a statistically significantly higher proportion of patients maintaining response versus placebo defined by EASI50/75/90, IGA ≤ 1, DLQI ≥ 4, and SCORAD75. As can be seen in Figure 6 and Figure 7, the proportion of patients continuing to respond on baricitinib gradually decreases over time. At Week 24, the only modelled definition by which patients on baricitinib maintained a superior response to placebo was ≥ 4-point Itch NRS improvement . Notably, the proportion of patients achieving the primary study outcome (IGA ≤ 1) fell from 21.7% [redacted] at Week 16 to [redacted] at Week 24, in comparison to 9.7% [redacted] over the same period in placebo patients. By these outcome measures it appears any benefit associated with baricitinib over placebo are reducing by Week 24.

Furthermore, the assumption of equivalence with dupilumab does not appear reasonable based on available trial data. At Week 16, the EASI score of dupilumab (300mg q2w) patients in the CHRONOS trial had improved by 76.7%, versus 43.2% on placebo[20]. In contrast, JAIN patients achieved an improvement of [redacted] on baricitinib, and [redacted] on placebo at Week 16. Further, symptoms of patients treated with dupilumab in the CHRONOS trial appeared to improve over the duration of the study, with a mean improvement in baseline EASI score of 78.3% at Week 52. Week 24 data from JAIN does not compare favourably, with patients appearing to lose some improvement in symptoms. Baricitinib patients' percent improvement from baseline declined to [redacted], a reduction of [redacted] on Week 16 levels, while placebo patients' scores continued to improve over this period by [redacted]. These data are not supportive of an assumption of equal efficacy and discontinuation between baricitinib and dupilumab from Weeks 16-52: while there appears to be some effect waning as early as eight weeks into the maintenance period on baricitinib, the treatment effect of dupilumab appears to remain stable until at least Week 52. Even assuming similar levels of effect waning between the two treatments within the first year, the lower absolute improvement in EASI score on baricitinib means patients will lose their EASI50 response sooner than on dupilumab.

Secondly, the two technologies have vastly different mechanisms of action and modes of administration, and are thus likely to differ substantially with regards to long-term efficacy and adherence. The company states that they believe the assumption of equal discontinuation rates

between the two technologies to be a conservative assumption, as unlike for dupilumab, anti-drug antibodies will not be generated against baricitinib. However, the ERG considers treatment efficacy a far more important predictor of treatment discontinuation. Other small molecule inhibitors have previously been appraised by NICE in the context of established monoclonal antibody-based biologic therapies in related indications, and trials have typically found them to be less effective and to have higher annual discontinuation rates in comparison.[59, 60]

Comparison of baricitinib trial data with dupilumab equivalence

The company state that the observed discontinuation rates between Week 16 and Week 24 in the JAIN trial were not used as they would bias the results in favour of baricitinib; however, this is not the case. At Week 24, [redacted] of JAIN patients (4 mg BARI + TCS) had discontinued treatment, yielding a per cycle rate of [redacted], and thus a projected discontinuation probability at 52 Weeks of [redacted]. This means that the level of discontinuation observed between Weeks 16 and 24 in the JAIN trial was actually higher than that applied in the model between Weeks 16 and 52 (6.1%), based on the assumption of equivalence with dupilumab. Thus, this assumption also lacks face validity, and is likely to introduce substantial bias in favour of baricitinib.

The BREEZE-AD3 (JAHN) extension study represents another potentially superior source of data. As discussed above, the company provided discontinuation data for baricitinib 4 mg as a monotherapy for up to 52 weeks of treatment. Of [redacted] patients who achieved the composite response outcome at Week 16, [redacted] were still responding at Week 52, equating to a per cycle discontinuation rate of [redacted]. This is very substantially higher than the 0.7% rate modelled by the company.

Use of conditional response data

The ERG does not consider it appropriate to use conditional response data to model discontinuation between week 16 and 52. While the ERG acknowledges that this approach was accepted in TA534[20] and that loss of efficacy may be a primary driver of discontinuation in many patients, it is not the only factor which will lead to discontinuation. This is demonstrated in the discontinuation data from both JAIN and JAHN where several patients discontinued for reasons other than loss of efficacy.

Furthermore, the use of conditional response implies that a response-based stopping rule will be applied, such that the response criteria are utilised on an ongoing basis to evaluate whether patients are benefiting from treatment. The ERG considers this inconsistent with clinical practice and notes that TA534[20] recommendations do not impose any such formal stopping rule. Assessment of continued benefit is instead likely to be based on less formal criteria and is likely to be informed by combination of clinician judgement and patient experience, which may not align fully with the original response criteria applied.

The ERG also questions the consistency of using all-cause discontinuation rates in the post 52 week period, but not in the 16 to 52 week period. This is especially odd, given that the data used to model post 52 week rates is actually based on rates of discontinuation observed between week 16 and 52. The ERG therefore prefers to apply a single discontinuation rate across both the 16 to 52 week and post 52 week periods using available date on all cause discontinuation. Scenario analysis presented in Section 6.1 explore the application of several approaches to modelling discontinuation including the use of all-cause discontinuation data.

4.2.5.3 Long-term Treatment Discontinuation (Week 52 onwards)

The company's model accounts for long-term discontinuation, i.e. from Week 52 until the end of the model, using a fixed annual probability of treatment discontinuation. This rate represents discontinuation for any reason, and is applied to patients in the maintenance health state starting at the second year of the modelled time horizon. Patients who discontinue during this period are assumed to transition to BSC. When the comparator is BSC, or for patients who have already transitioned to BSC, those who discontinue are classed as non-responders for the remainder of the model.

As there were no placebo-controlled trial data available for baricitinib extending beyond 52 weeks of treatment, the company assumed that baricitinib would have the same long-term discontinuation rate as dupilumab. These data were likewise unavailable from the CHRONOS study; in TA534[20] the annual discontinuation rate for dupilumab was based on all-cause discontinuation by Week 52 among the Week 16 responders. The annual probability of discontinuation for the second and subsequent years of treatment on baricitinib and dupilumab is therefore 3.7% using the EASI50 and $\Delta DLQI \geq 4$ definition of response, and 5.1% using EASI75, as shown in Table 42. These figures are lower than the conditional response probabilities at Week 52, as there was no stopping rule for loss of response in the trial, and thus partially responding patients and those with no response continued treatment regardless. This may mean that the applied rates underestimate long-term discontinuation outside of the setting of a clinical trial, particularly if response-based stopping rules are applied.

The ERG reiterates that conditional response at Week 52 amongst responders to 4 mg baricitinib monotherapy in the JAHN study was [redacted]. This is substantially higher than the modelled rate, and is likely an underestimate due to discontinuation for other reasons not captured in this figure.

Table 42 Annual probability of discontinuation in company model (Year 2 onwards)

	Annual discontinuation probability (%)		
Response definition	**Baricitinib**	**Dupilumab**	**BSC**
EASI50 and DLQI≥4*	3.7%	3.7%	57.0%
EASI75*	5.1%	5.1%	57.0%

JAHN EASI50 and DLQI≥4 response probability	[redacted]	3.7%	57.0%

* Assumption based on CHRONOS/TA534

The company used the annual probability of withdrawal or use of rescue medication from the patients on the placebo arm in the CHRONOS trial as a proxy for the rate of long-term treatment discontinuation in patients who responded to BSC. The figure implemented in the model was 57%, which results in utility gains for BSC rapidly decreasing to zero over the first few years of the model. While this discontinuation rate is in line with the Committee's preferences in TA534[20], it is associated with a number of issues which may bias the results against BSC (see ERG Comment).

ERG Comment

Validity and inconsistency of long-term discontinuation rates

The ERG has concerns regarding the validity of the projections of long-term discontinuation applied in the company's model. The rates as currently applied lack simple face validity and have a substantial impact upon the relative cost-effectiveness of baricitinib and the comparator treatments. As discussed in Section 4.2.5.2, the ERG considers the assumption of equivalent long-term efficacy and discontinuation between baricitinib and dupilumab to be inherently flawed. This is taken a step further beyond Week 52, as beyond this point there is no evidence for the ongoing effectiveness of dupilumab, and health benefits accrued are based on an assumption of ongoing response and treatment adherence. The source of discontinuation rates for BSC also appears to be biased towards baricitinib, as patients requiring rescue therapy on BSC lose response, while those on baricitinib and dupilumab do not.

A major issue with the company's estimation of long-term discontinuation is the inconsistent sources of discontinuation rates between baricitinib and BSC. The company's model assumes that patients permanently lose utility gains achieved on BSC in response to a flare, moving to the non-responder health state until death. For patients on baricitinib and dupilumab, flare is not considered an indicator of loss of response or grounds to discontinue treatment, and thus patients remain in the maintenance health state and incur no reduction in HRQL. If flare rates used to estimate discontinuation for all treatments, annual discontinuation would be at least 16% for dupilumab, and substantially higher for baricitinib, based on the Week 24 JAIN figure of [redacted].

Secondly, as was discussed in TA534, use of rescue therapy is not a good indication of patients losing response, and thus the HRQL gains achieved on treatment. The ERG's clinical advisor explained that a trigger factor, such as exposure to allergens or a stressful event, may overcome the symptom control provided by medication and result in flare. Clinicians would expect that, in the case of such a flare, control can be re-established on the same medication following rescue, thus having a flare should not be conflated with loss of response. Rescue therapy forms part of BSC with the objective of improving

symptoms and returning a patient to pre-flare QoL. Therefore, the assumption that rescue therapy indicates permanent loss of any HRQL improvement does not appear to be clinically justified.

Thirdly, the rate of 57% used by the company appears to be derived from the full BSC population in the CHRONOS trial over the 52 weeks of treatment. It therefore inappropriately includes patients who did not respond to BSC, those who required rescue prior to Week 16, and the 33 patients who withdrew from the study prior to Week 16. Finally, the committee's preference was for this rate of discontinuation to apply between years 2-5, after which point any patients remaining on BSC receive a utility benefit for the remainder of the model. This was to reflect the small proportion of patients who have an enduring response to BSC; however, as the discontinuation rate applied is so high initially, this has a negligible impact on the proportion of patients achieving a long-term improvement in symptoms.

As described in Section 4.2.1, the ERG considers the issue of uncertainty around long-term health effects to be inadequately explored or captured in the company's model. The model structure lacks the flexibility to capture the relapsing-remitting nature of AD, and assumes that patients remaining on treatment during the maintenance phase are continuously responding. Patients who lose response do so permanently, and are modelled to have a continuously poor HRQL for the remainder of the model, rather than reflecting periods of good and bad symptom control over time. The model doesn't allow for continuing treatment through a response that appears to fluctuate due to the nature of the condition, and the use of more intense background therapies to suppress flares.

Furthermore, the assumption of significantly different and rapidly diverging extrapolations of long-term efficacy between baricitinib and BSC appears to lack simple face validity, based on available evidence from the JAIN trial. As discussed in Section 4.2.5.1, patients on baricitinib were not (statistically significantly) more likely than those on BSC to have an enduring clinical response defined in terms of EASI50/75/90, IGA ≤ 1, DLQI ≥ 4, and SCORAD75. The ERG therefore questions the assumption underpinning the majority of differential QALY generation between baricitinib and BSC, which is based on a significant improvement in long-term durability of response. This is assumption is based on long-term efficacy data for dupilumab, but available trial evidence for baricitinib is very limited and suggests that the ability for baricitinib to generate durable benefits relative to BSC may be limited given the observed trends in response rates.

4.2.5.4 Flare Control

Patients with moderate-to-severe AD suffer periods of acute exacerbation of symptoms, known as flares. The company state that flares occur roughly 10 times per year in patients with moderate-to-severe AD. Worsening of erythema during these periods can have a substantial effect on quality of life, disrupting sleep for 7.3 to 14.6 nights per flare, which leads to distress and anxiety among

patients. The company's model appears to assume that health benefits associated with reducing the rate of flares would be captured in the HRQL data collected in the clinical trials. The impact of flares is reflected only in the cost of topical and systemic steroids, and calcineurin inhibitors (see Section 4.2.7).

A key element of management of AD is the control of symptom exacerbation during flares. During these periods, brief but aggressive treatment using high potency TCS and systemic agents (e.g. ciclosporin, methotrexate) can be necessary. In the CHRONOS trial, dupilumab demonstrated flare suppression over long-term continuous use versus placebo (16% vs 52% respectively), and thus a significantly reduced need for rescue therapy. Flare frequency was not an endpoint recorded in the baricitinib clinical trials, but in their submission, the company state that the receipt of rescue medication can be considered a proxy for flare. At Week 24 in the JAIN trial, [redacted] placebo patients had been 'rescued', compared with [redacted] patients in the baricitinib 4mg arm, which implies no substantial flare control is associated with baricitinib.

The company have assumed that baricitinib is equally effective as dupilumab with regards to flare control in their model, with an annual flare rate of 0.18, compared to 0.78 on placebo. It was also assumed that patients on dupilumab and baricitinib cannot experience flares in the induction period, while patients on BSC can.

ERG Comment

The ERG does not agree that the company used the most plausible available estimates of flare frequency. The company's model assumes a > 4-fold improvement in flare frequency associated with baricitinib over placebo. However, more patients in the baricitinib arm of the JAIN trial required rescue therapy than those on placebo, therefore the ERG considers it plausible that the rate of flares between baricitinib and placebo could be equal. The ERG also notes that flare was common on baricitinib during the induction period of the JAIN trial, and thus it is inappropriate to assume patients do not experience flares during this period.

Given the company's own estimate of ~10 flares per year suffered by patients with moderate-to-severe AD, it is also likely that the rates modelled represent more than a tenfold underestimate of annual flare rate. The ERG have explored this in a scenario in Section 6.

4.2.5.5 Adverse events

The adverse event rates included in the model were described by the company as being based on the most frequent and serious events reported in the baricitinib AD trials. Rates for dupilumab and BSC were derived from TA534[20]. The company also provided revised AE rates at the clarification stage to

include allergic conjunctivitis for baricitinib, as these had been excluded from their original analysis (see Table 43).

Baricitinib AE rates were only available for the 16-week trial period, from which annual rates were calculated. Rates of adverse events remain constant throughout treatment, and are varied according to the treatment received. The only consequences of adverse events in the model are the costs associated with their treatment (discussed in Section 4.2.7), but the impact of serious adverse events is also captured in the treatment discontinuation rates derived from the trials (applied post-52 weeks). Any disutilities associated with AEs were assumed to be included in the trial-derived HRQL data.

The adverse events considered by the company's model, and the probabilities of their occurrence are presented in Table 43. These include revised rates presented as part of the company's clarification response.

Table 43 Adverse event rates (annual) used in the company model (adapted from CS Table 88, Page 145, and company clarification response Table 25, page 42)

Adverse event	AE probability			
	Baricitinib 4mg (original)	Baricitinib 4mg (revised)	Dupilumab	BSC
Injection site reaction	[redacted]	[redacted]	0.091	0.000
Allergic conjunctivitis	[redacted]	[redacted]	0.401	0.188
Infectious conjunctivitis	[redacted]	[redacted]	0.255	0.033
Oral herpes	[redacted]	[redacted]	0.055	0.110
Upper respiratory tract infection	[redacted]	[redacted]	0.000	0.000

ERG Comment

Internal consistency of adverse event rates

The ERG is concerned that the rates modelled for a number of AEs applied to the baricitinib arm were zero, and [redacted] those associated with BSC, given that the baricitinib combination therapy under assessment also includes BSC. As there is no evidence to suggest baricitinib reduces the rate of AEs associated with concomitant BSC, and no clinical rationale as to why this would be the case, the ERG requested justification for this assumption from the company. In their response, the company restated that AE rates for baricitinib were based on the JAIN and JAIN-like JAIY population, while BSC and dupilumab rates were taken from TA534[20]. The company presented adjusted AE rates as described above, which the ERG did not consider to have addressed the issue raised at the clarification stage, as the method used by the company halved the allergic conjunctivitis rate on baricitinib compared to the observed value. The ERG, however, further notes that AE rates have

almost no impact on the results of the economic analysis representing < 0.5% of total costs in the baricitinib arm.

Other potentially important adverse events

There are several other adverse effects of treatment with baricitinib that may have implications for resource use and patient/clinician preferences. The ERG's clinician suggested the observed rises in LDL cholesterol, seen in 12.2% of patients in the baricitinib safety dataset,[33] would likely mean ongoing monitoring of lipid levels would be necessary. Clinicians may also consider use of statins to control LDL levels in some patients. The ERG considers a scenario analysis including these potential additional monitoring and management costs in Section 6.2.

Other AEs not included in the model include headaches, which were recorded in 7.6% of patients, the reported character and frequency was not sufficient to assess any potential effect on treatment adherence, however. Infections were reported in [redacted] of patients up to 16 weeks in the 4mg treatment group, vs [redacted] of patients on placebo.[33] Upper respiratory tract infections were observed in [redacted] of baricitinib patients.[33] The draft SmPC states that baricitinib treatment should be withdrawn until infection resolves, and a decision not to re-initiate treatment may be considered. This may have the effect of increasing the rate of treatment discontinuation seen in clinical practice, as discontinuation due to adverse events and patient/clinician preference was not included in the rate applied up to Week 52.

4.2.6 Health related quality of life

Health outcomes of the model were expressed using quality adjusted life years (QALYs). The utility values used in the economic analysis were derived from EQ-5D-5L data collected from the pivotal trials and cross walked to EQ-5D-3L using a mapping algorithm presented in van Hout *et al*. 2012.[61] The values used in the base-case analysis were drawn from the JAIN and JAIN-like JAIY trial populations. In scenario analyses presented by the company, alternative utility values were also generated using data from the JAIN Europe and JAIN-only subgroups, as well as JAIN-like patients recruited to JAHL and JAHN.

To estimate utility values applied in the model, the company developed a multivariable risk equation to predict utility values according to response status. In line with the base case model structure, response was defined with respect to EASI50 and a four-point reduction in DLQI. Scenario analysis explored alternative response criteria including EASI50 alone and EASI75. The variables included in the regression analysis were response status, age, sex, visit time and baseline EASI score. Results of the regression analysis are presented in Table 90 of the CS.

Modelled utilities were based on the results of the regression analysis, with patients in the induction phase of the model along with non-responders assigned utility scores based on baseline utility (0.5979). For responders, utility scores were estimated by applying a coefficient of 0.1821 to baseline scores (0.5979 + 0.1821). An important feature of the applied utility values is that they are based on a within group analysis, rather than between group analysis of the available data. This results in the data for non-responders to treatment not being used in the model. The company did not model the impact of any treatment-related adverse events on quality of life, assuming that the impact of such events was accounted for within the trial utilities.

ERG comment

Impact of response status on HRQL

An important observation from the regression analysis is that the reported utility values for responders to treatment at Week 16 do not differ fundamentally from those classified as non-responders (change from baseline EQ-5D 0.1821 and 0.2042 respectively). The ERG queried this apparent anomaly at the clarification stage to confirm the ERG's understanding of the presented results, and to seek an explanation for why there is no apparent HRQL benefit associated with response. The company's response confirmed the ERG's interpretation and noted that this may be a consequence of the small sample size. The ERG, however, does not consider this a reasonable explanation as similar results are reported in a larger sample including all JAIN-like patients ([redacted]).

The significance of the results of the company's regression analysis cannot be understated, as they undermine the validity of the adopted model structure and the meaningfulness of the applied response criteria. In essence, it suggests that the company's definition of a meaningful improvement in symptoms confers little or no improvement in HRQL. In the context of the model this implies that there are no health gains from treatment as valued by EQ-5D. The ERG can only speculate as to the reasons for this apparent lack of correlation between response status and HRQL gains. One explanation may be that it is a product of the analysis methods used by the company and how the regression model classifies patients as responders; the ERG notes that comparatively few observations are from responders despite a reported response rate of circa 50% in the baricitinib arm. The, ERG, however, do note that the utility values reported in TA534[20] show a similar pattern and so this may simply reflect an issue with the response criteria selected.

Importantly, the application of a within group analysis to circumvent this otherwise inconvenient result is fundamentally flawed as a within group analysis will be heavily confounded by regression to the mean effects. Further, this approach is inconsistent with methods adopted in the appraisal of biologics for other inflammatory conditions such as psoriasis and psoriatic arthritis. In all of these appraisals a between group, comparative approach has been used.[54, 56] In the absence of alternatives, the ERG explores applying utility values generated using between group analysis as well as exploring

the use of values reported in TA534, which are based on using treatment specific utilities (see Section 6.1).

Source of HRQL values

The ERG notes that the utility values are based on data from JAIN and the JAIN-like JAIY patients, even though relevant HRQL data were collected for JAIN-like patients in both the JAHL and JAHM studies.

In response to queries raised by the ERG at the PFC stage, the company justified the decision to limit their analysis to the JAIN and the JAIN-like JAIY populations by noting that JAHL and JAHM considered monotherapy regimens and therefore did not provide evidence directly relevant to the decision problem. As such, the company considered that the JAIN and JAIY studies, which evaluated combination therapy, were the most representative of the modelled population. Scenario analysis presented by the company as part of the clarification response using data on all JAIN-like patients (JAIN + JAIN-like JAIY, JAHL, and JAHM) demonstrated an increase in the ICER for baricitinib relative to BSC, see Section 5.2 for results.

The ERG acknowledges that using JAIN and the JAIN-like JAIY patients is consistent with the efficacy data used in the model, but notes the company's preference for using a pooled data set including both baricitinib and placebo patients to generate utility values. Such an approach implies that the treatment received is a not factor in determining HRQL, and instead assumes that response status is perfectly correlated with HRQL. This is therefore inconsistent with the company's justification that JAHL and JAHM patients are not relevant to the modelled population because they received a monotherapy regimen. On this point, the ERG notes that no patients in the JAIN-like JAIY population received dupilumab, yet the company still considers the values relevant to this group of patients.

Categorisation of response and use of pooled utility values

A potential limitation of the model structure adopted by the company is that it does not distinguish between different levels of response. For example, in psoriasis TAs it is common to distinguish between patients achieving PASI75, PASI90, and PASI100[54, 56]. Such an approach allows for a more precise estimate of the QALY gains associated with a specific treatment as it better represents the health gains attributable to the magnitude of the response achieved. The ERG recognises that this is potentially more complicated to implement in the present model due to the requirement for patients to also achieve a DLQI ≥ 4, but does not consider this issue insurmountable. Further, such an approach is potentially important when treatment specific utilities are not being used, as there is greater potential to overlook the benefits of a treatment consistently inducing a higher level of response, and thus potentially impacting upon model outcomes. In this respect the ERG notes that treatment specific

utilities were used in TA534[20], this potentially justifies the use of a simpler 3 state model structure used in that appraisal. Treatment specific utilities could in principle be applied in the current model, but would be significantly more problematic to generate given the lack of head to head data comparing all three treatment options. The ERG therefore considers that a model structure based around different levels of response would have been the preferable approach, and would have better captured differences in HRQL achieved on different treatments.

Age adjustment

The ERG considers the application of age adjustment appropriate and in line with assumptions made in TA534. The ERG is, however, unfamiliar with the method used to apply such adjustments and is not fully clear on how the adjustment factors were generated based on the information provided in the revised CS. At the clarification stage the ERG requested the company provide further details of its methodology and justification for its approach. The company's response outlined that the adjustment factors were based on data from Ara and Brazier[62] to which a linear trend was fitted, but provided no further details. While the ERG welcomes this clarification, the precise methods used to generate the adjustment factors remain unclear. The ERG therefore cannot comment on the appropriateness of the applied adjustment factors.

4.2.7 Resources and costs

The company's model included drug acquisition and administration costs, concomitant treatment costs, costs associated with the treatment of flares, health state costs which account for the management and monitoring of patients with AD, and costs associated with treating adverse events.

4.2.7.1 Drug acquisition and administration costs

Baricitinib acquisition costs were sourced from Lilly and estimated based on a dose of a single 2mg or 4mg tablet per day. No administration costs were included for baricitinib. Baricitinib acquisition costs presented were inclusive of a confidential PAS discount of [redacted], as such all analyses presented by the company are inclusive of this discount.

Acquisition costs for dupilumab were sourced from MIMS and estimated based on a dose of a single 300 mg subcutaneous injection every two weeks. In line with the SmPC, the model allows for a loading dose of 600 mg in the first cycle.[33] Consistent with TA534[20], administration costs for dupilumab were included in the first cycle and account for the training of patients to self-administer.[20] This training was costed based on the cost of 30 minutes of patient contact with a Band 6 hospital-based nurse.[63] No further administration costs were included thereafter – implying all patients can successfully self-administer dupilumab after the first cycle. A confidential PAS discount is available for dupilumab. All analyses presented by the company are exclusive of this discount.

Table 44: Drug acquisition and administration costs for baricitinib and dupilumab (Source: Table 94 of the CS)

Treatment	Pack cost	Administration frequency	Cost per 4 week cycle	Administration costs	Total cost per cycle
Baricitinib 2 or 4mg	List price: £805.56 PAS Price: [redacted]	Daily	List price: £805.56 PAS Price: [redacted]	£0.00	[redacted]
Dupilumab 300mg	£1,264.89	Once every two weeks	First cycle: £1,897.35 Subsequent cycles: 1,264.90	£56.50	£1953.85 in the first cycle £1,264.90 thereafter.

PAS, patient access scheme.

Best supportive care (both 1st and 2nd line) was modelled as a blended comparator consisting of one of the following Mometasone (TCS), Tacrolimus (TCI) or prednisolone (corticosteroid). Unit costs were obtained from MMIS, with dosing informed by expert clinical advice. As a blended comparator, the company estimated the average cost of BSC according to the proportion of each treatment option used in practice. The company attributed these proportions to expert advice. No administration costs were associated with BSC. Table 45 describes the unit costs and per cycle costs associated with BSC.

Table 45 Drug acquisition costs for BSC (Adapted from Table 95 of the CS)

Treatment	Pack cost, £	Administration frequency	Administration costs	Proportion of use, %
Mometasone (class II TCS)	9.50	Daily	£0.00	66.70
TCI (Tacrolimus)	47.28	Twice per week	£0.00	22.20
Oral corticosteroids (Prednisolone)	1.48	Daily	£0.00	5.00
Total cost per cycle: £58.92*				

*Calculated by the ERG. **Abbreviations:** BSC: best supportive care; N/A: not applicable; TCI: topical calcineurin inhibitor TCS: topical corticosteroids.

ERG comment

Dosing of Dupilumab

In the model the costs of all treatments are calculated for the initial 16-week induction period and as annual cost applied in the maintenance period. In the model the costs of dupilumab during the induction period are estimated assuming 10 doses. This, however, incorrect, accounting for the loading dose there will only be 9 doses of dupilumab with doses given in weeks 0, 2, 4, 6, 8, 10,12, 14; the dose given week 16 will only be received by responders to treatment and is accounted for in the maintenance dose applied in subsequent cycles of the model. In Section 6.1, the ERG presents scenario analysis correcting this error.

Composition of BSC

The ERG is concerned about the company's approach to modelling BSC which is modelled both as a distinct comparator, while also assumed to be part of background supportive medications included as part of modelled health state costs. See Section 4.2.7.3 for details of health state costs included. This is problematic as it results in the model double counting the costs of providing BSC and is inconsistent with the approach adopted in TA534[20] where BSC costs are confined to health state costs. As a consequence of this approach, the model assumes that 122% of responders to BSC (1st line BSC) will receive Mometasone (TCS), while in non-responders (2nd line BSC) this rises to 166%. Similarly, Tacrolimus (TCI) is assumed to be received by 122% of non-responders (2nd line BSC), with a further 6% of patients assumed to receive tacrolimus as part of acute treatment (4 weeks only) received for symptom flares.

Related to the above, the ERG also questions the inclusion of prednisolone as part of BSC as well as part of the acute medications used to control symptom flares. Prednisolone and other corticosteroids are not frequently prescribed for extended periods of time due to the well documented side effects associated with long-term use, with NICE guidelines recommending the reservation of such medications for the intensive treatment of severe flares [13]. The application of long-term prednisolone costs even in a minority of patients is therefore inconsistent with UK clinical practice.

Because of these issues and for consistency with the approach taken in TA534[20], the ERG considers that BSC described in Table 45 should be removed from the model, such that the acquisition costs associated with BSC are confined those included as part of health state and flare treatment costs. The implication of this proposed adjustment for cost-effectiveness is considered in the ERG exploratory analyses presented in Section 6.1.

Costs and dosing of BSC

The ERG further notes that the reported dosing and frequency of administration reported in Table 96 of the CS are different to those implied when the same treatments are considered as part of concomitant treatments, see Table 46 for comparison. The CS does not outline any justification for this apparent disparity though it states the dosing and frequency of BSC were based on clinical advice.

Clinical advice received by the ERG highlighted the difficult of assigning an average dose for such treatments given the heterogeneity of the condition, with our clinical advisor considering both set of values plausible. The ERG, however, favours the dosing and frequency rates applied in concomitant treatments as these were drawn from TA534[20] and have been previously been accepted by the NICE committee. Given the limitations in the reporting, the ERG is not entirely clear on the company's intentions regarding the estimation costs for BSC and explores assumptions regarding the dosing and composition of BSC in Section 6.1.

Table 46: Comparison of dosing of BSC

	Dosing and frequency of administration BSC	Dosing and frequency of administration concomitant treatment
Mometasone 0.1% (TCS)	32g daily	16g daily
Tacrolimus 0.1% (TCI)	1.75g twice weekly	1.75g weekly
Prednisolone (Corticosteroid)	10mg daily	NA

Concomitant treatments

In addition to the drug acquisition costs described in Section 4.2.7.1, all patients were assumed to receive supportive care consisting of bathing products, emollients and background medications taken by patients with AD. The use of concomitant supportive care was assumed to vary in accordance with response status, with patients classified as responders assumed to use fewer concomitant treatments than non-responders. Acquisition costs of concomitant products were sourced from MIMS. Resource utilisation for specific elements of supportive care were based on TA534 and reported in Table 93 of the CS. In line with TA534[20] responders to treatment were assumed to use 50% fewer bathing products emollients and TCSs. Responders to treatment were assumed to not require any usage of TCIs.

ERG comment

Notwithstanding the issue raised above regarding the duplication of costs, the ERG notes two further issues with the modelling of concomitant treatments.

The first relates to the composition of concomitant treatments, and the inclusion of emollient bathing additives. While at the time of TA534[20] bathing additives were frequently used in practice, recent practice has changed following the publication of an NIHR funded HTA[64]. This study conducted an RCT considering the effectiveness and cost-effectiveness of bathing additives in children with atopic dermatitis and concluded that such products offer no benefit. Consequently, there has been a significant reduction in the use of these products in practice. Scenarios analysis is presented Section 6.1 explores removing bathing additives from the list of concomitant treatments.

A second issue relates to the assumed reduction in concomitant treatments for responders to treatment, which broadly speaking allow for 50% reduction in all concomitant treatments. While the ERG recognises that similar assumptions were accepted in TA534[20], the ERG notes that this assumption was informed largely by expert opinion with limited supporting evidence on TCS use from the CAFÉ study, and therefore was not directly informed by company data on concomitant treatment use. An examination of the literature identified in the resource review, however, reveals no alternative sources with which to inform these assumptions and the ERG was not able to identify any relevant

information in additional searches undertaken. Independent clinical advice sought by the ERG however, validated the assumed reductions as a reasonable reflection of their own experience in using dupilumab and they considered it reasonable to assume similar reduction in patients responding to baricitinib.

4.2.7.2 Treatment of Flares

To account for the relapsing-remitting nature of AD, the economic model includes costs associated with the acute treatment of flares. Acute medications used included TCSs (potent and very potent), TCIs and systemic steroids. Acute medications used in the treatment of flares were assumed to be used for a period of four weeks, with costs drawn from MIMS. Rates of acute medications were based on data used in TA534[20] which utilised data from the long-term follow-up study CHRONOS. Rates applied were not linked to response, but instead assumed to vary according to whether a patient received biologic treatment or BSC; higher rates were assumed in patients receiving BSC. Annual flare rate applied in the model are reported in Table 97 of CS, and the estimates for flare treatment costs are reported in Table 98 of the CS. The modelled flare rate during the induction period was assumed to be zero for baricitinib and dupilumab, while patients still experienced flares and incurred associated costs on BSC.

ERG comment

As outlined in Section 4.2.5.4, the ERG has some concerns regarding the rate of flares applied in the model and specifically that they may underrepresent the frequency of flares. The ERG also notes that the assumption that patients do not experience flares during the induction period on baricitinib to be contrary to evidence from the JAIN trial. With regards to the composition and costs of flare treatments used in the model, the ERG clinical advisor is relatively satisfied that the listed treatments are reflective of practice, they however noted that less frequently, a short course of immunosuppressants may also be used in the treatment of more serious flare ups. The impact of this omission, is however, likely to be minimal given the nominal costs associated with immunosuppressant drugs.

4.2.7.3 Health state costs: disease management costs

Disease management and monitoring costs identified by the company as supportive of the condition were: dermatologist consultant consultations, dermatologist nurse visits, GP consultations, emergency department visits, hospitalisations, day-case hospital visits and full blood counts. Additionally, in patients classified as non-responders, costs associated with supportive phototherapy and psychological support were accounted for. Unit costs were obtained from the most recent NHS reference cost schedule[65] and the Personal Social Services Research Unit (PSSRU) handbook.[66] The rates of consumption of these resources were sourced from TA534[20] where costs were principally informed by a retrospective review of retrospective a cohort of 60 UK patients with uncontrolled AD recruited

from 6 secondary and tertiary centres. In line with TA534[20], resource use varied in accordance with response status with non-responders assumed to incur greater disease management and monitoring costs.

Unit costs associated with the management of AD are described in Table 99 of the CS and model cycle consumption rates for responders and non-responders are reported respectively in Tables 100 and 101 of the CS.

ERG comment

Alignment with TA534[20]

The values used in non-responders appear to reflect resource use assumptions adopted in TA534, and the ERG is satisfied that these are likely to represent the most relevant data to populate these inputs, as the resource review identified no alternatives. The ERG, however, highlights concerns previously raised in TA534, namely that that the dataset used to generate these values is small (based on 60 patients) and that there is some uncertainty regarding the generalisability of these data to the modelled population. The population recruited to the resource study is simply described as having uncontrolled AD.

In contrast with the non-responder values, the resource use estimates applied to responders differ to those applied in TA534[20], with greater resource use assumed across several components of health state costs, see Table 47. As such, response leads to a smaller reduction in health state costs than was assumed in TA534[20]. The implication of these alternative assumptions is that less effective treatments are favoured, meaning baricitinib would be favoured over dupilumab, and BSC over baricitinib.

Table 47: Comparison of resource rates for responders to biologics

	Baricitinib Model		Dupilumab model	
	1st year*	2nd year onwards	1st year	2nd year onwards
Dermatologist outpatient consultation (consultant led)	4.98	4.30	4	2
Dermatologist nurse visit	0.35	0.35	0.42	0.42
GP consultation	6.20	6.20	2.00	2.00
Accident & Emergency visit	0.02	0.02	0.02	0.02
Hospitalisation	0.02	0.02	0.02	0.02
Day case	0	0.00	0.00	0.00
Full blood count (FBC)	0	0.00	NA	NA

*These values have been estimated by the ERG to allow a meaningful comparison with the values reported in TA534.

Committee preferences regarding resource use in TA534[20] are unclear, as they were not well documented in the FAD. However, it is apparent from the ACD that some concerns were raised regarding the magnitude of assumed reduction in resource use, with the ERG exploring several

alternative scenarios. The more conservative values adopted by the company arguably therefore address these concerns. As such, the ERG is satisfied that the values align with the spirit of those previously accepted in TA534[20] and consider them are broadly acceptable given the paucity of alternative data sources.

4.2.7.4 Adverse events

Adverse events modelled included injection site reactions, conjunctivitis (allergic and infectious), oral herpes and upper repository infections. The rates of AEs were drawn from the pivotal trials – see Sections 3.2.5 and 4.2.5.5 for details. The costs associated with AEs were drawn from NHS reference costs and PSSRU and are reported in Table 102 of the CS.

ERG comment

The ERG is satisfied with the unit costs applied in respect to AE. For critique of AE rates applied see Section 4.2.5.5.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The list price of a 28-tablet pack of 2 or 4mg baricitinib is £805.56, resulting in an approximate annual price of £10,508.24. Baricitinib is currently licensed for patients with rheumatoid arthritis and has a PAS discount of [redacted], reducing the cost to £[redacted] per pack; around £[redacted] a year. If baricitinib is recommended for use for patients with AD, a revised PAS discount of [redacted]% will be used. This reduces the price of a 28-tablet pack of 2 or 4mg baricitinib to £[redacted], with an average annual cost of £[redacted]. A confidential PAS discount is also available for dupilumab. All results presented below are exclusive of this discount, and are presented in a confidential appendix generated by the ERG.

5.1.1 Base Case Results

The results of the base case cost-effectiveness analysis are summarised in Table 48. Compared with BSC, the results suggest baricitinib is associated with increased costs (cost difference of £[redacted]) but an improved quality of life (QALY difference of [redacted]). The company's base case ICER comparing baricitinib with BSC is £17,941 per QALY gained.

When comparing baricitinib with dupilumab, baricitinib is associated with lower costs compared to dupilumab (incremental cost of -£[redacted]), but also generates fewer QALYs (is less effective; incremental QALYs of [redacted]). The company's base case comparing baricitinib and dupilumab is therefore in the southwest quadrant and the reported £[redacted] ICER represents the cost of QALYs forgone i.e. the additional QALYs generated by dupilumab would cost £[redacted] per QALY gained.

Overall, the results suggest baricitinib is the cost-effective treatment option assuming a WTP threshold of £20,000 per QALY.

Table 48. Company base case results: baricitinib vs BSC and baricitinib vs dupilumab

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,941	£17,941
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,525[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

The ERG requested clarification on one parameter in the model that affected the company's base-case results at the clarification stage, relating to the conditional response rates applied for the BSC arm. In response, the company provided a corrected version of the model, which revised the conditional

response rate between weeks 16 and 52 for BSC. The updated incremental results for the base case are presented in Table 49. This update results in very slight change to the base-case less cost-effective compared to both BSC.

Table 49. Company base case results: baricitinib vs BSC and baricitinib vs dupilumab. Updated EASI50+DLQI ≥4 conditional response.

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,996	£17,996
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£89,048	£203,968[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

The CS also present the base case results as the net monetary benefit (NMB), presenting the monetary value of baricitinib to provide a single unit of benefit at a given WTP threshold. The results are presented in Figure 41 of the CS.

Figure 8.

[redacted]

5.1.2 Company Scenario Analysis

The company presented a scenario analysis where secondary censoring of patients was used to explore an alternative definition of non-responder. Secondary censoring censors patients as non-responders after permanent study drug discontinuation or after the initiation of systemic rescue

therapies. In this scenario, response is defined as EASI50 and ΔDLQI≥4. The incremental cost-effectiveness ratio for this scenario is presented in Table 50.

Table 50. Company scenario analysis: baricitinib vs BSC and baricitinib vs dupilumab using the secondary censoring rule (CS Table 109, Scenario 2a).

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,736	£13,736
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£68,392	£192,238[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

5.1.3 Probabilistic Sensitivity Analysis

The company performed a probabilistic sensitivity analysis (PSA), running 3,000 iterations for the pairwise and fully incremental comparisons (Table 51). The mean probabilistic ICER for baricitinib compared to best supportive care was £17,853/QALY. When compared to dupilumab, the mean probabilistic ICER equalled £199,001/QALY foregone. Figure 9 presents the cost-effectiveness plane and cost-effectiveness acceptability curve from the probabilistic sensitivity analysis.

Table 51. Company Probabilistic Sensitivity Analysis base case results: baricitinib vs BSC and baricitinib vs dupilumab (Source: CS Table 107)

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,853	£17,853
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,866	£199,001[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

Figure 9.

The ERG requested clarification on the methods used to calculate the standard errors in the PSA, as all standard errors included in the PSA were 10% of the mean. The company provided a PSA with updated utility SE values, which were based on the output of the MMRM regression models. An error in the SE corresponding to the response rates for EASI50 + ΔDLQI ≥ 4 in dupilumab was also corrected. The updated PSA results, generated using 3,000 iterations are presented in Table 52.

Table 52. Updated Probabilistic Sensitivity Analysis base case results: baricitinib vs BSC and baricitinib vs dupilumab

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,965	£17,965
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£89,879	£208,938[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

Overall, the change in the SE associated with response rates of dupilumab and the utility values made a minimal difference to the ICER. The probability of cost-effectiveness at the willingness to pay threshold at £20,000 and £30,000/QALY is presented in Table 53. There is a [redacted]% probability that baricitinib is the most cost-effective treatment at a £20,000/QALY threshold and a [redacted]% probability at £30,000/QALY threshold.

Table 53. Probability of cost-effectiveness at a £20,000 and £30,000/QALY WTP threshold (Source: Table 106 of CS).

	WTP threshold £20,000	WTP threshold £30,000
BSC	[redacted]	[redacted]
Baricitinib	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]

5.1.4 Deterministic Sensitivity Analysis

The company performed deterministic sensitivity analysis (DSA) on ten of the most influential variables in the analysis of baricitinib vs dupilumab and baricitinib vs BSC. The DSA for the pairwise comparison of baricitinib and dupilumab is presented in Figure 10 and suggests that discount rates, efficacy of baricitinib and the pack cost of dupilumab are the most influential parameters. Figure 11 presents the results of DSA for the pairwise comparison of baricitinib and BSC. The most influential parameters were the discount rates, the utility value assigned to the induction/non-responder health state and several elements of health state costs attributed to non-responders.

Figure 10.

[redacted]

Figure 11.

5.2 Additional sensitivity analyses

The ERG requested several scenario analyses at the clarification stage. The company's results and impact on the ICER for each scenario are presented below.

Health Related Quality of Life

The ERG requested the use of a pooled utility for JAIN-like patients across all trials, rather than using the smaller population of JAIN and JAIN-like JAIY patients alone. The resulting utilities from this imply a smaller HRQL benefit from response to treatment. The impact of this change is to favour less efficacious treatments and results in both baricitinib and dupilumab generating fewer incremental QALYs compared with BSC. The results of this analysis are presented in Table 54.

Table 54. Company scenario analysis of pooled utility values based on all JAIN-like patients (Source: Company Response to PFCs, Table 19),

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£25,092	£25,092
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£124,256	£284,654[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

In addition to the above, the ERG requested a scenario in which the BSC arm showed a sustained HRQL response. The company presented two scenarios in response, one where a lower

discontinuation rate is assumed for BSC, while also assuming a loss of HRQL for all treatments over time, and a second where a lower discontinuation rate is assumed for BSC, while holding HRQL values constant for all treatments. The results of these scenarios are respectively presented in Table 55 and Table 56 and both result in the increased ICER for baricitinib compared with BSC. This is because patients receiving BSC transition to the 2nd line BSC (non-response) health state more slowly and consequently this results in reduced costs and improved quality of life for patients in the BSC arm of the model.

Table 55. Company's scenario analysis with lower discontinuation rate for BSC and loss of utility applied over time for all treatments. (Source: Company response to PFCs, Table 21)

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£20,005	£20,005
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£96,267	£220,020[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

Table 56. Company's scenario analysis with lower discontinuation rate for BSC and there is no loss of utility applied over time for all treatments. (Source: Company response to PFCs, Table 22)

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£20,475	£20,475
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£98,162	£222,989[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

Adverse Events

In response to the ERG's concerns regarding the discrepancies in the annual adverse events reported in the original submission, the company provided three revised scenario analyses applying updated AE probabilities. The first scenario applied revised rates of TAEs from the JAIN and JAIN-like JAIY trial patients to include patients who experienced allergic conjunctivitis, in line with the incidence reported in Table 74 of the CS. Overall, there was a minimal difference between the updated and original base case ICER.

Table 57. Company revised scenario analysis applying AE probabilities seen in the trial. (Source: Company response to PFCs, Table 26)

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,897	£17,897
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,596[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

A second scenario analysis was conducted to correct for differences in the frequency of allergic conjunctivitis, infectious conjunctivitis and oral herpes between baricitinib (combination) and best supportive care (given that BSC is included in the baricitinib combination therapy). This was achieved by applying the relative risk of adverse events observed in the comparison of BSC and dupilumab (taken from TA534) to the baricitinib population. The results of the second scenario analysis are reported in Table 58 and show a very slight increase in the ICER. This is a result of an increase in the frequency of AEs associated with baricitinib.

Table 58. Company's scenario analysis using relative risk estimates from TA534 to determine AE frequency rate for baricitinib. (Source: Company response to PFCs, Table 28)

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,948	£17,948
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,513[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

Finally, a third scenario was presented where the adverse events used in the model are based on the long-term follow-up study (JAHN). The result of this analysis shows a small reduction in the ICER compared with the company base-case (see Table 59).

Table 59. Company's scenario analysis where adverse event rates are based on JAHN. (Source: Company response to PFCs, Table 30)

	Total Costs	Total QALYs	Incremental Costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)[a]
BSC	[redacted]	[redacted]	-	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,880	£17,880
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£88,842	£203,622[b]

[a] ICER of baricitinib vs comparator, [b] represents cost per QALY forgone

5.3 Model validation and face validity check

Validation undertaken by the company

The company state that they were unable to complete a clinical validation of the adopted model structure and assumptions due to the COVID-19 pandemic.

The internal validity of the model was checked by an independent third party who undertook a technical validation of the model. This included working through two separate technical and stress test checklists, and a review of all model calculations including standalone formulae, equations and Excel macros programmed in VBA.

Validation undertaken by the ERG

The ERG was unable to implement a full clinical validation of the economic model in lieu of such a validation by the company. The ERG was, however, able to consult with a clinical advisor to review many of the model assumptions and the general approach adopted by the company. As part of the ERG's critique, the ERG also compared the model assumptions to previous economic analyses of AD, particular those adopted in TA534[20] which addressed a similar decision problem.

As part of the ERG's assessment of the economic analysis, the ERG checked the internal validity of the model and considered the face validity of the model's predictions. This included a series of model calculation checks, including pressure tests and formula auditing. The ERG also completed the Drummond quality assessment checklist (Appendix 2) The ERG felt that the executable model was in general well presented, but contained a degree of redundancy, in that it contained calculations that did not contribute to model function. The heavy reliance on macros and code embedded into several sheets also made editing the model substantially more complicated.

Several minor model errors were identified as part of the ERG's validation checks. The most important of these errors concerned the discontinuation rates applied for second and third-line treatments. The ERG also identified a number of inconsistencies in the values used to model the rate of discontinuation for BSC. The company was able confirm that these were typographical errors at the clarification stage and supplied a revised corrected model. All identified errors were corrected by the ERG, and a revised model supplied to the company with altered cells highlighted to aid verification. These corrections did not impact substantively on the model's predictions. Revised results are presented in Section 6.

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

In Section 4.2.3, the ERG concluded that treatment sequences were an important omission from the company's economic analysis and considered the company's primary analysis to be potentially misleading and unsuitable for decision making, given that it does not permit sequential use of baricitinib and dupilumab. Implementation of treatment sequences requires several assumptions to be made regarding the effectiveness of the treatments as the available trial data for baricitinib cannot provide separate estimates of relative effectiveness for dupilumab naïve and dupilumab experienced patients. Similarly, there is a lack of data for dupilumab in baricitinib naïve and baricitinib experienced patients. The analysis presented therefore makes the simplifying assumption that effectiveness of baricitinib and dupilumab is the same regardless of the position in the treatment pathway. The ERG acknowledges that this assumption limits the value of this analysis, but considers that it is important to evaluate treatment sequences including both baricitinib and dupilumab given the company's position of baricitinib as a replacement to dupilumab.

To aid in the interpretation of the sequence analysis the ERG also presents the incremental net monetary benefit (NMB = λ x $\Delta E - \Delta C$) for each comparator versus BSC, at a £20,000 and £30,000 per QALY threshold. BSC is used as baseline as it is the cheapest treatment sequence in all scenarios. Using this approach if an intervention has an incremental net monetary benefit (NMB) > 0, then it would be considered more cost-effective than BSC. Further, the incremental net-benefit of each treatment versus BSC can be used as a basis for establishing the most cost-effective treatment sequence, and allows the ranking by cost-effectiveness without estimating fully incremental ICERs. The application of the net-benefit also has the specific advantage that it provides an unambiguous decision rule i.e. where NMB > 0 implies a technology is cost-effective relative to BSC and avoids complications created by negative ICERs. This approach is taken in all subsequent scenario analyses, with treatment sequences including both baricitinib and dupilumab evaluated on the company base case assumptions as well as the ERG's preferred base-case.

All results present in this Section are inclusive of PAS discount for baricitinib, but exclude the confidential PAS discount for dupilumab. Results inclusive of the dupilumab PAS are presented in a confidential appendix which also includes a brief discussion of how the PAS impacts on the interpreting the analysis presented by the ERG.

6.1 Exploratory and sensitivity analyses undertaken by the ERG

1) Removal of Discontinuation rates from first-line BSC

As discussed in Section 4.2.1, the ERG does not consider the company's approach to modelling BSC to be appropriate as it does not account for the waxing waning nature of AD. The ERG, therefore

proposes an alternative approach to the modelling BSC where the distinction between first- and second-line BSC is removed and all patients on BSC are assumed to receive utilities and costs corresponding to the average of both responders and non-responders. To reflect this alternative approach in the cost-effectiveness model provided by the company, a scenario is presented where the discontinuation rates for BSC have been changed to 0%, the initial placebo response observed in the JAIN and JAIN-like JAIY population is therefore attributed to regression to the mean and the observed response rates are assumed to be reasonable indicator of the proportion of patients on BSC who will have good disease control at any one time.

2) Using relative effectiveness from the ITC to model absolute response

In the model, the company calculated response for dupilumab and baricitinib by adding the absolute risk difference between placebo and each active treatment to the placebo response rate the observed in JAIN and JAIN-like JAIY population. As described in Section 4.2.5.1 this approach can result in bias. The company's approach also deviates from the recommendations outlined in NICE DSU TSD5,[3] which recommends that absolute response rates for the control arm are pooled with relative effects estimated for the active treatments vs control on the log-odds scale. In addition, different studies can be used to estimate the absolute placebo response rate and rates can also be calculated using different censoring rules.

To explore this, the ERG present four scenarios to evaluate the use of relative treatment effects to obtain absolute response probabilities for each treatment. These scenarios differ based on the type of censoring used, and the source of data used to model the response rate seen in the placebo arm. The four scenarios modelled are:

a) Primary censoring, using rate of response in the JAIN-like patients in placebo arm.
b) Primary censoring, using rate of response in the JAIN-like and CAFÉ-like patients in placebo arm.
c) Secondary censoring, using rate of response in the JAIN-like patients in placebo arm.
d) Secondary censoring, using rate of response in the JAIN-like and CAFÉ-like patients in placebo arm.

The absolute probabilities of response on placebo were pooled using Bayesian fixed-effect meta-analysis. The relative treatment effects (log-odds ratios and their standard errors) for baricitinib and dupilumab compared to placebo were then added to the log-odds of response on placebo, to obtain means and standard errors for the absolute probabilities of response on each intervention. Further details and OpenBUGS code are given in Appendix 4. These were then included in the model which assumes they follow a Beta distribution. The response rates estimated by the ERG and applied in the model are reported in Table 60.

Table 60 Absolute response rates applied

	Response probability, % (SE%)				
	Company base case	**Scenario 2a)**	**Scenario 2b)**	**Scenario 2c)**	**Scenario 2d)**
BSC	31.25 (3.86)	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinb	48.99 (4.09)	[redacted]	[redacted]	[redacted]	[redacted]
Dupilumab	79.25 (3.00)	[redacted]	[redacted]	[redacted]	[redacted]

3) Conditional response between 16-52 weeks based on JAHN.

In the model presented by the company, conditional response at week 52 (that is the probability of response at 52 weeks if an individual showed a response at 16 weeks) is used to model the rate of discontinuation with values for both baricitinib and dupilumab, based on rates for dupilumab reported in TA534. The ERG does not consider the assumption of equivalence of discontinuation rates to be appropriate and it is likely to overestimate the response to baricitinib. Further, the ERG considers the use of conditional response to model discontinuation rates undesirable, because it does not include other reasons for discontinuation, such AEs, and implies a formal stopping rule unlikely to be applied in practice. The ERG therefore considers three scenarios exploring alternative discontinuation rates.

In the first two scenarios a) and b) conditional rates for baricitinib are sourced from the JAHN extension study, as this represents the only source of long-term effectiveness data for baricitinib. The ERG is aware that the JAHN extension study is based on the use of baricitinib as a monotherapy but given the absence of the requisite data for combination therapy, deem it to be more accurate than using data for dupilumab. In scenario a) post week 52 rates of discontinuation for baricitinib continue to be based on dupilumab values, while in scenario b) post week 52 rates of discontinuation for baricitinib are based on all cause discontinuation rates for responders from JAHN (classified by IGA score rather than EASI50 and DLQI ≥ 4 as the latter was not available). In both scenarios dupilumab discontinuation rates are left unchanged.

In scenario c), the ERG's preferred scenario, rates of discontinuation for both the week 16 to 52 and post week 52 periods are based on relevant all-cause discontinuation rates. For baricitinib, JAHN data are used, while dupilumab rates are drawn from the CHRONOS study. The discontinuation rates used in company's base-case and the two scenarios are presented in Table 61.

Table 61. Per cycle discontinuation rates applied in the company's base-case, and the ERG's exploratory scenario analyses

	Baricitinib		Dupilumab	
	16-52 weeks	Post 52 weeks	16-52 weeks	Post 52 weeks
Company base-case	0.7%	0.29%	0.7%	0.29%
Scenario 3 a)	[redacted]	0.29%	0.7%	0.29%
Scenario 3 b)	[redacted]	[redacted]	0.7%	0.29%
Scenario 3 c)	[redacted]	[redacted]	0.29%	0.29%

4) Flare rates for baricitinib set to BSC

In their base-case analysis, the company assume that the patients receiving baricitinib and dupilumab have fewer flares than patients on BSC, with the same rates applied to both baricitinib and dupilumab. Evidence from the JAIN study, however, does not support such an improvement in flare control with baricitinib. The ERG therefore explores a scenario where the flare rates for baricitinib are set equivalent to the BCS group. The flare rates applied in this scenario are reported in Table 62.

Table 62. Modelled flare rates for baricitinib for the company base case and ERG scenario analysis based on flare rates of dupilumab and BSC respectively.

	Company base-case: flare rates equivalent to dupilumab		ERG Scenario Analysis: flare rates equivalent to BSC	
	Induction	Annual	Induction	Annual
TCS (Potent)	2%	8%	13%	42%
TCS (Very potent)	1%	4%	6%	21%
Systemic steroid	2%	5%	3%	10%
TCI	0%	0%	1%	5%

5) Alternative utility values

As noted in Section 4.2.6, an important feature of the regression analysis used to generate the utility values used in the model is that responders to treatment at Week 16 do not differ fundamentally from those classified as non-responders. The company attempt to avoid this issue by applying utility values based on a within group analysis of the available HRQL, however, as noted in Section 4.2.6, such an approach will be heavily confounded by regression to the mean effects. In an attempt to explore the impacts of this issue and to provide more meaningful and appropriate utility values, the ERG conducts two scenario analysis exploring the use of alternative utility values. In scenario a) utility value are drawn from HRQL data from the JAIN and JAIN-like JAIY patients and modelled considering a more appropriate comparative analysis. This scenario intends to illustrate the issues with the values provided and how they serve to undermine the model structure used by the company. In scenario b) a

more realistic set of values is used based on those reported in TA534.[20] In this scenario, treatment specific utilities are applied such that patients on maintenance baricitinib and dupilumab are assigned the reported utility of responders to dupilumab. Patients on BSC, including patients classified as non-responders are assigned a single utility value based on the average of all placebo patients at week 16. In contrast to TA534, no waning of utility was modelled for BSC. The utility values applied are reported in Table 63.

Table 63 Modelled utility values for the company base case and ERG scenario analysis

	Company base case		Scenario 5a)		Scenario 5b)	
	Baricitinib/ Dupilumab	**BSC**	**Baricitinib/ Dupilumab**	**BSC**	**Baricitinib/ Dupilumab**	**BSC**
Induction	0.5979	0.5979	0.5979	0.5979	0.66	0.66
Responders/ Maintenance	0.7800	0.7800	0.7800	0.7800	0.898	0.797
Non-responders	0.5979	0.5979	0.8021	0.8021	0.797	0.797

6) Removing drug acquisition costs for Best Supportive Care

As discussed in 4.2.7.1 the ERG has several concerns regarding the costing of BSC, noting that the assumed dosing of several elements of BSC differ depending upon whether they are as part of usual care or as a concomitant treatment. The ERG also considers that the modelling of BSC as a distinct comparator, while also including additional costs as part of health state costs leads to the double-counting of costs associated with BSC. The ERG therefore implements two scenarios. In scenario a) the ERG retains the treatment costs associated with BSC but modifies the dosing so that it aligns with the dosing regimens applied when the same therapies are considered as part of BSC. This reduces the costs of BSC from £14.73 to £7.92 per week. In scenario b), the ERG's preferred scenario, all treatment costs associated with BSC are removed and included only as part of modelled concomitant treatments. This latter approach aligns with the assumptions made in TA534[20] and is consistent with the concept that BSC is not a treatment aimed at achieving disease control in the same way as either dupilumab or baricitinib, but rather the treatment of choice to manage disease-related symptoms when disease control cannot be achieved with existing treatment.

7) Removal of Costs Associated with Bathing Products

The ERG considers the inclusion of costs associated with bathing products to be inappropriate, given recent evidence that these are of limited benefit resulting a considerable reduction in the use of bathing products in practice. The ERG therefore considers a scenario where the resource use rates for bathing products are set to zero.

8) Additional monitoring for patients receiving baricitinib

The company base-case assumes that costs of monitoring and management are the same for baricitinib and dupilumab. However, as discussed in Section 4.2.5.5, there may be additional monitoring required for patients receiving baricitinib that were not included in the model. To account for this additional monitoring, the ERG have changed the frequency of full blood count tests to align with assumptions made regarding BSC where 4 blood tests per annum are assumed.

9) Number of Dupilumab Injections

The ERG have corrected an apparent error in the model regarding the number of dupilumab injections provided in the first 16 weeks of treatment. According to the EMA product profile, cited by the company in the model, the administration of dupilumab is given as double 600 mg dose in week 0 followed by bi-weekly 300 mg at weeks 2, 4, 6, 8, 10, 12 and 14. Therefore, 9 dupilumab 300 mg doses are required in the first 16-weeks, rather than the 10 injections modelled.

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

A summary of the ERG's exploratory analyses of the cost-effectiveness of baricitinib for the treatment of atopic dermatitis is presented in Table 64. The results presented are inclusive of the PAS available for baricitinib, but exclude the PAS discount for dupilumab. Results including the PAS discount for dupilumab are presented in the confidential appendix.

Table 64. Exploratory analyses performed by the ERG.

Analysis	Intervention	Discounted Costs	Discounted QALYS	Fully incremental ICER	Change from Base Case	Net Monetary Benefit	
						£20,000 WTP threshold	£30,000 WTP threshold
Company Base Case (revised at the clarification stage)	BSC	[redacted]	[redacted]	[redacted]	-	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	-	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£203,968	-	[redacted]	[redacted]
ERG Correction of Model Errors	BSC	[redacted]	[redacted]	-		-	-
	Baricitinib	[redacted]	[redacted]	£18,003	+£7	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£204,046	+£78	[redacted]	[redacted]
Treatment Sequencing	BSC	[redacted]	[redacted]		-	-	-
	Baricitinib	[redacted]	[redacted]	£18,003	NA	[redacted]	[redacted]
	Baricitinib + Dupilumab	[redacted]	[redacted]	£90,446	NA	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	Dominated	NA	[redacted]	[redacted]
	Dupilumab + Baricitinib	[redacted]	[redacted]	£3,597,452*	NA	[redacted]	[redacted]
Scenario 1: No discontinuation on first-line BSC	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£29,595	+£11,599	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£291,428	+£87,460	[redacted]	[redacted]
Scenario 2: a) Relative Effect Primary Censoring JAIN-like population	BSC	[redacted]	[redacted]	-		-	-
	Baricitinib	[redacted]	[redacted]	£18,009	£13	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£205,062	£1094	[redacted]	[redacted]
b) Relative Effect Primary Censoring JAIN/ CAFÉ population	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£17,959	-£37	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£253,917	£49,949	[redacted]	[redacted]
c) Relative Effect	BSC	[redacted]	[redacted]	-	-	-	-

Secondary Censoring JAIN-like population	Baricitinib	[redacted]	[redacted]	£18,046	£50	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£182,592	-£21,376	[redacted]	[redacted]
d) Relative Effect Secondary Censoring JAIN/ CAFÉ population	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£17,954	-£42	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£237,490	£33,522	[redacted]	[redacted]
Scenario 3: a) Conditional Response JAHN	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£18,413	£447	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£144,144	-£59,824	[redacted]	[redacted]
b) Conditional Response and Discontinuation Rates JAHN	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£21,465	£3,499	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£98,746	-£105,222	[redacted]	[redacted]
c) JAHN discontinuation rates	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£20,543	£2,577	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£100,909	-£103,059	[redacted]	[redacted]
Scenario 4: Flare rates for baricitinib based on BSC	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£18,070	£74	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£203,938	-£30	[redacted]	[redacted]
Scenario 5: a) Utilities based on Company's regression analysis	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	Dominated	-	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	Dominated	-	[redacted]	[redacted]
b) Utilities based on dupilumab (TA534) submission.	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£33,451	£15,455	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£352,831	£148,863	[redacted]	[redacted]
Scenario 6 a) BSC drug costs amended	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£20,223	£2,257	[redacted]	[redacted]

	Dupilumab	[redacted]	[redacted]	£206,159	£2,191	[redacted]	[redacted]
b) BSC drug costs not included	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£22,808	£4,812	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£208,619	£4,651	[redacted]	[redacted]
Scenario 7: Remove the costs of bathing products	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£17,330	-£636	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£203,407	-£561	[redacted]	[redacted]
Scenario 8: Monitoring costs for baricitinib	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£18,078	£112	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£203,925	-£43	[redacted]	[redacted]
Scenario 9: Correction of number of dupilumab injections	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£18,003	£37	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	£203,056	-£912	[redacted]	[redacted]

*This ICER is estimated relative to the sequence baricitinib + dupilumab as the sequence including dupilumab alone is strongly dominated.

6.3 ERG's preferred assumptions

The ERG's base case applies several of the exploratory analyses performed. This includes Scenarios 1, 2d, 3c, 4, 5b, 6b, 7, 8, and 9. The results of the ERG base case are presented in Table 65. The ERG's base case ICER for baricitinib compared to BSC is £64,710/QALY.

Table 65. ERG base case analysis

Analysis	Intervention	Discounted Costs	Discounted QALYS	Fully incremental ICER	Net Monetary Benefit	
					£20,000 WTP threshold	£30,000 WTP threshold
ERG Base Case: Sequencing & Scenarios 1, 2d, 3c, 4, 5b, 6b, 7, 8 and 9	BSC	[redacted]	[redacted]			
	Baricitinib	[redacted]	[redacted]	£64,710	[redacted]	[redacted]
	Baricitinib + Dupilumab	[redacted]	[redacted]	£174,071	[redacted]	[redacted]
	Dupilumab	[redacted]	[redacted]	Dominated	[redacted]	[redacted]
	Dupilumab + Baricitinib*	[redacted]	[redacted]	£334,999	[redacted]	[redacted]

*This ICER is estimated relative to the sequence baricitinib + dupilumab as the sequence including dupilumab alone is strongly dominated.

Two ERG analyses resulted in a considerable increase in the company's base case ICER between baricitinib and BSC: removing discontinuation to second-line BSC increased the ICER of baricitinib compared to BSC by £11,599/QALY; the use of the dupilumab utilities resulted in an increase of £15,455/QALY for baricitinib compared to BSC.

Three ERG analyses resulted in considerable changes in the company's base case ICER between baricitinib and dupilumab. Removing discontinuation to second-line BSC increased the ICER (in this case, represented by cost/QALYs forgone) by £87,460/QALY; the use of dupilumab utilities also increased the ICER by £148,863/QALY. However, the use of JAHN discontinuation rates resulted in a reduction in the ICER by £104,476. As the PAS discount rates for dupilumab have not been included, the ICERs between baricitinib and dupilumab will reduce.

The ERG performed exploratory analyses where treatment sequences are modelled. In the treatment sequence where baricitinib is provided as a first-line therapy, followed by dupilumab, the ICER is £162,953/QALY. A treatment sequence where dupilumab is used as a first-line therapy followed by baricitinib, results in an ICER of £166,751. These ICERs will reduce when the PAS discount rates for dupilumab have been applied.

6.4 Conclusions of the cost effectiveness section

6.4.1 Summary of company's cost-effectiveness evidence

The company performed a targeted literature review to identify cost-effectiveness evaluations of systemic treatments for people with AD. No prior economic evaluations of baricitinib were identified in the review, but several relevant studies were identified for other treatments including dupilumab, the principal comparator in the company's economic analysis. The studies identified included economic evaluations carried out as part of the NICE appraisal of dupilumab,[20] as well as the Institute for Clinical and Economic Review report[37] which considered the cost effectiveness of dupilumab. The company states that both studies were used to inform inputs and assumptions in the company's economic analysis.

The company developed a *de novo* economic analysis to appraise the cost and benefits of baricitinib combination treatment in adult patients with moderate-to-severe AD who have failed at least one systemic immunosuppressant due to intolerance, contraindication, or inadequate disease control. This is a subgroup of the population covered by the anticipated marketing authorisation. The comparators considered were dupilumab combination therapy and BSC. The model structure developed was similar to that used in previous NICE appraisal of dupilumab and comprised a set of tunnel states representing the 16-week induction period, followed by a Markov model representing the remainder of the patient's life.

The company's base-case economic analysis suggested that baricitinib is more costly, but also more effective than BSC. The company's base-case ICER comparing baricitinib with BSC is £17,941 per QALY gained. When compared with dupilumab, baricitinib is associated with lower costs, but was also less effective. The company's base-case ICER for this comparison was in the south-west quadrant, at £203,525 per QALY forgone, i.e. additional QALYs generated by dupilumab cost £203,525 per QALY. Assuming a £20,000 per QALY willingness to pay threshold, baricitinib is the most cost-effective treatment. Note that these results are inclusive of a patient access scheme discount (PAS) for baricitinib, but exclusive of the PAS for dupilumab. At the £20,000 per QALY threshold, probabilistic analysis suggests a [redacted] probability that baricitinib is the most cost-effective treatment versus dupilumab, and at a threshold £30,000 this decreases to [redacted].

6.4.2 Conclusions of ERG's Critique

The ERG identified substantive structural uncertainties associated with the company's approach that potentially limit the reliability of the company's analysis and bring into question its suitability for decision making. These included concerns regarding the use of EASI scores and DQLI as indicator of response. While consistent with the previous appraisals in atopic dermatitis, it is not clear that this

composite outcome would be recognised or treated as clinically meaningful in practice. The ERG notes that the primary trial outcome was EASI75, which was also considered a clinically significant improvement by the British Association of Dermatologists in their submission. Importantly, based on the company's analysis of the HRQL data there is no apparent correlation between response status and HRQL. The model therefore does not appear to be fit for the purpose of demonstrating the clinical benefits of baricitinib.

The company's modelling approach also ignores important aspects of AD, such as the waxing and waning nature of the disease, and in doing so mischaracterises the aims of systemic therapy, focusing on the short-term alleviation of symptoms rather the ability of treatment to reduce the severity and frequency of exacerbations. This particularly impacts the modelling of patients who don't respond, or lose their response to treatment as they are assumed to remain in a state of chronic and severe AD until death. This is inconsistent with clinical reality and misrepresents the effectiveness of BSC.

The ERG also has substantive concerns regarding the company's positioning of baricitinib as a direct alternative to dupilumab. The mode of action of baricitinib, which is more broadly immunosuppressive than dupilumab, potentially places it as more natural alternative to immunosuppressive agents such as cyclosporin and methotrexate. Immunosuppressants are, however, not modelled as comparators or considered in the clinical evidence. Furthermore, the current positioning implies that baricitinib will directly replace dupilumab, precluding the use of dupilumab in patients who have failed baricitinib. This is clinically undesirable given that baricitinib is less effective than dupilumab. Furthermore, the economic analysis does not consider potentially relevant treatment sequences that include both baricitinib and dupilumab. It is therefore not clear from the company's analysis that a treatment sequence involving only baricitinib would represent the most cost-effective option.

In addition to the largely structural issues described above, the ERG also identified several uncertainties relating to key model inputs. Foremost among these is the approach adopted by the company to modelling treatment response rates. The company's approach applied response rates derived using absolute measures of the treatment effect rather than relative effects. This approach can lead to bias and where the response rates in the common comparator (placebo +TCS) differ across studies and departs from recommendations made in NICE DSU TSD5.[3] In line with the NICE DSU TSD5 the ERG favours applying the relative effects to the response probabilities on the log-odds scale, which also correctly captures the uncertainty. The ERG also considers the secondary censoring analysis to be more reflective of practice.

The ERG further questions the assumption of equivalence with dupilumab in terms of long-term efficacy and discontinuation. These assumptions are a major driver of cost-effectiveness, but are not

supported by the available data, with evidence from both the JAIN and JAHN studies suggesting higher rates of discontinuation and few patients retaining response.

Other major uncertainties relate to the company's approach to modelling HRQL which was based on within group analysis of the available HRQL data. Such an approach, is however, highly subject to bias and is inconsistent with the modelling of similar chronic conditions such psoriasis. The company's approach also ignores relevant HRQL data from the JAHL and JAHM studies, which the ERG believes should have been considered when generating the relevant utilities.

The impact of the uncertainties and structural issues described above, along with several more minor issues was considered in a series of exploratory analyses. The results of this illustrated that several of the ERG's alternative assumptions impacted significantly on the results of the economic analysis. Specifically, assumptions made regarding discontinuation rates, the modelling of BSC patients and the utilities applied in the model, were shown to be important drivers of cost-effectiveness. In the ERG base-case which consider several alternative treatment sequence, BSC was found to be the most cost-effective treatment at WTP of £30,000 per QALY. The pairwise ICER for baricitinib compared to BSC was £64,710 per QALY. Importantly, dupilumab was not found to be cost-effective compared to BSC in any analysis and was dominated by the sequence baricitinib, dupilumab in ERG's base case. These results are however, exclusive of PAS discount for dupilumab, see confidential appendix for details and brief discussion of the impact of the dupilumab PAS.

7 REFERENCES

1. Chopra R, Vakharia PP, Sacotte R, Patel N, Immaneni S, White T, et al. Severity strata for Eczema Area and Severity Index (EASI), modified EASI, Scoring Atopic Dermatitis (SCORAD), objective SCORAD, Atopic Dermatitis Severity Index and body surface area in adolescents and adults with atopic dermatitis. *Br J Dermatol* 2017;**177**:1316-21.

2. Leshem YA, Hajar T, Hanifin JM, Simpson EL. What the Eczema Area and Severity Index score tells us about the severity of atopic dermatitis - an interpretability study. *Br J Dermatol* 2015;**172**:1353-7.

3. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU technical support document 5: evidence synthesis in the baseline natural history model*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2012.

4. Weidinger S, Novak N. Atopic dermatitis. *Lancet* 2016;**387**:1109-22.

5. Kapur S, Watson W, Carr S. Atopic dermatitis. *Allergy Asthma Clin Immunol* 2018;**14**:52.

6. Thomsen SF. Atopic dermatitis: natural history, diagnosis, and treatment. *ISRN Allergy* 2014;**2014**:354250.

7. Abuabara K, Margolis DJ, Langan SM. The long-term course of atopic dermatitis. *Dermatol Clin* 2017;**35**:291-7.

8. Office for National Statistics (ONS). *Analysis of population estimates tool*. ONS; 2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/analysisofpopulationestimatestool [accessed 15th July 2020].

9. Barbarot S, Auziere S, Gadkari A, Girolomoni G, Puig L, Simpson EL, et al. Epidemiology of atopic dermatitis in adults: Results from an international survey. *Allergy* 2018;**73**:1284-93.

10. Drucker AM, Ellis AG, Bohdanowicz M, Mashayekhi S, Yiu ZZN, Rochwerg B, et al. Systemic immunomodulatory treatments for patients with atopic dermatitis: a systematic review and network meta-analysis. *JAMA Dermatology* 2020;**156**:659-67.

11. Boguniewicz M, Alexis AF, Beck LA, Block J, Eichenfield LF, Fonacier L, et al. Expert Perspectives on Management of Moderate-to-Severe Atopic Dermatitis: A Multidisciplinary Consensus Addressing Current and Emerging Therapies. *J Allergy Clin Immunol Pract* 2017;**5**:1519-31.

12. Schmitt J, Spuls P, Boers M, Thomas K, Chalmers J, Roekevisch E, et al. Towards global consensus on outcome measures for atopic eczema research: results of the HOME II meeting. *Allergy* 2012;**67**:1111-7.

13. National Institute for Health and Care Excellence (NICE). *Atopic eczema in under 12s: diagnosis and management. Clinical guideline [CG57]*. NICE; 2007. Available from: https://www.nice.org.uk/guidance/cg57 [accessed 27th August 2020].

14. Nowicki R, Trzeciak M, Wilkowska A, Sokołowska-Wojdyło M, Ługowska-Umer H, Barańska-Rybak W, et al. Atopic dermatitis: current treatment guidelines. Statement of the experts of the Dermatological Section, Polish Society of Allergology, and the Allergology Section, Polish Society of Dermatology. *Postepy Dermatol Alergol* 2015;**32**:239-49.

15. National Institute for Health and Care Excellence (NICE). *Clinical Knowledge Summary: Eczema - atopic. What is it? Available at: https://cks.nice.org.uk/eczema-atopic#!backgroundSub [Last accessed: 24th January 2020]*. [accessed

16. Eichenfield LF, Tom WL, Berger TG, Krol A, Paller AS, Schwarzenberger K, et al. Guidelines of care for the management of atopic dermatitis: Section 2. Management and treatment of atopic

dermatitis with topical therapies. *J Am Acad Dermatol* 2014;**71**:116-32. Available from: https://doi.org/10.1016/j.jaad.2014.03.023

17. Patrizi A, Raone B, Ravaioli GM. Management of atopic dermatitis: safety and efficacy of phototherapy. *Clinical, cosmetic and investigational dermatology* 2015;**8**:511-20. Available from: https://pubmed.ncbi.nlm.nih.gov/26491366

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4599569/

18. Taylor K, Swan DJ, Affleck A, Flohr C, Reynolds NJ. Treatment of moderate-to-severe atopic eczema in adults within the UK: results of a national survey of dermatologists. *Br J Dermatol* 2017;**176**:1617-23.

19. Simpson EL, Bruin-Weller M, Flohr C, Ardern-Jones MR, Barbarot S, Deleuran M, et al. When does atopic dermatitis warrant systemic therapy? Recommendations from an expert panel of the International Eczema Council. *J Am Acad Dermatol* 2017;**77**:623-33.

20. National Institute for Health and Care Excellence (NICE). *Dupilumab for treating moderate to severe atopic dermatitis: Technology appraisal guidance [TA534]*. 2018. Available from: https://www.nice.org.uk/guidance/ta534 [accessed 28th January 2020].

21. Bao L, Zhang H, Chan LS. The involvement of the JAK-STAT signaling pathway in chronic inflammatory skin disease atopic dermatitis. *JAK-STAT* 2013;**2**:e24137. Available from: https://doi.org/10.4161/jkst.24137

22. Newsom M, Bashyam AM, Balogh EA, Feldman SR, Strowd LC. New and emerging systemic treatments for atopic dermatitis. *Drugs* 2020;**80**:1041-52.

23. Kantor R, Thyssen JP, Paller AS, Silverberg JI. Atopic dermatitis, atopic eczema, or eczema? A systematic review, meta-analysis, and recommendation for uniform use of 'atopic dermatitis'. *Allergy* 2016;**71**:1480-5.

24. *Eczema therapies. Living systematic review with network meta-analysis*. Women's College Hospital; 2020. Available from: http://eczematherapies.com/research/ [accessed 10th July 2020].

25. Simpson EL, Lacour JP, Spelman L, Galimberti R, Eichenfield LF, Bissonnette R, et al. Baricitinib in patients with moderate-to-severe atopic dermatitis and inadequate response to topical corticosteroids: results from two randomized monotherapy phase III trials. *Br J Dermatol* 2020.

26. Guttman-Yassky E, Silverberg JI, Nemoto O, Forman SB, Wilke A, Prescilla R, et al. Baricitinib in adult patients with moderate-to-severe atopic dermatitis: A phase 2 parallel, double-blinded, randomized placebo-controlled multiple-dose study. *J Am Acad Dermatol* 2019;**80**:913-21.e9.

27. *I4V-MC-JAIN (BREEZE-AD4) Clinical Study Report (13th April 2020)*.

28. *I4V-MC-JAIY (BREEZE-AD7) Clinical Study Report (25th October 2019)*.

29. *I4V-MC-JAHN (BREEZE-AD3) Clinical Study Report (7th November 2019)*.

30. Centre for Reviews and Dissemination. *Systematic Reviews. CRD's guidance for undertaking reviews in healthcare*. York: CRD, University of York; 2009.

31. ClinicalTrials.gov. *A study of baricitinib (LY3009104) in participants with moderate-to-severe atopic dermatitis*. ClinicalTrials.gov; Available from: https://clinicaltrials.gov/ct2/show/results/NCT02576938 [accessed 3rd August 2020].

32. Brunner PM, Guttman-Yassky E, Leung DY. The immunology of atopic dermatitis and its reversibility with broad-spectrum and targeted therapies. *J Allergy Clin Immunol* 2017;**139**:S65-s76.

33. Electronic Medicines Compendium (EMC). *Olumiant 2 mg Film-Coated Tablets: SmPC*. 2019. Available from: https://www.medicines.org.uk/emc/product/2434/smpc [accessed 20th April 2020].

34. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;**50**:683-91.

35. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;**22**:153-60.

36. National Institute for Health and Care Excellence (NICE). *Final Appraisal Determination: Dupilumab for treating moderate to severe atopic dermatitis*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta534/evidence [accessed 18th June 2020].

37. Institute for Clinical and Economic Review (ICER). *Dupilumab and crisaborole for atopic dermatitis: effectiveness and value*: Institute for Clinical and Economic Review; 2017.

38. Ellis CN, Drake LA, Prendergast MM, Abramovits W, Boguniewicz M, Daniel CR, et al. Cost-effectiveness analysis of tacrolimus ointment versus high-potency topical corticosteroids in adults with moderate to severe atopic dermatitis. *J Am Acad Dermatol* 2003;**48**:553-63.

39. Ellis CN, Kahler KH, Grueger J, Chang J. Cost effectiveness of management of mild-to-moderate atopic dermatitis with 1% pimecrolimus cream in children and adolescents 2-17 years of age. *Am J Clin Dermatol* 2006;**7**:133-9.

40. Pitt M, Garside R, Stein K. A cost-utility analysis of pimecrolimus vs. topical corticosteroids and emollients for the treatment of mild and moderate atopic eczema. *Br J Dermatol* 2006;**154**:1137-46.

41. Healy E, Bentley A, Fidler C, Chambers C. Cost-effectiveness of tacrolimus ointment in adults and children with moderate and severe atopic dermatitis: twice-weekly maintenance treatment vs. standard twice-daily reactive treatment of exacerbations from a third party payer (UK National Health Service) perspective. *Br J Dermatol* 2011;**164**:387-95.

42. Ameen M, George S, Harwood C, Weller R, Lear J, Hudson R, et al. PSS44 - An evaluation of the current treatment pathways and associated NHS resource use for the management of uncontrolled moderate-to-severe atopic dermatitis in secondary care. *Value Health* 2018;**21**:S430.

43. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095-108.

44. Blauvelt A, de Bruin-Weller M, Gooderham M, Cather JC, Weisman J, Pariser D, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corticosteroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. *Lancet* 2017;**389**:2287-303.

45. Silverwood RJ, Forbes HJ, Abuabara K, Ascott A, Schmidt M, Schmidt SAJ, et al. Severe and predominantly active atopic eczema in adulthood and long term risk of cardiovascular disease: population based cohort study. *BMJ* 2018;**361**:k1786. Available from: https://www.bmj.com/content/bmj/361/bmj.k1786.full.pdf

46. Brunner PM, Guttman-Yassky E. Racial differences in atopic dermatitis. *Ann Allergy Asthma Immunol* 2019;**122**:449-55. Available from: http://www.sciencedirect.com/science/article/pii/S1081120618314200

47. Kaufman BP, Guttman-Yassky E, Alexis AF. Atopic dermatitis in diverse racial and ethnic groups—Variations in epidemiology, genetics, clinical presentation and treatment. *Exp Dermatol* 2018;**27**.

48. Noda S, Suárez-Fariñas M, Ungar B, Kim SJ, de Guzman Strong C, Xu H, et al. The Asian atopic dermatitis phenotype combines features of atopic dermatitis and psoriasis with increased TH17 polarization. *J Allergy Clin Immunol* 2015;**136**:1254-64. Available from: http://www.sciencedirect.com/science/article/pii/S0091674915011963

49. Sanyal RD, Pavel AB, Glickman J, Chan TC, Zheng X, Zhang N, et al. Atopic dermatitis in African American patients is TH2/TH22-skewed with TH1/TH17 attenuation. *Ann Allergy Asthma Immunol* 2019;**122**:99-110.e6. Available from: http://www.sciencedirect.com/science/article/pii/S1081120618307464

50. Williams HC, Pembroke AC, Forsdyke H, Boodoo G, Hay RJ, Burney PGJ. London-born black caribbean children are at increased risk of atopic dermatitis. *J Am Acad Dermatol* 1995;**32**:212-7. Available from: http://www.sciencedirect.com/science/article/pii/0190962295901280

51. Kim Y, Blomberg M, Rifas-Shiman SL, Camargo CA, Jr., Gold DR, Thyssen JP, et al. Racial/Ethnic Differences in Incidence and Persistence of Childhood Atopic Dermatitis. *The Journal of Investigative Dermatology* 2018;**139**:1523-747 (Electronic).

52. de Bruin-Weller M, Thaci D, Smith CH, Reich K, Cork MJ, Radin A, et al. Dupilumab with concomitant topical corticosteroid treatment in adults with atopic dermatitis with an inadequate response or intolerance to ciclosporin A or when this treatment is medically inadvisable: a placebo-controlled, randomized phase III clinical trial (LIBERTY AD CAFE). *Br J Dermatol* 2018;**178**:1083-101.

53. Blauvelt A, de Bruin-Weller M, Gooderham M, Cather JC, Weisman J, Pariser D, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corticosteroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. *Lancet* 2017;**389**:2287-303. Available from: http://www.sciencedirect.com/science/article/pii/S0140673617311911

54. National Institute for Health and Care Excellence. *Brodalumab for treating moderate to severe plaque psoriasis [TA511]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta511 [accessed 20th April 2020].

55. National Institute for Health and Care Excellence. *Certolizumab pegol for treating moderate to severe plaque psoriasis [ta574]*. NICE; Available from: https://www.nice.org.uk/guidance/ta574/resources/certolizumab-pegol-for-treating-moderate-to-severe-plaque-psoriasis-pdf-82607142805189 [accessed 20th April 2020].

56. National Institute for Health and Care Excellence. *Tildrakizumab for treating moderate to severe plaque psoriasis [TA575]*. NICE; 2019. Available from: https://www.nice.org.uk/guidance/ta575 [accessed 20TH aPRIL 2020].

57. Yosipovitch G, Reaney M, Mastey V, Eckert L, Abbé A, Nelson L, et al. Peak Pruritus Numerical Rating Scale: psychometric validation and responder definition for assessing itch in moderate-to-severe atopic dermatitis. *Br J Dermatol* 2019;**181**:761-9.

58. Kimball AB, Naegeli AN, Edson-Heredia E, Lin C-Y, Gaich C, Nikai E, et al. Psychometric properties of the Itch Numeric Rating Scale inpatients with moderate-to-severe plaque psoriasis. *Br J Dermatol* 2016;**175**:157-62.

59. National Institute for Health and Care Excellence (NICE). *Apremilast for treating moderate to severe plaque psoriasis: Technology appraisal guidance [TA368]*: NICE; 2015.

60. National Institute for Health and Care Excellence (NICE). *Tofacitinib for treating active psoriatic arthritis after inadequate response to DMARDs [TA543]*; 2018.

61. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15**:708-15.

62. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;**14**:539-45.

63. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Personal Social Services Research Unit (PSSRU); 2019. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/ [accessed 9th June 2020].

64. Santer M, Rumsby K, Ridd MJ, Francis NA, Stuart B, Chorozoglou M, et al. Adding emollient bath additives to standard eczema management for children with eczema: the BATHE RCT. *Health Technol Assess* 2018;**22**:57.

65. NHS Improvement. *National schedule of NHS costs -Year 2018-2019*. Available from: https://improvement.nhs.uk/resources/national-cost-collection/ [accessed 20th April 2020].

66. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury, Kent: Personal Social Services Research Unit, University of Kent 2020. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/

APPENDICES

Appendix 1: Adverse events

Table 66 Commonly observed treatment-emergent adverse events (TEAEs) observed in JAIN at 24 weeks (adapted from Table 185 of the CSR)

TEAEs affecting > 3% of patients, n (%)	PBO + TCS ([redacted])	BARI 1 mg + TCS ([redacted])	BARI 2 mg +TCS ([redacted])	BARI 4 mg +TCS ([redacted])
TEAEs at 24 Weeks				
≥ 1 TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]
Influenza	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]
Upper abdominal pain	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Oral herpes	[redacted]	[redacted]	[redacted]	[redacted]
Herpes simplex	[redacted]	[redacted]	[redacted]	[redacted]
Peripheral oedema	[redacted]	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]	[redacted]	[redacted]
Acne	[redacted]	[redacted]	[redacted]	[redacted]
Conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]
Erysipelas	[redacted]	[redacted]	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]	[redacted]	[redacted]
Skin infection	[redacted]	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]
Oropharyngeal pain	[redacted]	[redacted]	[redacted]	[redacted]
Asthma	[redacted]	[redacted]	[redacted]	[redacted]
Blood creatine phosphokinase increased	[redacted]	[redacted]	[redacted]	[redacted]
Bronchitis	[redacted]	[redacted]	[redacted]	[redacted]

TEAE: treatment-emergent adverse event; PBO: placebo; BARI: baricitinib; TCS: topical corticosteroid

Appendix 2: Additional ITC data and results tables

Table 67 Relative treatment effects for EASI50 at week 16 for the base-case population: JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients, the full JAIN versus CAFÉ population and the European patient population (adapted from Table 58, 60 and 64 of the CS)

Source	n/N (%) Placebo	n/N (%) Active Treatment	OR†	95% CI
Primary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
Pooled: JAIN+ JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	47/108 (43.5)	91/107 (85)	[redacted]	[redacted]
CAFÉ-like CHRONOS	NA/61 (NA)	NA/23 (NA)	[redacted]	
Pooled: CAFÉ + CAFÉ-like CHRONOS	64/169	108/130	[redacted]	[redacted]
ITC: Pooled JAIN+ JAIN-like JAIY vs. Pooled CAFÉ + CAFÉ-like CHRONOS (fixed-effects model)			[redacted]	[redacted]
ITC: Full JAIN vs. CAFÉ (fixed-effects model)			[redacted]	[redacted]
Secondary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
Pooled: JAIN+ JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	54/108 (50)	95/107 (88.8)	7.92*	3.90 to 16.10*
CAFÉ-like CHRONOS	NA/61 (NA)	NA/23 (NA)	NA	NA
Pooled: CAFÉ + CAFÉ-like CHRONOS	82/169	115/130	[redacted]	[redacted]
ITC: Pooled JAIN+ JAIN-like JAIY vs. Pooled CAFÉ + CAFÉ-like CHRONOS (fixed-effects model)			[redacted]	[redacted]
European Patients (Primary Censoring)				
BARI + TCS vs PBO + TCS				
JAIN (European Patients)	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	47/108 (43.5)	91/107 (85)	[redacted]	[redacted]
ITC: European JAIN vs. CAFÉ (fixed-effects model)			[redacted]	[redacted]

* Calculated by the ERG

† ORs presented were for fixed effects meta-analyses

OR: odds ratio; CI: confidence interval; BARI: baricitinib; PBO: placebo; TCS: topical corticosteroids; NA: not available

Table 68 Relative treatment effects for EASI75 at week 16 for the base-case population: JAIN + JAIN-like JAIY patients versus CAFÉ + CAFÉ-like CHRONOS patients, the full JAIN versus CAFÉ and the European patient population (adapted from Table 59, 61 and 65 of the CS)

Source	n/N (%)	n/N (%)	OR†	95% CI
	Placebo	Active Treatment		
Primary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
Pooled: JAIN+ JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	32/108 (29.6)	67/107 (62.6)	[redacted]	[redacted]
CAFÉ-like CHRONOS	NA/61 (NA)	NA/23 (NA)	NA	NA
Pooled: CAFÉ + CAFÉ-like CHRONOS	43/169 (0.3)	83/130 (0.6)	[redacted]	[redacted]
ITC: Pooled JAIN+ JAIN-like JAIY vs. Pooled CAFÉ + CAFÉ-like CHRONOS (fixed- effects model)			[redacted]	[redacted]
ITC: Full JAIN vs. CAFÉ (fixed-effects model)			[redacted]	[redacted]
Secondary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
Pooled: JAIN+ JAIN-like JAIY	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	35/108 (0.3)	69/107 (64.5)	3.79*	2.84 to 5.05*
CAFÉ-like CHRONOS	NA/61 (NA)	NA/23 (NA)	NA	NA
Pooled: CAFÉ + CAFÉ-like CHRONOS	51/169 (0.3)	87/130 (66.9)	[redacted]	[redacted]
ITC: Pooled JAIN+ JAIN-like JAIY vs. Pooled CAFÉ + CAFÉ-like CHRONOS (fixed-effects model)			[redacted]	[redacted]
European Patients (Primary Censoring)				
BARI + TCS vs PBO + TCS				
JAIN (European Patients)	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	32/108 (29.6)	67/107 (62.6)	[redacted]	[redacted]
ITC: European JAIN vs. CAFÉ (fixed-effects model)			[redacted]	[redacted]

***** Calculated by the ERG

† ORs presented were for fixed effects meta-analyses

OR: odds ratio; CI: confidence interval; BARI: baricitinib; PBO: placebo; TCS: topical corticosteroids; NA: not available

Table 69 Relative treatment effects for EASI90 at week 16 for the full JAIN versus CAFÉ trial populations and the European patient population (adapted from Table 62 and 66 of the CS)

Source	n/N (%) Placebo	n/N (%) Active Treatment		OR†	95% CI
	Primary Censoring				
	BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]		[redacted]	[redacted]
	DUPI + TCS vs PBO + TCS				
CAFÉ	49/107 (45.8)	13/108 (12)		[redacted]	[redacted]
ITC: Full JAIN vs. CAFÉ (fixed-effects model)				[redacted]	[redacted]
	European Patients (Primary Censoring)				
	BARI + TCS vs PBO + TCS				
JAIN (European)	[redacted]	[redacted]		[redacted]	[redacted]
	DUPI + TCS vs PBO + TCS				
CAFÉ	13/108 (12)	49/107 (45.8)		[redacted]	[redacted]
ITC: European JAIN vs. CAFÉ (fixed-effects model)				[redacted]	[redacted]

† ORs presented were for fixed effects meta-analyses

OR: odds ratio; CI: confidence interval; BARI: baricitinib; PBO: placebo; TCS: topical corticosteroids; NA: not available

Table 70 Relative treatment effects for Itch NRS ≥4-point improvement at week 16 for the full JAIN versus CAFÉ trial populations and the European patient population (adapted from Table 63 and 67 of the CS)

Source	n/N (%) Placebo	n/N (%) Active Treatment	OR†	95% CI
Primary Censoring				
BARI + TCS vs PBO + TCS				
JAIN	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	13/91 (14.3)	43/94 (45.7)	5.06	2.48 to 10.33
ITC: Full JAIN vs. CAFÉ (fixed-effects model)			[redacted]	[redacted]
European Patients (Primary Censoring)				
BARI + TCS vs PBO + TCS				
JAIN (European)	[redacted]	[redacted]	[redacted]	[redacted]
DUPI + TCS vs PBO + TCS				
CAFÉ	13/91 (14.3)	43/94 (45.7)	5.06	2.48 to 10.33
ITC: European JAIN vs. CAFÉ (fixed-effects model)			[redacted]	[redacted]

† ORs presented were for fixed effects meta-analyses

OR: odds ratio; CI: confidence interval; BARI: baricitinib; PBO: placebo; TCS: topical corticosteroids; NA: not available

Appendix 3: Drummond checklist

Table 71 Quality assessment of included CEA study using Drummond et al. checklist completed by the ERG

	CEA quality assessment questions	Answer (Yes/No/Unclear)	Notes/Explanation for No or Unclear
1	Was the research question stated?	Yes	
2	Was the economic importance of the research question stated?	Yes	
3	Was/were the viewpoint(s) of the analysis clearly stated and justified?	Yes	
4	Was a rationale reported for the choice of the alternative programmes or interventions compared?	No	Immunosuppressive agents (e.g. ciclosporin, methotrexate) were excluded as comparators despite being potentially relevant treatment options in the modelled population. The model also did not consider treatment sequences including multiple active treatments.
5	Were the alternatives being compared clearly described?	Yes	
6	Was the form of economic evaluation stated?	Yes	
7	Was the choice of form of economic evaluation justified in relation to the questions addressed?	Yes	
8	Was/were the source(s) of effectiveness estimates used stated?	Yes	
9	Were details of the design and results of the effectiveness study given (if based on a single study)?	Yes	
10	Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of a number of effectiveness studies)?	Yes	

11	Were the primary outcome measure(s) for the economic evaluation clearly stated?	Yes	
12	Were the methods used to value health states and other benefits stated?	Yes	
13	Were the details of the subjects from whom valuations were obtained given?	Yes	
14	Were productivity changes (if included) reported separately?	N/A	Productivity changes were not reported
15	Was the relevance of productivity changes to the study question discussed?	No	Implication upon productivity were not discussed.
16	Were quantities of resources reported separately from their unit cost?	Yes	
17	Were the methods for the estimation of quantities and unit costs described?	Yes	
18	Were currency and price data recorded?	Yes	
19	Were details of price adjustments for inflation or currency conversion given?	N/A	No inflation of prices was necessary.
20	Were details of any model used given?	Yes	
21	Was there a justification for the choice of model used and the key parameters on which it was based?	Yes	
22	Was the time horizon of cost and benefits stated?	Yes	
23	Was the discount rate stated?	Yes	
24	Was the choice of rate justified?	Yes	

25	Was an explanation given if cost or benefits were not discounted?	N/A	All costs and benefit were discounted.
26	Were the details of statistical test(s) and confidence intervals given for stochastic data?	No	Parameters for measurement and uncertainty were provided, but were not based on confidence intervals for stochastic data. A revised model provided at the clarification stage.
27	Was the approach to sensitivity analysis described?	Yes	
28	Was the choice of variables for sensitivity analysis justified?	Yes	
29	Were the ranges over which the parameters were varied stated?	Yes	
30	Were relevant alternatives compared? (That is, were appropriate comparisons made when conducting the incremental analysis?)	Yes	
31	Was an incremental analysis reported?	Yes	
32	Were major outcomes presented in a disaggregated as well as aggregated form?	Yes	
33	Was the answer to the study question given?	Yes	
34	Did conclusions follow from the data reported?	Yes	
35	Were conclusions accompanied by the appropriate caveats?	Yes	
36	Were generalisability issues addressed?	No	Subgroup analysis demonstrated differential efficacy in some groups including patients with more/less severe disease and patients from different ethnic groups. These differences were, however, not explored in the economic analysis. The available trial evidence included no patients with back skin limiting the conclusions that can be drawn about the cost-effectiveness of baricitinib in this group.

Appendix 4: Calculation of absolute probabilities using relative effects from the ITC

Absolute probabilities of response were calculated in R version 4.0.2 using the R2OpenBUGS (v 3.2.3.2.1). The number of patients in the JAIN, JAIN-like JAIY, CAFÉ and/or CAFÉ-like CHRONOS populations who responded in the placebo arm were pooled on the log-odds scale to form a baseline probability of response. A Bayesian fixed effect meta-analysis model with non-informative prior distributions was used.

The log-odds ratios of baricitinib and dupilumab compared to placebo were assumed to follow a normal distribution and added to the log-odds of response on placebo to calculate the log-odds of response on baricitinib and dupilumab, respectively (results using primary or secondary censoring were used). These were then transformed to the probability scale to obtain posterior distributions for the absolute probabilities of response on placebo (as a proxy for BSC) baricitinib and dupilumab. The mean and standard error of these probabilities were included in the economic model, where they were assumed to follow a Beta distribution.

Four different scenarios were run using two chains with a burn-in of 10,000 iterations, by which convergence had occurred, and posterior samples based on a further 20,000 iterations. OpenBUGS code, initial values and data for each scenario are given below.

OpenBUGS code

```
# Code to meta-analyse probabilities on log-odds scale
# Outputs:
#   pooled posterior probability (mean and SE) for placebo and active
treatments
#
# Fixed effect
model{
for (i in 1:ns){
  r[i] ~ dbin(p[i], n[i])    # likelihood for probability
  logit(p[i]) <- mu          # model on log-odds scale
 }
mu ~ dnorm(0, .0001)         # prior distribution for pooled log-odds
logit(R.pl) <- mu            # calculate posterior probability: Placebo
#
# calculate active treatment absolute effects
for (k in 1:nt){
  LOR[k] ~ dnorm(eff[k], prec[k])
  prec[k] <- pow(se[k], -2)
  logit(R.trt[k]) <- mu + LOR[k] # calculate posterior probability: Active
 }
}
```

Initial values

chain 1

list(mu=0.00000E+00, LOR=c(0.00000E+00, 0.00000E+00))

chain 2

list(mu=-5.00000E+00, LOR=c(2.00000E+00, -2.00000E+00))

Data

Scenario 2a: Primary censoring, Placebo probability from JAIN-like patients
list(nt=2.00000E+00, ns=2.00000E+00, r=c(2.70000E+01, 1.80000E+01), n=c(9.30000E+01, 5.10000E+01), eff=c(7.41976E-01, 2.11756E+00), se=c(2.44817E-01, 2.67694E-01))

Scenario 2b: Primary censoring, Placebo probability from JAIN-like & CAFE-like patients
list(nt=2.00000E+00, ns=3.00000E+00, r=c(2.70000E+01, 1.80000E+01, 3.50000E+01), n=c(9.30000E+01, 5.10000E+01, 1.69000E+02), eff=c(7.41976E-01, 2.11756E+00), se=c(2.44817E-01, 2.67694E-01))

Scenario 2c: Secondary censoring, Placebo probability from JAIN-like patients
list(nt=2.00000E+00, ns=2.00000E+00, r=c(3.00000E+01, 1.80000E+01), n=c(9.30000E+01, 5.10000E+01), eff=c(8.62525E-01, 1.95240E+00), se=c(2.41874E-01, 2.61861E-01))

Scenario 2d: Secondary censoring, Placebo probability from JAIN-like & CAFE-like patients
list(nt=2.00000E+00, ns=3.00000E+00, r=c(3.00000E+01, 1.80000E+01, 4.70000E+01), n=c(9.30000E+01, 5.10000E+01, 1.69000E+02), eff=c(8.62525E-01, 1.95240E+00), se=c(2.41874E-01, 2.61861E-01))

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies, you must inform NICE by **5pm on Monday 14 September** using the below comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Section 1: Major issues

Issue 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 12, first sentence states: "*The population considered in the submission was adults with moderate to severe atopic dermatitis (AD) who are candidates for systemic therapy and who have a history of inadequate response to topical therapy as well as a history of intolerance to, contraindication to, or inadequate response to ciclosporin.*" Page 44 states: "*The population of interest in the CS is adult patients who are candidates for systemic therapy who had a history of inadequate response to topical therapy and a history of intolerance to, contraindication to, or inadequate response to ciclosporin.*" Page 85 states: "*The population considered in the submission was adult patients who are candidates for systemic therapy who had a history of inadequate response to topical therapy and a history of intolerance to, contraindication to,*	Please amend this wording to: "The population considered in the submission was adults with moderate to severe atopic dermatitis (AD) who are candidates for systemic therapy and **who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control**". Please include this wording wherever the population of interest is referred to in the ERG report.	As per Section B.1.1 of the Company Submission (CS), the population considered in the submission was adults patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control.	Amended

or inadequate response to ciclosporin."			

Issue 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 12 states: "*Therefore, patients on the lower end of the moderate scale may be excluded, biasing the trial populations towards more severe disease. The ERG therefore considers that the population presented in the clinical evidence may not represent all moderate to severe patients in the NHS population (Section 3.2.2).*" Page 17 states: "*The baricitinib trials (and thus the modelled population) limited inclusion to patients with more severe disease than would be expected in the moderate-to-severe AD population seen in NHS practice.*" Page 85 states: "*Therefore, in terms of age and disease severity, the ERG considers the population in the clinical evidence presented may not represent all moderate to severe patients in the*	Please consider amending this wording to reflect that this issue is not limited to baricitinib, but also dupilumab: *"Therefore, patients on the lower end of the moderate scale may be excluded, biasing the trial populations towards more severe disease.* ***This issue also applies to the clinical trials (CAFÉ and CHRONOS) supporting dupilumab, the main comparator considered in the appraisal, where the eligibility criteria included an EASI score of ≥20.*** *The ERG therefore considers that the population presented in the clinical evidence* ***and considered in the NICE appraisal of dupilumab (TA534)*** *may not represent all moderate to severe patients in the NHS population (Section 3.2.2)."* Please consider adding relevant wording wherever this issue is raised in the ERG report.	We acknowledge that several published strata for the EASI score indicate that the baricitinib trial populations may be biased towards more severe disease. However, this statement is factually inaccurate by omission and potentially misleading, since this issue is not specific to baricitinib, but also applies to the clinical trials supporting dupilumab (CAFÉ and CHRONOS).	Page 12 amended "This includes patients with EASI scores far below the trial cut off of 16. ***This issue also applies to the clinical trials (CAFÉ and CHRONOS) supporting dupilumab, the main comparator considered in the appraisal, where the eligibility criteria included an EASI score of ≥20.*** Therefore, patients on the lower end of the moderate scale may be excluded ***from the evidence presented for baricitinib and dupilumab***, biasing the trial populations towards more severe disease.[1]" Page 17, 85, 100 - Not a factual error as those statements are referring to baricitinib trials, not comparing

NHS population." Page 100 states: "*The inclusion criteria applied in all the pivotal trial evidence presented by the company require patients to have IGA ≥ 3 and EASI score > 16. This minimum requirement is, however, potentially overly restrictive and may mean that the recruited population has more severe disease than patients treated in practice*."			to dupilumab trials.

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 12 states: "*As a Janus kinase (JAK) 1/JAK2 inhibitor, baricitinib works similarly to other systemic immunosuppressants, whereas dupilumab has a more targeted mode of action.*" Page 18 states: "*Further, the mode of action of baricitinib, potentially places it as a more natural comparator to the immunosuppressants. This is because as a JAK1/JAK2 inhibitor, baricitinib is more broadly immunosuppressive than dupilumab which has a more focused mode of action distinct from both baricitinib and*	Please amend this wording to make clear that baricitinib has a more targeted mode of action than general immunosuppressants: Page 12 and Table 4, Comparator(s) row: "*As a Janus kinase (JAK) 1/JAK2 inhibitor, baricitinib ~~works similarly to~~* ***has a more targeted mode of action as compared with*** *other systemic immunosuppressants, ~~whereas dupilumab has a more targeted mode of action~~* ***but a less targeted mode of action as compared with dupilumab****."* Page 18: "*Further, the mode of action of baricitinib, potentially places it as a more natural comparator to the immunosuppressants* ***than dupilumab****. This is because as a JAK1/JAK2 inhibitor,* ***whilst*** *baricitinib* ***has a more targeted***	We acknowledge that dupilumab has a more specific mode of action than baricitinib. However, these statements are factually inaccurate by omission and potentially misleading, since baricitinib has a more specific mode of action than general systemic immunosuppressants.	Page 12. Amended Table 4 amended to "*baricitinib* ***acts in a similar manner*** *to other systemic immunosuppressants such as methotrexate and ciclosporin in targeting a broad****er*** *range of cellular processes and mediators* ***than dupilumab***." Page 18 amended to "Further, the mode of action

immunosuppressants." Table 4 in the Population row states: "*baricitinib works similarly to other systemic immunosuppressants such as methotrexate and ciclosporin in targeting a broad range of cellular processes and mediators*." Table 4 in the Comparator(s) row states: "*As a JAK1/JAK2, baricitinib works similarly to other systemic immunosuppressants, whereas dupilumab has a more targeted mode of action*." Page 102 states: "*As outlined in Table 4 and further discussed in Section 4.2.3, there are several reasons to consider that immunosuppressive agents such as ciclosporin, methotrexate, and azathioprine represent a more natural comparator to baricitinib than dupilumab due to the similarities in their mode of action*." Page 106 states: "*Further, the mode of action of baricitinib potentially places it as a more a natural comparator to the immunosuppressants such as ciclosporin and methotrexate, than dupilumab. This is because as a JAK1/JAK2 inhibitor, baricitinib is more broadly immunosuppressive*	***mode of action as compared with other systemics, it** is more broadly immunosuppressive than dupilumab which has a more focused mode of action distinct from both baricitinib and immunosuppressants*." Table 4, Population row: "*baricitinib ~~works similarly to~~ **has a more targeted mode of action as compared with** other systemic immunosuppressants such as methotrexate and ciclosporin ~~in targeting~~ **but acts in a similar manner to target** a broad**er** range of cellular processes and mediators **than dupilumab***." Page 102: "*As outlined in Table 4 and further discussed in Section 4.2.3, there are several reasons to consider that immunosuppressive agents such as ciclosporin, methotrexate, and azathioprine represent a more natural comparator to baricitinib than dupilumab due to the similarities in their mode of action, **although baricitinib has a more specific mode of action as compared with these systemic immunosuppressants**."* Page 106: "*Further, the mode of action of baricitinib potentially places it as a more natural comparator to the immunosuppressants such as ciclosporin and methotrexate, than dupilumab. This is because as a JAK1/JAK2 inhibitor, baricitinib **has a more targeted mode of action as compared with other systemics, but** is more broadly immunosuppressive than dupilumab which has a more focused mode of action distinct from both baricitinib and immunosuppressants*."		of baricitinib, potentially places it as a more natural comparator to the immunosuppressants **than dupilumab**. This is because as a JAK1/JAK2 inhibitor, baricitinib is more broadly immunosuppressive than dupilumab which has a more focused mode of action distinct from both baricitinib and immunosuppressants" Page 102: Not a factual error. The ERG is stating that the mode of action of immunosuppressants and baricitinib is **similar** not identical. Page 108: amended Page 149: Not a factual error. The ERG does not consider the additional detail to be relevant in this sentence.

than dupilumab which has a more focused mode of action, distinct from both baricitinib and immunosuppressants." Page 149 states: "*The mode of action of baricitinib, which is more broadly immunosuppressive than dupilumab…*"	Page 149: "*The mode of action of baricitinib, which* ***has a more specific mode of action as compared with systemic immunosuppressants but*** *is more broadly immunosuppressive than dupilumab…"*		

Issue 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 17 states: "*The ERG found it particularly concerning that there were no black patients included in the evidence base comprising the company's ITC, given the greater prevalence and severity of AD in the Black British population.*" Page 102: "*The ERG also finds concerning that there were no black patients included in the JAIN and JAIN-like, JAIY, JAHL, and JAHM populations comprising the company's ITC.*" Page 102: "The reasons for the absence of black patients from the trial are unclear, though the ERG notes that several centres were located in countries where black people represent a non-	Please consider amending this wording to reflect that this issue pertains to both baricitinib and comparator trials: Page 17: "*The ERG found it particularly concerning that there were no black patients included in the evidence base comprising the company's ITC, given the greater prevalence and severity of AD in the Black British population.* ***This issue pertains to both baricitinib and comparator trials.***" Page 102: "*The ERG also finds concerning that there were no Black patients included in the JAIN and JAIN-like, JAIY, JAHL, and JAHM populations comprising the company's ITC.* ***Similarly, no black patients were included in the placebo + TCS or dupilumab Q2W + TCS arms of the CAFÉ trial, and only 4 black patients were included in the CAFÉ + CHRONS CAFÉ-like pooled population.*** " Page 102: *"The reasons for the absence of*	We acknowledge the ERG's concerns surrounding the lack of black patients included in the JAIN and JAIN-like, JAIY, JAHL, and JAHM populations comprising the ITC. However, this statement is factually inaccurate by omission and potentially misleading, since this issue is not specific to baricitinib, but also applies to the evidence base supporting dupilumab. No black patients were included in the placebo + TCS or dupilumab Q2W + TCS arms of the CAFÉ trial, and only 4 black patients were included in the CAFÉ + CHRONS CAFÉ-like pooled population. We also appreciate that the ERG have acknowledged the lack of US centres included in the trials. This	Page 17 - amended Page 102: Not a factual error. This section is discussing baricitinib trials, so it is not relevant to include detail about the comparator trials.

negligible proportion of the population.”	*black patients from the trial are unclear, though the ERG notes that several centres were located in countries where black people represent a non-negligible proportion of the population.* ***However, patients were not recruited from the US; if they had been, it would be expected that a higher proportion of black patients would have been included.****”*	important context should be reported when discussing the locations of the centres included in the trials.	

Issue 5

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 17 states: “*It is not clear whether the post hoc composite of EASI50 and ΔDLQI ≥ 4 outcomes to define response would be recognised or treated as clinically meaningful in practice.*”	Please consider amending this wording to acknowledge that the composite endpoint of EASI50 and ΔDLQI ≥ 4 was accepted in TA534: *“****Whilst the committee in TA534 concluded that the post hoc composite of EASI50 and ΔDLQI ≥ 4 outcomes to define response was appropriate for decision-making****, it is not clear whether this would be recognised or treated as clinically meaningful in practice.*”	As reported the Appraisal Consultation Document for TA534, the committee concluded that the composite endpoint of EASI 50 plus an improvement in the DLQI of at least 4 was appropriate for decision-making. As such, whilst we acknowledge the ERG’s concerns surrounding the lack of correlation between response and HRQL in the regression analysis, it should be noted that the use of the composite endpoint to define response was in line with what has been previously accepted for a similar decision problem.	Not a factual error We disagree this a factual error but consider that it is relevant to acknowledge that the committee have previously accepted this. We have edited accordingly.

Issue 6

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 18 states: "*Furthermore, a substantial proportion of the eligible population will be dupilumab-experienced, which would preclude patient access to baricitinib if sequences are not permitted.*" Page 105 states: *"Further, by the company's logic it would be necessary for any NICE recommendation for baricitinib to explicitly prohibit the use of dupilumab as subsequent therapy, as there is currently nothing in the NICE recommendations or marketing authorisation for dupilumab that precludes its use following another biologic agent."* Page 149 states: *"Furthermore, the current positioning implies that baricitinib will directly replace dupilumab, precluding the use of dupilumab in patients who have failed baricitinib."*	Please consider removing the discussion surrounding a NICE recommendation for baricitinib influencing the use of dupilumab.	Contrary to the statement made by the ERG, the Company did not make any argument in support of this statement. Is not an accurate reflection of the NICE technical appraisal process, since this outcome of this appraisal cannot impact the status or wording of the NICE recommendation for dupilumab. Thus, approving baricitinib cannot prohibit the use of dupilumab as a subsequent therapy.	Not a factual error The company modelled dupilumab as a comparator to baricitinib and were very clear in the clarification response that they considered an analysis where baricitinib is used in a sequence including dupilumab to be irrelevant. This suggests that patients will either receive baricitinib **or** dupilumab not both.

Issue 7

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 20 states: "*While this population is consistent with the effectiveness data used in the model, it ignores available data on JAIN-like patients recruited to other pivotal trials (JAHL and JAHM). It is the ERG's view that utility should be drawn from the largest possible sample and should include all JAIN-like patients, particularly given the small number of responders providing data (see Section 4.2.6).*"	Please consider amending this wording to acknowledge that patients in the JAHL and JAHM trials were not receiving TCS in combination with baricitinib. "*While this population is consistent with the effectiveness data used in the model, it ignores available data on JAIN-like patients recruited to other pivotal trials (JAHL and JAHM). It is the ERG's view that utility should be drawn from the largest possible sample and should include all JAIN-like patients, particularly given the small number of responders providing data (see Section 4.2.6).* ***However, it should be noted that, unlike the patient population considered in the economic model, patients recruited to the JAHL and JAHM trials did not receive concomitant TCS.***"	We acknowledge the ERG's concerns presented in Section 4.2.6 surrounding the justification that JAHL and JAHM patients are not relevant to the modelled population. However, unlike the patients considered in the economic model who were all assumed to receive concomitant TCS (across those receiving baricitinib, placebo and dupilumab), patients in the JAHL and JAHM studies did not receive concomitant TCS. It is important to acknowledge this additional context.	Not a factual error This is a summary of the ERG's position and the argument that these are monotherapy patients is discussed and rejected in Section 4.2.6.

Issue 8

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 21 describes ERG Scenario 5b as: "*Applying the dupilumab utilities to all treatment arms.*" Page 23 describes ERG Scenario 5b as: "*Utilities based on dupilumab (TA534) submission.*" Page 145 describes ERG	Please update this wording to: Page 21: "*Applying the dupilumab* ***health state utility values*** *to each treatment arm.* ***In contrast to TA534, no waning of utility was modelled for BSC.***" Page 23 and Page 145: "***Health state utility values*** *based on dupilumab (TA534)*	This ERG scenario does not apply the waning of HRQoL as was performed in the dupilumab appraisal (TA534). Therefore, the application of the health state utility values from the dupilumab submission does not replicate the entire approach taken in TA534,	Not a factual error. The ERG have not updated the wording in the Tables on Page 21, 23 or 145. For clarity, the ERG have updated the wording to describe the utilities on Page

Scenario 5b as: "*Utilities based on dupilumab (TA534) submission.*"	*submission*. ***In contrast to TA534, no waning of utility was modelled for BSC.***" Please update this wording wherever this scenario analysis is referred to in the ERG report.	which will result in the overestimation of BSC utility in the long-term.	142: *"Patients on BSC, including patients classified as non-responders are assigned a single utility value based on the average of all placebo patients at week 16.* ***In contrast to TA534, no waning of utility was modelled for BSC****."*

Issue 9

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 28 states: "*The marketing authorisation for baricitinib is expected in* [redacted] *and positive opinion from the Committee for Human Medicinal Products (CHMP) is expected in* [redacted]."	Please amend the wording to: "*The marketing authorisation for baricitinib is expected* [redacted] *and positive opinion from the Committee for Human Medicinal Products (CHMP) is expected on* [redacted]."	[redacted]	Amended

Issue 10

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 72 states: "*It was unclear to the ERG why the company chose to look at the AEs observed in monotherapy trials and combination therapy trials collectively for the integrated safety analysis and further why*	Please consider adding the context that these data and explanations have not previously been requested from the company: "~~*It was unclear to the ERG why the company chose to look at the AEs observed in monotherapy trials and combination therapy*~~	The integrated safety analysis was included to provide an analysis set with higher number of patients. The 2 mg data, which are presented in the summary of clinical safety, were not provided for NICE and were not requested by the ERG at the	Page 72 – amended "The company did not **present** safety outcomes for patients who received 1 mg or 2 mg baricitinib, **in the CS**. As patients in these groups do not

the Phase II trial was included in the safety analysis. The company did not explain why they excluded safety outcomes for patients who received 1 mg or 2 mg baricitinib. As patients in these groups do not receive a higher dose of baricitinib than is recommended, the ERG considers the excluded data useful for safety analysis, particularly as the company proposes a 2 mg dose for patients aged over 75 years."	*~~trials collectively for the integrated safety analysis and further why the Phase II trial was included in the safety analysis.~~ The company did not ~~explain why they excluded~~* ***present*** *safety outcomes for patients who received 1 mg or 2 mg baricitinib* ***in the CS****. As patients in these groups do not receive a higher dose of baricitinib than is recommended, the ERG considers this data useful for safety analysis, particularly as the company proposes a 2 mg dose for patients aged over 75 years.* ***The Company subsequently provided these data****."*	clarification questions stage. This additional reference has been provided alongside this document,	receive a higher dose of baricitinib than is recommended, the ERG considers these data useful for safety analysis, particularly as the company proposes a 2 mg dose for patients aged over 75 years. **The Company subsequently provided these data."**

Issue 11

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 102 states: "*the company response outlining that the trial program was not designed to investigate baricitinib efficacy in black patients.*"	Please amend this wording to: "*the company response outlining that the trial program was not designed to investigate baricitinib efficacy in black patients* **compared with other patient populations**."	It is not stated in the clarification question response that the trial program was not designed to investigate baricitinib efficacy in black patients. The trials were not designed to investigate differential efficacy across ethnic groups, as per the wording of the response to clarification questions: "the trial program was not designed to investigate baricitinib efficacy in black patients <u>compared with other patient populations</u>." As such, this statement is factually inaccurate by omission and potentially misleading.	Amended as suggested.

Issue 12

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 106 states: "*Given these concerns, the ERG requested the company provide further justification for excluding immunosuppressants as a comparator and to provide a relevant scenario analysis; however, the company declined to attempt such an analysis.*"	Please consider amending this wording to reflect that this analysis was found to be infeasible: "*Given these concerns, the ERG requested the company provide further justification for excluding immunosuppressants as a comparator and to provide a relevant scenario analysis; however,* ***the feasibility assessment conducted by the company found the available data to be insufficient for a robust analysis to be performed***."	We acknowledge that the Company did not present this analysis in response to the ERG's request. However, this statement is factually inaccurate by omission and potentially misleading, since the Company performed a feasibility assessment which found the analysis was not possible.	Not a factual error. The ERG is aware of limitations of the efficacy data supporting ciclosporin It is however, possible to generate a comparison with baricitinib. This is evidence in TA534 where scenario were presented including ciclosporin as a comparator.

Issue 13

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 122 states: "*The ERG therefore considers that a model structure based around different levels of response would have been the preferable approach, and would have better captured differences in HRQL achieved on different treatments.*"	Please consider adding further context to this statement: "*The ERG therefore considers that a model structure based around different levels of response would have been the preferable approach, and would have better captured differences in HRQL achieved on different treatments.* ***However, the ERG notes that for many outcomes, the lack of comparator data would preclude this analysis from being undertaken****."*	We acknowledge the ERG's concerns surrounding the use of a model which does not include different levels of response. However, this statement is factually inaccurate by omission and potentially misleading, since in several cases, the comparator data necessary for these analyses are not available, and therefore it was not feasible for the company to provide this model structure.	Not a factual error. The ERG consider that it would be possible to adapt the model structure to account for different levels of response. Data on EASI75 response and discontinuation were reported in both this submission and TA534 and could be used to model different levels of response to baricitinib and dupilumab respectively. The ERG is aware that data on patients achieving ΔDLQI≥4 +

			EASI75 may not be available for dupilumab. However, the ERG consider that it is not an insurmountable issue, as it is likely that most patients who achieve EASI75 will also achieve ΔDLQI≥4.

Section 2: Minor comments

Issue 14

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 2 and 44 reports the discounted QALYs for dupilumab to be [redacted] in the "ERG Correction of Model Errors" analysis	Please update to: "[redacted]"	Minor typographical error	Amended

Issue 15

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 4, Subgroups row states: "*Furthermore, subgroups by region (Europe and Japan) were presented in Section 3.2.4.3, which may be considered a reasonable proxy for ethnicity. They reported a significant interaction, which indicate that*	Please consider amending to acknowledge the variety of potential explanations for the significant interaction observed in the subgroup analysis by region: "*Furthermore, subgroups by region (Europe and Japan) were presented in Section 3.2.4.3, which may be considered a reasonable proxy*	We appreciate that the ERG acknowledge that additional evidence presented in the response to clarification plausibly outlines alternative explanations for the observed differences between region subgroups. This additional context should be presented in	Table 4 – text added "**However, the evidence provided in the company's response to clarification suggests that the differences are not driven primarily by ethnicity, but**

outcomes for patients with different skin types are not the same."	*for ethnicity. They reported a significant interaction, which indicate that outcomes for patients with different skin types are not the same.* ***However, the evidence provided in the company's response to clarification suggests that the differences are not driven primarily by ethnicity, but rather by differences in the characteristics of the recruited patients and treatment practices.***"	Table 4, to comprehensively contextualise the conclusion that outcomes for patients with different skin types are not the same.	**rather by differences in the characteristics of the recruited patients and treatment practices**."

Issue 16

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 4, Special consideration row states: "*The British Association of Dermatologists state that effects on different skin type (e.g. BAME skin) should be considered as an equality issue for this indication.*"	Please consider amending to: "*The British Association of Dermatologists state that effects on different skin type**s** (e.g. BAME skin **types**) should be considered as an equality issue for this indication.*"	Whilst we acknowledge that the ERG are quoting The British Association of Dermatologists, a group of skin types are being referred to here. It is not appropriate to conflate skin types, and thus this statement should be corrected accordingly.	Amended

Issue 17

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 24 reports SAEs in the BARI 1 mg+ TCS group as: "[redacted]"	Please update to: "[redacted]"	Minor typographical error.	Amended

Issue 18

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 28 reports patients with ≥1 TEAE in the PBO arm of the BREEZE-AD2 (JAHM) trial as: "[redacted] *([redacted])*"	Please update to: "[redacted] *([redacted])*"	Minor typographical error	Amended

Issue 19

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 36, Induction Period to Treatment row states for NICE TA534: "*Patients who respond to ciclosporin at 16 weeks remain on the treatment until 1 year.*"	Please update to: "*Patients who respond to* ***dupilumab*** *at 16 weeks remain on the treatment until 1 year.*"	Minor typographic error	Amended

Issue 20

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Table 44 reports the baricitinib pack cost at PAS price as: £[redacted].	Please update to £[redacted] for each of the three times this is listed.	Minor typographical error	Amended

Issue 21

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG Response</th></tr>
<tr>
<td>Table 53 is reported as follows:
<table>
<tr><th></th><th>WTP threshold £20,000</th><th>WTP threshold £30,000</th></tr>
<tr><td>BSC</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Baricitinib</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Dupilumab</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Please update the table as follows:
<table>
<tr><th></th><th>WTP threshold £20,000</th><th>WTP threshold £30,000</th></tr>
<tr><td>BSC</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Baricitinib</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Dupilumab</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Apologies, this was a typographical error in the CS, which has been carried over into the ERG report.</td>
<td>Table amended</td>
</tr>
</table>

Issue 22

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16 states: *"However, in the CHRONOS trial, dupilumab reported greater flare suppression when compared with placebo (16% vs 52% respectively), significantly reducing the need for rescue therapy. This indicates that baricitinib treatment may not be effective at reducing flares (Section 3.4.2)."* Page 19 states: *"In the CHRONOS trial, dupilumab demonstrated flare suppression over long-term continuous use*	Please consider amending this wording to reflect that fact that there may be heterogeneity between the placebo arms of the CHRONOS and JAIN trials that is contributing to the observed difference in the flare suppression between dupilumab and baricitinib. *"However, in the CHRONOS trial, dupilumab reported greater flare suppression when compared with placebo (16% vs 52% respectively), significantly reducing the need for rescue therapy.* ***Whilst there may be heterogeneity between the placebo arms of the CHRONOS and JAIN trials that contributes to the observed difference in the***	The considerable difference in placebo response rates between the CHRONOS and JAIN trials indicates that there may be heterogeneity between the placebo arms of the two trials, which may be contributing to the observed difference in the flare suppression between dupilumab and baricitinib. This important context should be reported alongside the ERG's judgement.	Not a factual inaccuracy

versus placebo (16% vs 52% respectively), and thus a significantly reduced need for rescue therapy. The company assumed that baricitinib is equally effective as dupilumab with regards to flare control. However, in the JAIN trial, more patients in the baricitinib arm required rescue therapy for flares than did those on BSC, implying no substantial flare control is associated with baricitinib treatment (see Section 4.2.5.4)." Page 88 states: *"The ERG also notes that in the CHRONOS trial, dupilumab reported greater flare suppression when compared with placebo (16% vs 52% respectively), which significantly reduced the need for rescue therapy in patients treated with dupilumab. This indicates that baricitinib treatment is not as effective at controlling flares and reducing the need for high potency TCS."*	***flare suppression, these data indicate that, relative to dupilumab, baricitinib may not be as effective at controlling flares (Section 3.4.2)."*** Please consider adding relevant wording wherever this issue is raised in the ERG report.		

Issue 23

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 47 states: "*Intention-to-treat analysis, with non-responder*	Please amend to: *"Intention-to-treat analysis, with non-responder imputation (NRI) or* ***mixed***	Minor typographical error.	Amended

imputation (NRI) or Markov Chain Monte Carlo (MCMC) for missing data, was used for all analyses."	***model repeated measures (MMRM)*** *for missing data, was used for all analyses."*		

Issue 24

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 67 states: "*In the 1 mg baricitinib group [redacted]% and the 2 mg baricitinib group [redacted]% of patients experienced at least one TEAE.*"	Please update to: "*In the 1 mg baricitinib group [redacted]% and the 2 mg baricitinib group [redacted]% of patients experienced at least one TEAE.*"	Minor typographical error.	Amended

Issue 25

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 70 states: "At week 24, an additional [redacted] patients had discontinued due to AEs."	Please update to: *"At week 24, an additional [redacted] patients had discontinued due to AEs."*	Minor typographical error	Amended

Issue 26

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 76 states: "To maximise sample sizes, data were pooled for the baricitinib monotherapy studies JAHL and JAHM by pooling patients from the JAHL study with patients in the JAHM study's JAIN-like subgroup."	Please add additional context: *"To maximise sample sizes, data were pooled for the baricitinib monotherapy studies JAHL and JAHM by pooling* ***JAIN-like*** *patients from the JAHL study with patients in the JAHM study's JAIN-like subgroup."*	Additional context should be added for clarity.	Amended

Issue 27

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 74 states: "[redacted]% of patients in the 4 mg baricitinib arm discontinued due to AEs."	Please update to: "[redacted]*% of patients in the 4 mg baricitinib arm discontinued due to AEs."*	Minor typographical error	Amended

Issue 28

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 77 states: "*The company also did not provide separate baseline characteristics for the CHRONOS CAFÉ-like, SOLO1 CAFÉ-like or SOLO2 CAFÉ-like populations and they are not publicly available.*"	Please consider amending this wording to further acknowledge that these data were not available to us for presentation: "*The company also **could** not provide separate baseline characteristics for the CHRONOS CAFÉ-like, SOLO1 CAFÉ-like or SOLO2 CAFÉ-like populations **since** they are not publicly available."*	These data were not presented since the Company only had access to publicly available data from these trials. This wording is therefore unclear, and it may be mis-construed that these data were not presented due to a choice made by the Company.	Amended to "The company also did not provide separate baseline characteristics for the CHRONOS CAFÉ-like, SOLO1 CAFÉ-like or SOLO2 CAFÉ-like populations, **since they** are not publicly available"

Issue 29

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 124 states: "While the CS outlines the dosing and frequency of the components of BSC, the ERG notes that the proportions reported in Table 96 of the CS do not add up to 100%, perhaps suggesting an error."	Please remove this sentence.	The listed proportions in the CS Table 96 (33%, 25%, 17%, 15% and 10%) sum to 100%.	Sentence removed

Section 2: Confidentiality highlighting amendments

Issue 30

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 74 states: "The proportion of patients who experienced at least 1 TEAE was [redacted] in patients receiving 4 mg baricitinib ([redacted]%) compared to 2 mg baricitinib ([redacted]%), 1 mg baricitinib ([redacted]%) and placebo ([redacted]%). While [redacted] patients in the placebo and 1 mg baricitinib arms experienced AEs resulting in permanent discontinuation…"	Please update the confidentiality highlighting: *"The proportion of patients who experienced at least 1 TEAE was [redacted] in patients receiving 4 mg baricitinib ([redacted]%) compared to 2 mg baricitinib ([redacted]%), 1 mg baricitinib ([redacted]%) and placebo ([redacted]%). While [redacted] patients in the placebo and 1 mg baricitinib arms experienced AEs resulting in permanent discontinuation…"*	Minor typographical error	Amended

Issue 31

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Figure 6 Proportion of patients in BREEZE-AD4 (JAIN) achieving EASI75 over trial period (reproduced from PFC Response Figure 8a) And Figure 7 Proportion of patients in BREEZE-AD4 (JAIN) achieving a > 4-point improvement in DLQI score over the trial period (reproduced from CS PFC	Please add academic in confidence highlighting to these figures.	These figures present confidential data. Apologies that this was not made clear in the Company response to the clarification questions, since these figures were supplied in an embedded file.	Academic in confidence highlighting added to figures

response Figure 17a)			

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Technical report

Baricitinib for treating moderate-to-severe atopic dermatitis [ID1622]

This document is the technical report for this appraisal. It has been prepared by the NICE technical team.

The technical report and stakeholder's responses to it are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the appraisal committee meeting.

This report is based on:

- the evidence and views submitted by the company, consultees and their nominated clinical experts and patient experts and
- the evidence review group (ERG) report.

The technical report should be read with the full supporting documents for this appraisal.

1 Key issues summary

Issue	Summary	Technical Team Preliminary Judgement
1. Patient population	• The population in the company submission is 'adult patients with moderate-to-severe atopic dermatitis (AD) who are candidates for systemic therapy who have failed at least 1 systemic immunosuppressant due to intolerance, contraindication or inadequate disease control'. The company considers the main comparator to be dupilumab. • This is narrower than the population in the NICE scope, which is 'adults with moderate-to-severe AD who are candidates for systemic therapy that had an inadequate response or intolerance to existing topical treatments'. • The company has proposed the narrower population based on the likely positioning of baricitinib in clinical practice. It also matches the eligibility criteria of the company's main source of clinical effectiveness evidence, the BREEZE-AD4 (JAIN) trial. • Clinical expert advice to the ERG is that baricitinib would be offered to patients prior to dupilumab in clinical practice, at the same point as other immunosuppressants *(ERG report Section 1.1)*. • A submission from the National Eczema Society indicates that patients inadequately controlled on topical treatments would benefit from a new alternative treatment option to existing systemic immunosuppressants.	• Adults with moderate-to-severe AD who have not yet had systemic immunosuppressants are also a relevant population for baricitinib. **What additional information at engagement would help address the issue?** • Clinical advice on where baricitinib is most likely to fit in the treatment pathway.
2. Comparators	• The company's model has 2 comparators: best-supportive care (BSC) and dupilumab. The NICE scope includes 3 other comparators:	• In TA534, the expert stated that people would not be offered every available systemic immunosuppressant before being offered dupilumab. However, this does not

	o Phototherapy including ultraviolet B radiation or psoralen-ultraviolet A o Systemic immunosuppressive therapies (azathioprine, ciclosporin, methotrexate and mycophenolate mofetil) o Alitretinoin (in people with AD affecting the hands) • The ERG agrees that phototherapy and alitretinoin are not relevant comparators *(ERG report Table 4)*. • The ERG and an expert submission to NICE consider immunosuppressants a relevant comparator to baricitinib *(ERG report Table 4)*. The ERG notes that as several alternative systemic immunosuppressants are available, the failure of one should not mean that another cannot be used as second-line systemic therapy *(ERG report Section 4.2.3)*. • The company argues that in TA534 (dupilumab for treating moderate-to-severe atopic dermatitis), an expert advised the committee that in clinical practice, patients are unlikely to be offered every systemic immunosuppressant before dupilumab. As such, systemic immunosuppressants are not a relevant comparator for the narrower population in the company submission (see Issue 1). • The company states that systemic immunosuppressants could not be included as a comparator due to a lack of evidence on their relative efficacy in the patient population.	mean that more than 1 systemic immunosuppressant would not be offered. Furthermore, the company in TA534 provided a comparison with ciclosporin as a scenario. An analysis of patterns of systemic treatment in adults with moderate-to-severe AD in the UK indicates that methotrexate is commonly used as a second-line systemic immunosuppressant. • The technical team therefore considers systemic immunosuppressants a relevant comparator. **What additional information at engagement would help address the issue?** • The company is asked to provide a comparison of baricitinib with systemic immunosuppressants (azathioprine, ciclosporin, methotrexate or mycophenolate mofetil individually, or as a blended comparator). • Clinical advice on whether systemic immunosuppressants are a relevant comparator in people who have failed at least 1 systemic immunosuppressant due to intolerance, contraindication or inadequate disease control.
3. Disease severity of patient population	• The inclusion criteria in the company's clinical trials (JAIN, JAIY, JAHL and JAHM) defined moderate-to-severe AD as an Eczema Area and Severity Index (EASI) score ≥ 16, Investigator's Global Assessment (IGA) score ≥ 3 and body surface area (BSA) involvement ≥ 10% at screening.	• In TA534, dupilumab was appraised for treating moderate-to-severe AD, the same population as in the current baricitinib appraisal. The committee concluded that patients with the following scores reflected people who would be treated with dupilumab in the NHS:

	• The mean EASI scores in the 4 baricitinib trials ranged from [redacted] which is within published definitions of severe AD which range from 21.1 to 50. • Although there is no widely accepted definition of moderate disease the ERG notes that the company's EASI inclusion criteria excluded patients on the lower end of the moderate scale (EASI of 6-22.9) which may bias the population towards more severe disease *(ERG report Section 2.3.1)*. However, the ERG acknowledges that this issue also applied to the dupilumab CAFÉ and CHRONOS trials *(ERG response to company FAC, Issue 2)*. • The potential impact of the patient population in the baricitinib trials being skewed towards more severe disease is unclear. The company was unable to provide a scenario requested by the ERG in which separate utilities are applied for patients with moderate and severe disease.	o EASI: 34 o Dermatology Life Quality Index (DLQI): 15 o Patient Oriented Eczema Measure (POEM): 20 • The mean baseline EASI score for the baricitinib population included in the company's basecase indirect treatment comparison (ITC) was slightly lower than that of dupilumab ([redacted] versus 33.6). The mean baseline DLQI score was comparable ([redacted] for baricitinib versus 14.6 for dupilumab). The mean baseline POEM score was slightly higher ([redacted] for baricitinib versus 19.8 for dupilumab) *(Company submission Table 56, Company clarification response Table 7)*. • The baricitinib population in the company's ITC therefore largely aligns with what the committee previously agreed was an appropriate moderate-to-severe population for dupilumab, being slightly less severe in terms of baseline EASI score. However, there is uncertainty as to whether baricitinib and dupilumab would be used in the same patient populations (see Issue 1). **What additional information at engagement would help address the issue?** • Clinical advice on whether the population in the baricitinib trials reflects the patients in NHS practice who would be treated with baricitinib. • Clinical advice on whether the efficacy of baricitinib is likely to differ in moderate and severe patients.

4. Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice	• The company used a composite outcome of EASI50 plus a ≥ 4-point improvement in DLQI in its base-case cost-effectiveness analysis to define response. • The ERG is concerned that this outcome does not correlate with health-related quality of life (HRQoL) (see Issue 10). It is also unclear whether it is relevant in clinical practice *(ERG report Section 4.2.5)*. However, the ERG acknowledges that this outcome was considered appropriate for decision making in TA534 *(ERG response to company FAC, Issue 5)*. • The submission from the British Association of Dermatologists defines a clinically significant treatment response as a reduction in EASI score of 75%, or a fall in IGA of 2 points. • The ERG notes that the company's ITC showed that baricitinib has similar effectiveness to dupilumab in achieving a ≥ 4-point improvement on the Itch Numeric Rating Scale at week 16. This is considered to be a clinically meaningful improvement in itch symptoms in AD *(ERG report Section 4.2.5)*.	• The composite outcome of EASI50 plus a ≥ 4-point improvement in DLQI was considered appropriate for decision making by the committee in TA534. • In the ACD for TA534, clinical experts explained that EASI50 plus a ≥ 4-point improvement in DLQI is more sensitive to changes in treatment outcomes and more clinically relevant than EASI75 and IGA 0/1. • However, there appears to be some uncertainty over what is considered a clinically significant treatment response. • The technical team shares the ERG's concern that response defined by EASI50 plus a ≥ 4-point improvement in DLQI does not correlate with an HRQoL improvement based on the baricitinib trial data. **What additional information at engagement would help address the issue?** • Clinical advice on the relevance of this composite outcome (EASI50 plus a ≥ 4-point improvement in DLQI) in clinical practice in the NHS. • Clinical advice on whether any other outcomes (or composite outcomes) would be more clinically relevant.
5. Time to assessment of response	• In the company's model, response is assessed at 16 weeks. This aligns with the primary endpoint of JAIN (EASI75 at week 16), which the company considers reflects clinical practice in the NHS. • The ERG notes that the company's draft SmPC states that response to baricitinib should be assessed at 12 weeks. The clinical data for baricitinib	• The technical team agrees with the ERG that the baricitinib trial data appear to show a peak response at week 12. **What additional information at engagement would help address the issue?**

	show a peak response before week 12 across many outcomes *(ERG report Section 4.2.1)*. • Clinical opinion provided to the ERG indicates that response to baricitinib is likely to be assessed at 12 weeks, which is in line with other drugs used earlier in the treatment pathway. • The ERG is unable to present a scenario analysis with response assessed at 12 weeks due to a lack of data. However, the ERG considers that assessing response at 16 weeks may negatively impact baricitinib's cost-effectiveness, as non-responding patients would be incurring the cost of baricitinib for 4 weeks longer than necessary.	• Clinical advice on the timepoint at which response to baricitinib would likely be assessed in NHS clinical practice. • Clinical advice on any potential disadvantages of assessing treatment response to baricitinib at 12 weeks. • The company is asked to provide a scenario analysis of the impact of assessing response to baricitinib at 12 weeks.
6. Treatment sequencing	• The company's model currently assumes that patients transition from baricitinib or dupilumab to BSC upon discontinuation. The company does not provide a scenario where dupilumab is used after baricitinib, or vice versa. • The company does not consider the evaluation of treatment sequences relevant to the decision problem. The company considers that baricitinib will be used as an alternative to dupilumab, and that there is limited efficacy and safety data available for the 2 potential sequences (baricitinib after dupilumab, and dupilumab after baricitinib). • Clinical advice to the ERG is that baricitinib and dupilumab are likely to be used in a sequence, and that baricitinib may be used prior to dupilumab. The ERG notes that treatment sequencing is commonly modelled in appraisals in related therapy areas, such as psoriasis *(ERG report Section 4.2.3)*. • The ERG considers treatment sequencing relevant to the appraisal. • The ERG explored a sequence of baricitinib followed by dupilumab and vice versa *(ERG report section 6)*,	• Neither the NICE recommendation for dupilumab nor the company's proposed positioning for baricitinib precludes them from being used in a sequence. • There is plausible potential that clinicians would like to offer baricitinib and dupilumab in a sequence. • The technical team considers the sequence of baricitinib followed by dupilumab and vice versa to be relevant to the appraisal. **What additional information at engagement would help address the issue?** • Clinical advice on the most likely treatment sequence, assuming both baricitinib and dupilumab were available. • Clinical advice on whether the efficacy of baricitinib is likely to differ in people who have or have not previously received treatment with dupilumab, or vice versa.

	with the simplifying assumption that the effectiveness of baricitinib and dupilumab is not impacted by using them after one another. The sequence of baricitinib followed by dupilumab is more effective and less costly than dupilumab followed by BSC (i.e. it dominates dupilumab followed by BSC), and has an incremental cost-effectiveness ratio (ICER) of £57,034 versus BSC. Compared to baricitinib followed by BSC, baricitinib followed by dupilumab has an ICER of £90,446. All ICERs mentioned in the Technical Report are with the PAS applied for baricitinib, but not for dupilumab.	
7. Modelling of BSC	• In the company's model, patients transition onto BSC if they do not respond to induction treatment at week 16, or if they subsequently discontinue treatment. BSC is modelled as a separate line of therapy, with a 16-week induction in a similar way to baricitinib and dupilumab. Once patients transition onto BSC they quickly move into the subsequent 'non-response' health state, due to a high annual discontinuation rate of 57% from year 2 onwards (see Issue 8). The non-response state has poor HRQoL and a high cost compared to the BSC maintenance state (see Issue 10). Patients cannot then transition back into the previous states. • The ERG considers that this model structure does not accurately reflect the waxing and waning nature of AD, or clinical practice *(ERG Report Section 4.2.1)*. BSC is not intended to achieve disease control in the same way as baricitinib or dupilumab. Patients on BSC in reality are likely to have periods of good and bad disease control, whereas in the company's model they remain in a state of chronic and severe AD until death.	• The technical team agrees that the company's model does not capture the fact that patients having BSC are likely to have fluctuations in their HRQoL. • The technical team favours the ERG's scenario in which discontinuation from BSC is removed. This appears to more closely align with the week 16 – 52 data from the dupilumab CHRONOS study, and better reflects the fluctuating nature of disease control on BSC. However, the technical team notes that the ERG's scenario does not take into account a reduction in BSC efficacy outside of the clinical trial setting, when treatment adherence is likely to decrease. • The technical team agrees with removing the costs of BSC to avoid duplication and align with the approach taken in TA534. **What additional information at engagement would help address the issue?** • Clinical advice on which approach to modelling BSC is most appropriate.

	• In TA534 BSC was modelled as a single health state, with a weighted utility value. It was assumed that a proportion of the health benefit associated with BSC would be lost each year (96% - 97% by year 5 in the committee's preferred analyses). This was because the effect size for BSC was considered unlikely to persist outside of the clinical trial setting due to decreased treatment adherence (see section B3.3.6 of the dupilumab company submission). Patients losing benefit would revert to the baseline utility of 0.66, while patients retaining benefit had a BSC average utility value of 0.77. • The ERG highlights data from the placebo arm of the CHRONOS trial (dupilumab), which shows that *(ERG report Section 4.2.1)*: o 37% of patients achieved EASI50 by week 16 o 18.7% of these patients then lost these improvements by week 52, as shown by the conditional probability of an EASI50 response at 52 weeks in Table 3.5, page 179 of the dupilumab company submission o However, there was less of a difference in the overall proportion of patients in the CHRONOS BSC arm achieving EASI50 at week 16 (37%) compared to week 52 (30%). • These data suggest that for every patient losing disease control another achieves symptom improvement. The data do not support the assumption that patients receiving BSC have poor HRQoL indefinitely. • The ERG presents a scenario in which the discontinuation rate on BSC is set to 0%. In this scenario the proportion of patients on BSC with the higher utility value and lower costs at any one time following induction is equal to the BSC response rate	

	at week 16 (31.25%). This is intended to reflect the CHRONOS data above, where fluctuations in patients on BSC losing and gaining disease control meant that the proportion achieving EASI50 was approximately consistent from week 16 - 52. This scenario increases the ICER for baricitinib versus BSC by £11,599 compared with the company's base case. • Further, the ERG notes that the company applied a drug acquisition cost for BSC, in a similar manner to baricitinib and dupilumab. This was a weighted cost based on a combination of topical corticosteroids (TCS), topical calcineurin inhibitors (TCI) and oral corticosteroids, equating to £14.73 per week. The company also applied the cost of concomitant medications to BSC, with concomitant medications comprising of bathing products, emollients, TCS and TCI. • As there is an overlap between the composition of BSC and the composition of concomitant medications, more than 100% of responders to first-line BSC receive mometasone and tacrolimus in the company's model *(ERG report Section 4.2.7.1)*. • The ERG favours a scenario in which the drug acquisition cost for BSC are removed. This aligns with the approach taken in TA534, and also avoids some of the duplication of medication costs. In this scenario, the ICER for baricitinib versus BSC is increased by £4,812 compared with the company's base case.	
8. Long-term discontinuation rates for baricitinib	• The company models the probability of a patient discontinuing baricitinib treatment using discontinuation rates in the dupilumab company submission (TA534) because long-term discontinuation rates for baricitinib could not be	• The technical team considers that the company has not provided sufficient justification for assuming equivalence between baricitinib and dupilumab for long-term discontinuation.

	reliably estimated from the JAIN data. Baracitinib is therefore assumed to have equivalent discontinuation rates to dupilumab. • In the first year (following initial response, assessed at week 16) the discontinuation rate for patients on baricitinib in the company's model is 6.1%. The discontinuation rate for patients on BSC in the first year is 23.3%. These numbers were taken from Table 3.5, page 178 in the dupilumab company submission, which provides the probability of a sustained response at week 52 in patients responding at week 16 in the dupilumab CHRONOS trial. • The annual discontinuation rate is 3.7% from year 2 (week 53) onwards for patients on baricitinib in the company's model. This was taken from Table 3.6, page 178 in the dupilumab company submission, which provides the proportion of patients withdrawing from the CHRONOS study in the 52-week treatment period among those responding at week 16. • The annual discontinuation rate from year 2 onwards for patients having BSC in the company's model is 57%. This is based on the annual probability of study withdrawal or use of rescue medication from the BSC arm in the CHRONOS trial. • The ERG notes that the baricitinib trial data do not support equivalence for baricitinib with dupilumab for discontinuation rates from week 16-52 *(ERG report Section 4.2.5.2)*: o By week 24 in JAIN, baricitinib was no longer significantly better than placebo in the following outcomes: EASI50/75/90, IGA≥1, DLQI≥4, and SCORAD75. The ERG therefore questions the validity of assuming significantly	• The technical team agrees with the ERG that the available baricitinib trial data do not support equivalence for baricitinib with dupilumab for discontinuation rates from week 16-52. • The technical team prefers the ERG's scenarios, with the long-term discontinuation rates drawn from JAHN. **What additional information at engagement would help address the issue?** • Additional data cuts from JAIN and JAHN (e.g. an interim analysis) to reduce the uncertainty around the long-term discontinuation rates. • Clinical advice as to whether it is appropriate to assume equivalence for baricitinib and dupilumab in long-term discontinuation rates, and if not what discontinuation rates are most appropriate given the available data.

	diverging long-term efficacy for baricitinib compared with BSC. o Percent improvement in EASI score from baseline for baricitinib decreased from [redacted] ([redacted]) at week 16 to [redacted] at week 24 in JAIN ([redacted]). Conversely in CHRONOS, the mean change in EASI score from baseline for dupilumab increased from 76.7% at week 16 (p<0.0001 versus placebo) to 78.3% at week 52 (p<0.0001 versus placebo). o Of the patients in JAIN that had not discontinued at week 16, [redacted] in the 4mg baricitinib arm went on to discontinue at week 24. This equates to a discontinuation probability of [redacted] from week 16-52. o In the BREEZE-AD3 (JAHN) monotherapy extension study, [redacted] of patients who responded at week 16 were still responding at week 52. This equates to a discontinuation probability (due to loss of response) of [redacted] from week 16-52. • The ERG notes that baricitinib and dupilumab have different mechanisms of actions and routes of administration, and are therefore likely to differ in long-term efficacy and adherence. • The ERG explored 3 scenarios where the longer term discontinuation rates from JAHN were applied for baricitinib rather than the values from TA534. The ERG's base case includes a scenario where the discontinuation rates for patients on baricitinib from week 16-52 and week 53 onwards are based on the all-cause discontinuation rates from JAHN at week 36 (a per-cycle discontinuation rate for bariticinib of [redacted]). In this scenario, the ICER for baricitinib versus	

	BSC is increased by £2,577 compared with the company's base case.	
9. Loss of utility benefit (consistency with TA534)	• The company base case model in TA534 assumed that patients receiving dupilumab maintenance would lose 2% of the treatment benefit in year 2, 5% in year 3, 7% in year 4, and 8% in year 5. It used these estimates to adjust down the proportion of people who continued to have dupilumab. This was applied in addition to the annual discontinuation rate of 3.7%, based on the all-cause discontinuation rate from CHRONOS (see Issue 8). • These assumptions were based on feedback from the experience of 5 dupilumab trial investigators, supported by evidence from its open-label extension study that showed a sustained treatment effect for dupilumab. In response to the ACD, the company in TA534 highlighted that 92.5% of patients treated with dupilumab achieved EASI50 plus a ≥ 4-point improvement in DLQI at week 48. At week 76 this figure was 93.8%, and at week 100 this figure was 97.6%. • In the baricitinib company submission, no such treatment waning effect is assumed for baricitinib.	• Given that the company applies the dupilumab discontinuation rates from TA534 in the current appraisal, the technical team considers that a scenario should also be presented in which the treatment waning assumption from TA534 is also applied. **What additional information at engagement would help address the issue?** • Clinical advice as to whether the assumptions around the loss of quality of life benefit in TA534 for dupilumab should also be applied for baricitinib.
10. Utility values	• In the company's model, patients achieving a response at week 16 (EASI50 plus a ≥ 4-point improvement in DLQI) are assigned a utility of 0.7800. Patients during induction, patients not responding to BSC, or patients subsequently discontinuing from BSC are assigned a utility of 0.5979. The company did not break the utilities out further based on magnitude of response, and applied the same utilities regardless of treatment received. • The company derived the utility values above from data collected from the JAIN and BREEZE-AD7 (JAIY) trials. The company did not include the data	• The technical team shares the ERG's concern that response status as defined in the company's model does not correlate with improved HRQoL based on the baricitinib trial data. • The technical team prefers the ERG's revised scenario in which the utility values are drawn from TA534, as these appear more plausible than those of the company. **What additional information at engagement would help address the issue?**

	from BREEZE-AD1 (JAHL) and BREEZE-AD2 (JAHM) in its base-case analysis, as in these trials patients had bariticinib as a monotherapy. • The improvement in EQ-5D-3L at week 16 was higher in non-responders (0.2042) than responders (0.1821) based on the data from JAIN and JAIY. This was also consistent with the results if the data from JAHL and JAHM are included ([redacted] increase for non-responders versus [redacted] for responders). The company did not consider it clinically plausible for non-responders to have a higher utility than responders, so applied the baseline utility to non-responders. • The ERG is concerned that the company's definition of response (EASI50 plus a ≥ 4-point improvement in DLQI) appears to confer little or no HRQoL improvement based on trial data. However, the ERG notes that the sample size of the responder group ([redacted]) was small. In addition, a similar pattern was seen in TA534 *(ERG report Section 4.2.6)*, in that there was little difference in the utility value for all patients at week 16 (0.891) compared to week 16 responders (0.898). • The ERG considers that the current model structure with only 2 utility values may bias the model in favour of less efficacious treatments *(ERG report Section 1.3)*. A model structure with different levels of response would have better captured HRQoL differences between treatments. The ERG notes that while a simple 2-state approach was used in TA534, different utilities were applied for dupilumab compared with BSC which would better capture differences between treatments. In TA534, a utility value of 0.898 was assigned to dupilumab	• Clinical advice on whether a response as defined by EASI50 plus a ≥4 point improvement in DLQI would confer a benefit in HRQoL. • Clinical advice on whether the utility values of the company or ERG are more appropriate.

	responders, while a weighted utility value was assigned to patients on BSC as described in Issue 7. • The ERG considers that the data from JAHL and JAHM patients should be included to generate the utility values, although acknowledges that patients in JAHL and JAHM did not receive concomitant TCS *(ERG response to company FAC, Issue 7)*. • The ERG notes that no patients in the JAIN-like JAIY population received dupilumab, yet the company considered the values relevant to these patients *(ERG report Section 4.2.6)*. • The ERG favours a scenario in which utility values are derived from TA534. Patients on maintenance baricitinib and dupilumab are assigned a utility value of 0.898. Patients during induction are assigned a utility value of 0.66. Patients on maintenance BSC, or non-responders to baricitinib or dupilumab, are assigned a utility value of 0.797. In this scenario, the ICER for baricitinib compared with BSC is increased by £15,455 compared with the company's base case, to £33,451.	

2 Other issues

Issue	Summary	Technical Team Preliminary Judgement
11. Differences in clinical outcomes in Japanese patients	• A greater proportion of European patients achieved EASI75 with baricitinib compared with placebo in JAIY (risk ratio (RR) [redacted]) compared with those from Japan (RR [redacted]) and the rest of the world (ROW) (RR [redacted]). Around [redacted] of the people who had baricitinib in the JAIN plus JAIN-like JAIY population were from Japan.	• The technical team are concerned that the data from the Japanese patients recruited in the baricitinib trials are not generalisable to clinical practice in the NHS. The company used the data from the full pooled JAIN plus JAIN-like JAIY population in its base-case analysis (including patients from Europe, Japan and ROW).

	• The company considers that the greater response in European patients was due to:
	  o Higher rates of rescue therapy in Japan and ROW compared with Europe. Japanese patients are more likely to be rescued with high-potency TCS than European patients. Patients having rescue therapy were censored (meaning that they were assumed to be a non-responder), leading to higher response rates in European patients.
	  o Higher response rates in the placebo arm, possibly due to an insufficient use of TCS prior to study enrolment in Japan and ROW.
	  o Ethnic differences in the prevalence of certain cytokine pathways, although the impact on treatment response is unclear.
	• Japanese patients had more severe AD (had a higher EASI score, SCORAD score and BSA) at baseline than European patients. The ERG concludes that the differences in treatment efficacy in Japanese patients are likely to be based on a combination of differences in baseline disease severity and clinical practice, rather than ethnicity *(ERG response to company FAC, Issue 15)*. However, this is an area of uncertainty, and may limit the generalisability of data from Japanese patients to NHS clinical practice.
	• [redacted].
	The company stated that the trial program was not designed to investigate baricitinib efficacy in black patients compared with other patient populations, but noted that there is some evidence that the pathology of AD could be more severe and persistent in black patients.The ERG notes that it may not be appropriate to assume that the efficacy results
		• The technical team agrees with the company that very few black patients were recruited in CAFÉ and CHRONOS and that skin colour was not raised as an issue in TA534, other than in the context of using the EASI.
		• In TA534 the committee noted that the EASI might underestimate the severity of atopic dermatitis in people with darker skin and concluded that healthcare professionals should take into account skin colour and how this could affect the EASI score when following NICE's recommendations.
		
		What additional information at engagement would help address the issue?
		• Clinical advice on whether the data from the Japanese patients in the baricitinib trials is generalisable to NHS clinical practice, and whether the data from European patients only may be more clinically relevant.
		• Clinical advice on whether it is reasonable to assume comparable efficacy for baricitinib in black patients in the absence of data.

	observed in white patients are transferable to other ethnicities. This is a concern given the greater prevalence and severity of AD in the black British population *(ERG report Section 2.1)*. • The company notes that few black patients were recruited in the trials for dupilumab (CAFÉ and CHRONOS), which the ERG agreed with *(ERG response to company FAC, Issue 4)*.	
12. Indirect treatment comparison heterogeneity	• There is no direct evidence comparing baricitinib with dupilumab. The company submitted an ITC using data from 4 trials in its base case: o JAIN and JAIY (JAIN-like patients): baricitinib plus TCS versus placebo plus TCS. o CAFÉ and CHRONOS (CAFÉ-like patients): dupilumab plus TCS versus placebo plus TCS. • The ERG notes the following differences between the trials included in the ITC, increasing heterogeneity and reducing the reliability of the ITC results *(ERG report Section 2.1)*: o There was a higher proportion of Asian patients in JAIN ([redacted]) compared with CAFÉ (2%). Geographic region may be an effect modifier (see Issue 11). o Entry into the baricitinib trials required an EASI score ≥ 16, while entry into CAFÉ required an EASI score ≥ 20. The baseline EASI scores were higher in the dupilumab trials (patients had more severe AD) (see Issue 3). o Patients were permitted to apply TCS in the 2-weeks prior to randomisation in the dupilumab trials, while the baricitinib trials had a 2-week washout period for topical treatments.	• The technical team considers that the patient populations in the company's indirect comparison were reasonably comparable. • The technical team notes that while the baseline EASI scores were higher in the dupilumab trials, these patients may also have been less likely to experience a flare. **What additional information at engagement would help address the issue?** • Clinical advice on the extent to which the differences between the patient populations and censoring rules in the baricitinib and dupilumab trials are likely to bias the results of the ITC.

	Patients in the dupilumab trials may therefore have been less likely to experience a flare. o Different secondary censoring rules were applied in the baricitinib and dupilumab trials. In the baricitinib trials, data were subject to secondary censoring after permanent study drug discontinuation or after initiation of rescue therapy with systemic treatments (but not TCS). In the dupilumab trials, all observed data regardless of rescue treatment was used, including data collected after withdrawal (see page 16 of the dupilumab company submission). This may favour dupilumab *(ERG report Section 3.3)*, as data from additional rescued patients may be included, rather than categorised as a non-responder.	
13. Impact of baricitinib on flare control	• The company have assumed that baricitinib is equally effective as dupilumab with regards to flare control, with an annual flare rate of 0.18 compared to 0.78 for patients having placebo (from TA534). The company considers these estimates to be most plausible, as long-term baricitinib data is pending. • The ERG does not agree because while flare frequency was not recorded in the baricitinib trials, the receipt of rescue medication can be considered a proxy for flare *(ERG report Section 3.2.5.6)*. At week 24 in JAIN, more patients in the baricitib had rescue medication ([redacted] compared to [redacted] in the placebo arm). • In the CHRONOS trial at week 52 fewer patients in the dupilumab arm had rescue medication compared with the placebo arm (16% versus 52%). The company considers that this may be the result of heterogeneity between the placebo arms of the JAIN and CHRONOS trials, but the ERG considers that	• The company has not provided sufficient justification for assuming equivalence between baricitinib and dupilumab for flare control. • The technical team agrees with the ERG that the available baricitinib trial data do not support equivalence for baricitinib with dupilumab for flare control. **What additional information at engagement would help address the issue?** • The technical team requests that the company provide the additional data on long-term flare data from JAIN as soon as possible to reduce the uncertainty on this point. • Clinical advice on whether the flare control assumptions of the company (equivalence to

	this may be because dupilumab has greater flare suppression than baricitinib. • An ERG scenario in which flare rates for baricitinib are set to be equivalent to those for placebo had a minimal impact on the ICERs for baricitinib compared with dupilumab and BSC.	dupilumab) or ERG (equivalence to BSC) are more appropriate.
14. Concomitant treatments	• The company modelled concomitant treatment to include bathing products, such as aqueous creams and shower emollients. • Based on TA534, the company assumed that responders have a 50% reduction in concomitant treatment compared to non-responders. • The ERG understands that there has been a significant reduction in the use of bathing products in clinical practice since TA534 *(ERG report Section 4.2.7.1)*. As such, the ERG explored a scenario in which the costs of bathing products were removed from the model. This has minimal impact on the ICERs for baricitinib compared with dupilumab and BSC. • The ERG also notes that while a 50% reduction in concomitant treatment for responders was accepted in TA534, this was largely based on expert opinion with little supportive evidence. However, clinical advice to the ERG indicates that the company's assumptions were reasonable.	**What additional information at engagement would help address the issue?** • Clinical advice on the use of comcomitent bathing products in NHS practice to manage moderate to severe AD, to validate the ERG's conclusions. • Clinical advice on the reduction in concomitant treatments seen in NHS practice in people who respond to treatment.

3 Questions for engagement

Issue 1. Patient population

1. The company's proposed population is 'adult patients with moderate-to-severe atopic dermatitis (AD) who are candidates for systemic therapy who have failed at least 1 systemic immunosuppressant due to intolerance, contraindication or inadequate disease control'. Are there any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe AD who have not yet had a systemic immunosuppressant?

Issue 2. Comparators

2. Approximately what proportion of people would be offered the following treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants in the tables below (due to intolerance, contraindication or inadequate disease control)? Please provide a percentage for each treatment.

Following first-line **ciclosporin**	
Dupilumab (%)	
Azathioprine (%)	
Methotrexate (%)	
Mycophenolate mofetil (%)	

BSC (%)	
Other (%)	

Following first-line **methotrexate**	
Dupilumab (%)	
Azathioprine (%)	
Ciclosporin (%)	
Mycophenolate mofetil (%)	
BSC (%)	
Other (%)	

Following first-line **azathioprine**	
Dupilumab (%)	
Ciclosporin (%)	
Methotrexate (%)	
Mycophenolate mofetil (%)	
BSC (%)	
Other (%)	

Issue 3. Disease severity of patient population

3. Do you consider the patient population in the baricitinib trials to be generalisable to people with moderate-to-severe AD in the NHS?

Issue 4. Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice

4. Is EASI50 plus a ≥ 4-point improvement in DLQI a relevant outcome for determining treatment response in NHS clinical practice? Are there any other outcomes (or composite outcomes) that should be considered?

Issue 5. Time to assessment of response

5. Do you anticipate that response to baricitinib would be assessed at 12 or 16 weeks in NHS clinical practice?

6. Would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks?

Issue 6. Treatment sequencing

7. Assuming both baricitinib and dupilumab were available, how do you anticipate that they would likely be used in NHS practice? Please provide a rationale for your response.

    - Baricitinib followed by dupilumab
    - Dupilumab followed by baricitinib
    - No treatment sequencing (i.e. baricitinib or dupilumab followed by BSC)

8. Do you anticipate that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment? If so, why?

Issue 7. Modelling of BSC

9. Which of the following approaches (or combination of approaches) do you consider most appropriate for modelling BSC?

    - Company: an initial response rate of 31.25% at week 16, followed by a discontinuation rate of 23.3% from week 16 – 52, and a 57% annual discontinuation rate from then on. Patients not responding to BSC, or subsequently discontinuining from BSC, have a lower utility and higher costs than the remaining patients on BSC
    - ERG: an initial response rate of 31.25% at week 16, followed by no discontinuation from week 16 – 52, and no discontinuation annually from them on. The proportion of patients on BSC with a higher utility and lower costs is therefore constantly equal to the BSC response rate at week 16

- TA534 Sensitivity analysis 1: the proportion of patients on BSC losing quality-of-life benefit (and having a lower utility value) is as follows: Year 2: 82%, Year 3: 90%, Year 4: 94%, Year 5 and beyond: 96%

- TA534 Sensitivity analysis 2: the proportion of patients on BSC losing quality-of-life benefit (and having a lower utility value) is as follows: Year 2: 57%, Year 3: 82%, Year 4: 92%, Year 5 and beyond: 97%

Issue 8. Long-term discontinuation rates for baricitinib

10. Is the company's assumption of equivalence between baricitinib and dupilumab for long-term discontinuation rates reasonable, in the absence of longer-term data for baricitinib?

11. The ERG used data from JAHN to derive alternative discontinuation rates. Should the discontinuation rate from week 16-52 be based on the probability of response at week 52 conditional on response at week 16, or on all-cause discontinuation rates?

Issue 9. Loss of utility benefit (consistency with TA534)

12. Should an assumption be incorporated into the model that a proportion of patients discontinue baricitinib each year due to a loss of utility benefit, in addition to the all-cause discontinuation rates already applied? If so, is it reasonable to assume equivalence with dupilumab? (i.e. 2% of patients losing HRQoL benefit in year 2, 5% in year 3, 7% in year 4, 8% in year 5 and beyond)

Issue 10. Utility values

13. Do you anticipate that a patient achieving a response as defined by EASI50 plus a ≥4 point improvement in DLQI would experience an improvement in HRQoL?

14. Do you consider the utility values preferred by the company or ERG (derived from TA534) to be more appropriate?

- Company assumption: induction and non-response utility of 0.5979, response utility of 0.7800 (applied for all treatments)

- ERG assumption: induction utility of 0.66 (applied for all treatments), response utility of 0.898 for baricitinib / dupilumab and 0.797 for BSC, non-response utility of 0.797 (applied for all treatments)

Issue 11. Differences in clinical outcomes based on region and skin type

15. Do you consider the data from the Japanese patients recruited in the baricitinib trials to be generalisable to clinical practice in the NHS?

16. Would data from European patients only be more clinically relevant than the full intent-to-treat population?

17. Is it reasonable to assume comparable efficacy for baricitinib in black patients compared to white patients in the absence of robust data? If not, what adjustments should be made when assessing the effectiveness of treatment in black patients?

Issue 12. Indirect treatment comparison heterogeneity

18. Do you consider that the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the ITC in favour of either treatment?

Issue 13. Impact of baricitinib on flare control

19. Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable, in the absence of longer term data for baricitinib?

Issue 14. Concomitant treatments

20. Does the ERG's revised scenario in which the costs of bathing products are removed from the model best reflect current NHS clinical practice?

21. Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?

NICE National Institute for Health and Care Excellence

Technical engagement response form

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments **13 November 2020**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential

information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Eli Lilly & Company**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Patient population	
Are there any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe atopic dermatitis (AD) who have not yet had a systemic immunosuppressant?	• The company do not consider any further populations to be relevant for consideration in this appraisal due to: o **The expected positioning of baricitinib in the treatment pathway.** As stated in the company submission, baricitinib is expected to be positioned in UK clinical practice for the population in which the highest unmet clinical need exists: adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This is also aligned with the eligibility criteria for the main source of clinical effectiveness evidence, BREEZE-AD4 (JAIN), and with the population for which dupilumab was recommended by NICE in TA534.[1, 2] This population is narrower than the full marketing authorisation for baricitinib, which is for the treatment of moderate-to-severe atopic dermatitis (AD) in adult patients who are candidates for systemic therapy.[3] The company acknowledge that there is an unmet need in UK clinical practice for alternatives to current systemic immunosuppressants, which are associated with poor safety profiles. This is reflected by the submission from the National Eczema Society which indicated that patients inadequately controlled on topical treatments would benefit from a new alternative treatment option. The company therefore also understand the NICE Technical Team's position that adults with moderate-to-severe AD who have not yet had systemic immunosuppressants may be a relevant population for baricitinib. The comparators in this setting include current systemic immunosuppressants such as ciclosporin, methotrexate, azathioprine and mycophenolate mofetil, although ciclosporin alone of these holds marketing authorisation for use in this population. o **The paucity of evidence to conduct a robust indirect comparison of baricitinib versus systemic immunosuppressants.** Given that most systemic immunosuppressants used in AD are being used off-license, the identification in the clinical SLR of sparse evidence for systemic immunosuppressants in AD was as expected. In the absence of a comparative, double-blind, parallel, placebo-controlled study of ciclosporin versus dupilumab, an indirect comparison versus ciclosporin was attempted in a scenario analysis in TA534, but it should be acknowledged that this was considered by the ERG to not be robust.[2] Similarly, the company attempted to perform a matching-adjusted indirect comparison (MAIC) of ciclosporin versus baricitinib using

data from the BREEZE-AD7 JAIY trial and evidence on the efficacy of ciclosporin identified in the literature, which resulted in an effective sample size for baricitinib (following the application of weights) of [redacted] patients. Therefore, valid weighted comparisons of responses between baricitinib and ciclosporin could not be derived and a robust scenario analysis in the full population could not be conducted. Full details on the methodology and results of the MAIC are presented in the Company Technical Engagement Appendix.

- **The availability of direct evidence to inform a robust indirect comparison of baricitinib versus dupilumab and BSC.** Direct evidence for the relative efficacy of baricitinib versus BSC in patients who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control is available from the BREEZE-AD trials (placebo, with or without topical corticosteroids, can be considered a proxy for BSC).[1, 4-7] Similarly, post-hoc pooled analyses presented in TA534 facilitated a robust ITC of baricitinib versus dupilumab in this population.[2]
- **The unmet treatment need in the target population.** As discussed in the response to Issue 2, the population specified in the company submission represents those patients whose only remaining treatment options are dupilumab or best supportive care (BSC). Dupilumab is recommended by NICE for adults with severe-to-moderate AD who experience failure with, are intolerant to or have contraindication to at least one systemic therapy.[2] Whilst dupilumab is effective in controlling the disease, there are considerable limitations to its use. Dupilumab is administered via subcutaneous injection every other week. Many patients experience injection site reactions, with over 1 in 10 patients experiencing swelling at the site of injection, and more than 1 in 100 reporting redness, pain or itch at the injection site.[3] Eye disorders such as conjunctivitis are also common adverse events of dupilumab treatment. In the CAFÉ trial, 28% patients receiving dupilumab (every other week in combination with topical corticosteroids) experienced conjunctivitis, which was severe in 0.9% and moderate in 12.1% patients.[8] These adverse events result in additional health care resource use through the need for consultant ophthalmologist visits. If dupilumab fails to control the disease, or in patients for whom use of dupilumab is not recommended or contraindicated, no further safe and effective treatment options are available so patients are treated with BSC. There is a clear unmet clinical need for an effective, tolerable, easily-administered treatment option for those patients whose only alternatives are dupilumab or BSC.

- Given that baricitinib has the opportunity to simplify the treatment paradigm and address the considerable unmet need in this population, and that sufficient evidence is available to generate robust estimates of relative efficacy versus the relevant comparators, this was considered the most suitable position for baricitinib to be able to demonstrate cost-effectiveness and facilitate patient access to this innovative therapy in current UK practice. This

	is reflected in the company submission and the additional evidence provided in the Company Technical Engagement Appendix, which demonstrates that baricitinib is a cost-effective treatment option in this population.

Issue 2: Comparators

Approximately what proportion of people would be offered the following treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants in the tables below (due to intolerance, contraindication or inadequate disease control)? Please provide a percentage for each treatment.

In order to answer this question, Lilly sought the expert clinical opinion of a consultant dermatologist practicing in the UK. The feedback received was is that the approximate patient proportions presented in Table 1 would be treated with these systemic treatments following first-line failure with ciclosporin, methotrexate and azathioprine. The company note that feedback from the clinical expert is that azathioprine is very rarely used as a first-line treatment in general UK clinical practice.

Table 1: Proportion of patients offered various treatments in UK clinical practice following failure of three first-line treatment

Treatment	Proportion of patients treated (%)		
	Following first-line ciclosporin	Following first-line methotrexate	Following first-line azathioprine
Dupilumab	[redacted]	[redacted]	[redacted]
Azathioprine	[redacted]	[redacted]	[redacted]
Methotrexate	[redacted]	[redacted]	[redacted]
Mycophenolate mofetil	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]

- The company do not consider systemic immunosuppressants to be relevant comparators for baricitinib due to:
  - **The expected positioning of baricitinib in the treatment pathway.** As stated in the company submission and discussed further in response to Issue 1, baricitinib is expected to be positioned in UK clinical practice for patients whose only remaining treatment options are dupilumab or BSC, making dupilumab and BSC the only relevant comparators for this appraisal.

	○ **The positioning of dupilumab in the treatment pathway.** In TA534, dupilumab was recommended *"as an option for treating moderate to severe atopic dermatitis in adults, only if the disease has not responded to at least 1 other systemic therapy, such as ciclosporin, methotrexate, azathioprine and mycophenolate mofetil, or these are contraindicated or not tolerated."*[2] The company did not include ciclosporin as a comparator, with the justification that the evidence base of dupilumab compared to ciclosporin is sparse and that the treatments would not, in any case, occupy the same place in the treatment pathway.[2] The decision problem addressed by the company specified the relevant comparator as best supportive care, which was accepted by both the ERG and the NICE committee. Baricitinib is positioned as an alternative to dupilumab in this submission, and the population of relevance is in line with the recommendation for dupilumab in TA534. Thus, the relevant comparators in this population are dupilumab and BSC, and do not include systemic immunosuppressants. ○ **The treatment length of baricitinib versus systemic immunosuppressants.** Baricitinib is a long-term treatment option, unlike systemic immunosuppressants, which tend to be used for short-term episodic use due to their severe side effects. According to the treatment pattern survey by Taylor *et al.* (2016), the average length of a course of ciclosporin treatment is 5.8 months.[9] The average lengths of courses of azathioprine and methotrexate (which are not licensed for the treatment of AD in the UK) are 13.8 and 15.1 months, respectively.[9] Dupilumab therefore represents a more appropriate comparator for baricitinib. ○ **A lack of evidence to inform a comparison of systemic immunosuppressants versus baricitinib.** As discussed further in response to Issue 1, a scenario analysis informed by a MAIC of ciclosporin versus baricitinib was found not to be feasible. Similarly, it should be noted that the scenario analysis of comparison with ciclosporin in the full license population presented in the dupilumab appraisal was considered by the ERG to be not robust (TA534 Committee Papers, Pre-Meeting Briefing, page 15).[2]
Issue 3: Disease severity of patient population	
Do you consider the patient population in the baricitinib trials to be generalisable to people with moderate-to-severe AD in the NHS?	• The patient population in the BREEZE-AD baricitinib trials is generalisable to people with moderate-to-severe AD in the NHS: ○ **Representative baseline characteristics.** The baseline characteristics of patients in the BREEZE-AD trials were acknowledged by the ERG (ERG report, page 14) and the NICE Technical Team to be representative of patients in UK clinical practice and the company welcome the conclusion of the Technical Team that the baricitinib population in the ITC presented by the company therefore largely aligns with the population previously agreed by the committee to be an appropriate moderate-to-severe population for dupilumab. ○ **The distinction between moderate and severe AD is not well defined.** No definitive and widely accepted cut-off exists between moderate and severe AD, with classification based on a variety of clinical factors in UK clinical practice. As per the conclusions of a steering committee consisting of a multidisciplinary group of AD

	experts, including 8 dermatologists, 2 allergists, and a patient advocacy group representative, the factors for diagnosis of moderate-to-severe AD include consideration of body surface area (BSA) affected, individual lesion severity, lesion location and/or quality of life impairment.[10] Standardised scales, such as EASI score, can provide one classification system, but do not capture all of these criteria and therefore may not be reflective of all aspects of disease, particularly given that this scoring could provide an inconsistent classification of a flaring disease. This is a potential source of significant heterogeneity within clinical practice and ultimately, the steering committee concluded that these disease severity scales may not be practical for routine use in clinical practice for the classification of disease severity.[10] This conclusion is supported by evidence from a UK-based trial[9] and by advice received from UK clinicians consulted by Eli Lilly which suggests that clinical experience, rather than severity scales in isolation, is likely to inform UK clinical practice. o **A subgroup analysis is not feasible.** In addition to the clinical limitations of disease classification discussed above, analysis of moderate versus severe subgroups are prevented by a lack of available data for these populations for dupilumab. Together, these clinical and data limitations mean that it is not feasible to conduct any efficacy comparisons with the key comparator in these populations, as noted in response to ERG clarification question A8. • Therefore, the company conclude that the patient populations of the BREEZE-AD trials are representative of patients in UK clinical practice. The company did not consider analysis of a moderate versus severe subgroup based on EASI category to be clinically relevant, as it is an insufficiently holistic measure of severity, or practically feasible.
Issue 4: Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice	
Is EASI50 plus a ≥ 4-point improvement in DLQI a relevant outcome for determining treatment response in NHS clinical practice? Are there any other outcomes (or composite outcomes) that should be considered?	• The company initially selected EASI50 plus a ≥4-point improvement in DLQI as the response definition to align with the definition accepted by the Appraisal Committee as clinically relevant during the dupilumab NICE appraisal (TA534).[2] • Following feedback from the ERG and the NICE Technical Team, the company consider EASI75 to be a more relevant outcome for determining treatment response in NHS clinical practice than EASI50 plus a ≥ 4-point improvement in DLQI given that:

	○ EASI75 formed the primary or key secondary endpoint of all trials in the BREEZE-AD programme as it is considered of significant clinical importance by dermatologists. For this reason, use of EASI75 as the response definition was presented as a key scenario analysis in the CS. ○ This is supported by clinical expert opinion as sought by the ERG and the NICE technical engagement team as part of this appraisal and as obtained by Eli Lilly from a current consultant dermatologist in UK clinical practice, who confirmed that departments are moving more towards EASI75 as an assessment criterion for response to treatment. Additionally, EASI75 was considered to be a clinically significant treatment response by the British Association of Dermatologists as noted on Pages 17 and 79 of the ERG report. • Therefore, the company provide an economic model in which the revised company base case considers EASI75 as the definition of response. The additional input data for this response outcome and a narrative description of the changes made in the economic model are presented in the Company Technical Engagement Appendix. The analyses where response is based on achievement of EASI75 demonstrate that baricitinib is a cost-effective use of NHS resources.
Issue 5: Time to assessment of response	
Do you anticipate that response to baricitinib would be assessed at 12 or 16 weeks in NHS clinical practice?	• A panel of expert dermatologist advisors to Eli Lilly have confirmed that it is expected that the majority of clinical assessments will be carried out at 16 weeks in UK clinical practice. • The company consider Week 16 to be the most suitable timepoint for assessment of response to baricitinib in the economic model because, as stated in the company submission and discussed in response to ERG clarification question B4, assessment of response at Week 16 was selected in alignment with the primary timepoint of response assessment performed in the pivotal BREEZE-AD trials which provide the clinical evidence base for baricitinib.[1, 4-7] • Furthermore, the dupilumab trials used a 16-week timepoint for response assessment, and thus the ITC performed using data from the Week 16 timepoint provides the most robust comparative efficacy estimates for baricitinib versus dupilumab for use in the model. • The wording in the recently published SmPC for baricitinib has been updated since the draft version shared at the time of submission. The published SmPC for baricitinib now contains the following wording: “that consideration should be given to discontinuing treatment in patients who show no evidence of therapeutic benefit after 8 weeks

	of treatment."[11] Importantly, the SmPC does not suggest that a *response assessment* should be carried out at this earlier timepoint, only that patients should be considered for discontinuation if they have shown *no evidence* of therapeutic benefit, i.e. a *futility assessment*. Patients who show some improvement on baricitinib but have not yet achieved EASI75 by Week 8 would therefore not be considered for discontinuation, and thus is would be inappropriate to use this timepoint for assessment of response in the model. As a result of the changes since the draft SmPC provided at submission, consideration of response at Week 12 is no longer mentioned in the latest SmPC.
Would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks?	• As noted above, the recently published SmPC for baricitinib now contains the following wording: "that consideration should be given to discontinuing treatment in patients who show no evidence of therapeutic benefit after 8 weeks of treatment", but does not suggest that response assessment should be carried out at this timepoint.[11] • Feedback from an expert panel of dermatologists is that this timepoint is likely to be too early to enable accurate assessment of efficacy in UK clinical practice. An unwillingness to discontinue patients this early was highlighted due to the possibility that they may later respond, particularly given that in some AD patients, some areas of skin inflammation such as the face may take a longer time than 8 weeks to improve. A practical limitation of patients returning so quickly to the clinic was also raised, as this would put additional pressure on already overburdened dermatology departments.
Issue 6: Treatment sequencing	
Assuming both baricitinib and dupilumab were available, how do you anticipate that they would likely be used in NHS practice? Please provide a rationale for your response. • Baricitinib followed by dupilumab • Dupilumab followed by baricitinib	• As stated in the company submission, baricitinib is positioned in UK clinical practice for adult patients with moderate-to-severe AD who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control. This is also aligned with the eligibility criteria for the main source of clinical effectiveness evidence, BREEZE-AD4 (JAIN).[1] This population is narrower than the full marketing authorisation for baricitinib, which is for the treatment of moderate-to-severe atopic dermatitis (AD) in adult patients who are candidates for systemic therapy, and positions baricitinib for use in a population where dupilumab and BSC are the only remaining treatment options.[3] The positioning of baricitinib is further discussed in the response to Issue 1.

• No treatment sequencing (i.e. baricitinib or dupilumab followed by best supportive care)	• Treatment sequencing with dupilumab is not the intended positioning of baricitinib in UK clinical practice. The company consider that baricitinib should be recommended for use as a treatment option alongside dupilumab, positioned identically to dupilumab in the treatment pathway, as is further discussed in the response to Issue 1. • If baricitinib were to be considered in a sequence the target population would remain the same as the current positioning: patients for whom the only remaining treatment options are currently dupilumab and BSC. • No data are available to describe the efficacy of dupilumab following baricitinib treatment and vice versa, so the impact of prior use of one treatment on the expected efficacy of the other in the target population is unknown. Modelling baricitinib and dupilumab in sequences requires the simplifying assumption that efficacy is unchanged, regardless of positioning. The results of any analyses using this assumption are therefore subject to uncertainty and must be interpreted with caution. • Despite these considerable limitations, the company have explored scenarios in which baricitinib and dupilumab are considered in sequence for completeness. • Fully incremental analysis showed that the treatment sequences presented are not cost-effective. However, it should be noted that these analyses include dupilumab at list price and therefore these ICERs will not be considered for decision-making. Given the dupilumab PAS is confidential, the company cannot produce or comment on decision-making ICERs. The full results and additional input data of these scenarios are presented in the Company Technical Engagement Appendix.
Do you anticipate that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment? If so, why?	• The company are not aware of any data that could inform the efficacy of dupilumab or baricitinib in the target population following treatment with the other, so are unable to provide an answer to this question. This conclusion of uncertainty due to a current lack of evidence is supported by expert opinion sought by Lilly from a clinician practicing in the UK, who agreed that further comment is not possible until more data become available, including real-world evidence from registries such as the A-STAR and other international organisations.
Issue 7: Modelling of best supportive care	
Which of the approaches (or combination of approaches) do you	• The company acknowledge the view of the ERG and the NICE Technical Team that the original model did not take into account the waxing and waning nature of AD. As such, no discontinuation for BSC has been modelled in the revised company base case, to reflect that an approximately constant proportion of patients on BSC may be responding at any given time. However, given the efficacy-effectiveness gap expected for patients receiving

consider most appropriate for modelling best supportive care?	BSC and in alignment with TA534 and the preferences of the NICE Technical Team, the company only considers this approach to be reasonable when treatment waning is also applied to the BSC maintenance and non-response states (in line with TA534). Treatment waning has therefore been applied in the revised company base case. • **Removing treatment discontinuation for BSC may be appropriate:** The company acknowledge the concerns regarding how BSC was modelled given the waxing and waning nature of AD. As stated in the company submission and discussed further in the response to Issue 9, discontinuation from BSC from Year 2 onwards was modelled in the original company base case to reflect that patients who have the observed effect at the end of induction period discontinue BSC over time due to loss of efficacy. In the original base case, the baseline utility value was applied for patients who discontinue from BSC. The company acknowledge the view of the ERG and the NICE Technical Team that this does not take into account the waxing and waning nature of AD. The company therefore understand the ERG's preference to remove discontinuation from BSC from Week 16–Week 52 in Year 1 and in all subsequent years, to reflect that an approximately constant proportion of patients on BSC may be responding at any given time (based on evidence from the CHRONOS trial). Accordingly, no discontinuation for BSC has been modelled in the revised company base case. • **There is no consensus on modelling BSC in dermatology:** The definition of BSC is uncertain, and there is a lack of relevant effectiveness and cost estimates. As a result, BSC has sometimes been proxied by the placebo arms of clinical trials. In atopic dermatitis, this may be problematic given the non-trivial response to placebo. However, from a clinical standpoint, patients have previously failed the constituents of BSC and are therefore expected to have only limited improvement in symptoms in a real-world setting. • **An efficacy-effectiveness gap is expected for patients receiving BSC:** The assumption of sustained effectiveness beyond the trial period is highly implausible and represents a substantial departure from the TA534 model. In essence, it assumes that there is no efficacy-effectiveness gap for topical treatment. There is a wealth of evidence to the contrary. Collectively these findings, even though some of them are from dermatology more broadly, make it very unlikely that this assumption is correct: o In dermatological clinical practice, up to a third of patients with eczema do not collect the first prescription of a newly prescribed topical medicine.[12] For those who do collect the first prescription, secondary persistence is generally low, with most psoriasis patients redeeming only one or two prescriptions before discontinuing the treatment.[13] Within a clinical trial setting of atopic dermatitis, evidence from electronic monitoring suggests that adherence was driven by study visits, peaking and waning according to the study visit schedule.[14] Outside

	of this setting, there is evidence to suggest that patients significantly overstate their adherence to treatment, and 95% of patients with dermatologic conditions in one study underdosed on topical corticosteroids, using only 35% of the expected median dose.[15, 16] Finally, the application of topical medication in AD patients is associated with substantial disutility, with disutility increasing with potency.[17] o Taking these results together, the company consider the assumption that [redacted]% and [redacted]% of patients on BSC would maintain an EASI75 and EASI50 response (note that EASI50 is used to derive the non-response utility value, as per the response to Issue 10), based on Week 16 placebo data from the JAIN trial, to be a significant overestimation of the proportion of patients who would be responding to BSC in usual clinical practice. Given the efficacy-effectiveness gap expected for patients receiving BSC, an assumption of sustained effectiveness based on data derived from a clinical trial setting lacks face validity. Therefore, the company only considers this approach to be reasonable when treatment waning is also applied to the BSC maintenance and non-response states, in line with the committee-preferred approach in the dupilumab appraisal (TA534) and the views of the NICE Technical Team. Scenarios excluding a treatment waning assumption have limited applicability for decision making. • Therefore, the company presents a revised base case in which: o No discontinuation is modelled for BSC. o The utility values have been updated to align with the EASI75 response definition (see response to Issue 4), and a weighted utility value based on data for EASI75 non-responders is applied in the non-response state (see response to Issue 10). o To reflect the efficacy-effectiveness gap expected for BSC in clinical practice (e.g. due to reduced adherence outside the clinical trial setting), a utility waning assumption has also been applied to the BSC maintenance and non-response sates, in alignment with the sensitivity analyses presented in the dupilumab submission (TA534) and as suggested by the NICE Technical Team. The NICE Committee in TA534 considered two utility waning scenarios presented to be clinically plausible, and thus the company present both for consideration.[2] • These results demonstrate baricitinib to be cost-effective versus BSC and dupilumab, with ICERs of £27,037–£28,396/QALY gained and £113,459–£114,262/QALY foregone (South-West quadrant), respectively. • The additional input data and a narrative description of the changes made in the economic model are presented in the Company Technical Engagement Appendix.

Issue 8: Long-term discontinuation rates for baricitinib	
Is the company's assumption of equivalence between baricitinib and dupilumab for long-term discontinuation rates reasonable, in the absence of longer-term data for baricitinib?	• The company acknowledge the uncertainty surrounding the assumption of long-term discontinuation from baricitinib used in the original company base case and have provide additional data from a later data cut (Week 52) of the JAIN trial as part of this response to address this. • As discussed in the company submission, an assumption of equivalence of baricitinib to dupilumab was made since a reliable and valid estimate for discontinuation rates from baricitinib used in combination with TCS could not be generated based on the sustained effectiveness of baricitinib data available from the JAIN trial at the time of submission. This extrapolation would have been dependent on the single time-point of 16–24 weeks. Since submission, 52-week data from the JAIN trial have become available and these longer-term data have been used to derive discontinuation rates from baricitinib and BSC at Week 52 conditional on Week 16 response. • Therefore, the company have provided a revised company base case of the updated economic model to consider these new data, as described in the Company Technical Engagement Appendix, removing the need for an assumption of equivalence with dupilumab.
The ERG used data from JAHN to derive alternative discontinuation rates. Should the discontinuation rate from week 16-52 be based on the probability of response at week 52 conditional on response at week 16, or on all-cause discontinuation rates?	• The company consider conditional probability of response to be a more appropriate proxy of "sustained response" than discontinuation rates, since, unlike the trial setting, patients in clinical practice are unlikely to continue treatments that are not effective. This assumption is also in line with that taken in TA534, and thus the same approach was taken for baricitinib to ensure consistency with the discontinuation rate used for dupilumab. • Discontinuation rates between Week 16 and Week 52 were not shown to be influential variables in the DSA.
Issue 9: Loss of utility benefit (consistency with TA534)	
Should an assumption be incorporated into the model that a proportion of patients discontinue baricitinib each year due to a loss of	• In the original model, the all-cause discontinuation rate applied from Year 2 onwards was assumed to be equivalent to data presented in the dupilumab CHRONOS trial because no long-term data for baricitinib discontinuation were available at the time of submission.[18] As discussed in response to Issue 8, the revised

utility benefit, in addition to the all-cause discontinuation rates already applied? If so, is it reasonable to assume equivalence with dupilumab?	company base case considers the long-term data on discontinuation from baricitinib now available from the JAIN trial, removing the need to assume equivalence to dupilumab. • However, based on the feedback from the NICE Technical Team and given a utility waning assumption is applied to the BSC maintenance and non-response states in the revised company base case (see the response to Issue 7), utility waning assumptions have also been applied to the baricitinib and dupilumab treatment arms to fully align with the committee-preferred approach in TA534.[2]
Issue 10: Utility values	
Do you anticipate that a patient achieving a response as defined by EASI50 plus a ≥4 point improvement in DLQI would experience an improvement in health-related quality of life?	• The company acknowledge the concerns of the ERG and NICE Technical Team surrounding the utility values associated with the EASI50 plus a ≥4-point improvement in DLQI endpoint. Utility values based on the EASI response categories have been applied in the revised company base case. • As stated in the company submission and in response to Issue 4, the company initially selected EASI50 plus a ≥4-point improvement in DLQI as the response definition to align with the definition accepted by the Appraisal Committee as clinically relevant during the dupilumab NICE appraisal (TA534). The company acknowledge the ERG's concerns regarding the limitations of the utility values employed in the original base case which indicate that response is associated with little to no improvement in HRQoL in alignment with the pattern observed in the dupilumab appraisal. • As discussed in response to Issue 4 above, the company has reconsidered the clinical relevance of the composite outcome following feedback from the ERG and the NICE Technical Team and the clinical expert opinion received as part of this appraisal. Therefore, the revised company base case considers EASI75 as the response definition. • Accordingly, the revised company base case incorporates utility inputs based on EASI response categories, as presented in the Company Technical Engagement Appendix. In contrast to the original utility values (based on EASI50 with ΔDLQI ≥4 response categories), the revised utility inputs did not lack face validity, and as such, non-responders are assigned a weighted average of the EASI<50 and EASI50 to EASI<75 utility values.
Do you consider the utility values preferred by the company or ERG	• Following the update to the economic model to utilise the EASI75 response rate, the newly presented utility values based on the EASI75 response definition should be considered alongside those preferred by the ERG.

(derived from TA534) to be more appropriate?	
Issue 11: Differences in clinical outcomes based on region and skin type	
Do you consider the data from the Japanese patients recruited in the baricitinib trials to be generalisable to clinical practice in the NHS?	• The company consider the patients recruited to the BREEZE-AD trials to be representative of patients in UK clinical practice. This is supported by a consultant dermatologist practicing in the UK consulted by Eli Lilly who highlighted that the BREEZE-AD trial programme would likely be reviewed as a whole by the clinical community, and that overall global recruitment was considered to be well-balanced across all trials. It was further acknowledged that, in the absence of UK-specific clinical trials, it is not possible to collect data that wholly reflect the patient population in UK clinical practice. • The company acknowledge the concerns of the ERG and NICE Technical team regarding the generalisability of data from Japanese patients in the BREEZE-AD trials, but not that the BREEZE-AD trial programme was not designed to investigate baricitinib efficacy in Japanese patients compared with other patient populations. Therefore, response variation in Japanese and non-Japanese patients should be interpreted with caution given the limited the available data and potential differences in clinical practice.[1, 4-7] • As discussed in response to ERG clarification question A6, exposure-response analysis (conducted with data from Studies JAHL and JAHM, for which PK samples were collected) up to 16 weeks of treatment did identify the Japanese patient population as a significant covariate related to drug effect, but little separation between the [redacted]% prediction intervals in the response-time plots for Japanese versus non-Japanese patients was observed for the clinically relevant endpoints of EASI75, EASI50, EASI90 and 4-point improvement in Itch NRS.[4, 5] Interpretation of the covariate effect between Japanese and non-Japanese patients should be carried out with caution given that the Japanese subpopulation was relatively small ([redacted] patients in total, or [redacted]% of the total patients included in the pharmacodynamic dataset) and differences for several baseline disease characteristics were noted between Japanese and non-Japanese patients. These differences included EASI and BSA baseline characteristics (more severe AD in Japanese patients) and the use of TCS ([redacted]% by Week 16 in Japan compared with [redacted]% in the overall population). • As discussed in response to ERG clarification question A5b, rescue rates in Japan ([redacted]%) were higher than those observed in Europe ([redacted]%). It is not possible to determine from the available data whether this reflects differences in the disease itself and it remains likely that it is instead reflective, at least in part, of differences in

	clinical practice and investigator's choice. For example, Japanese clinical practice favours TCS use, including high potency TCS, rather than systemic agents, while European clinical practice broadly limits the use of high potency TCS. Therefore, Japanese patients would be more likely to be rescued with higher potency TCS, leading to non-responder imputation (when the primary censoring rule is applied) indicating lower response rates in these patients.[1, 4-7] • Additionally, the results of regional subgroup analyses for IGA 0 or 1 in Studies JAHL and JAHM and in Study JAIY do not support a lower response for baricitinib 4 mg in East Asian countries not including Japan, suggesting that there is not a specific effect of East Asian ethnicity. • Together, the company conclude that these results do not suggest a specific effect of Japanese ethnicity on the treatment effect of baricitinib. The overall low response rate in both PBO and baricitinib-treated patients in Japan likely reflect differences in baseline characteristics and treatment practices related to TCS rescue specific to Japan, but the company acknowledge that in the absence of robust data investigating the effect of ethnicity on baricitinib efficacy, no definitive conclusion can be offered on this topic.
Would data from European patients only be more clinically relevant than the full intent-to-treat population?	• For completeness, the company have conducted a scenario analysis based on European patients only from JAIN. • Use of the Europe-only population in this scenario analysis had a minimal impact on the cost-effectiveness results as compared with the revised company base case. Baricitinib was found to be cost-effective versus BSC, with an estimated ICER of £27,077/QALY which falls below the NICE willingness-to-pay (WTP) threshold of £30,000. Versus dupilumab, baricitinib was cost-effective in the South-West quadrant, accruing considerably fewer costs and slightly fewer QALYs (ICER: £128,407/QALY foregone). • The full results and additional input data for this scenario are presented in the Company Technical Engagement Appendix.
Is it reasonable to assume comparable efficacy for baricitinib in black patients compared to white patients in the absence of robust data? If not, what adjustments should be made when assessing the	• The company acknowledge the concerns of the ERG and the NICE Technical Team regarding the lack of data for black patients from the BREEZE-AD trials, but note, as discussed in response to ERG clarification question A7, that the trial program was not designed to compare efficacy between different racial or ethnic groupings and as such, the ethnicity distributions of the BREEZE-AD trials are reflective of the participating countries rather than of the occurrence of AD. In particular, the company highlight that patients were not recruited from the US and that if they had been, it would be expected that a higher proportion of black patients would have been recruited.

effectiveness of treatment in black patients?	• The company further acknowledge that the pathology of AD and its underlying cytokine pathways may be distinct in black patients, but note that the Appraisal Committee in the dupilumab appraisal (TA534) concluded there to be insufficient evidence to determine the extent to which differences in cytokine pathways could modify treatment effect.[2]
Issue 12: Indirect treatment comparison heterogeneity	
Do you consider that the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the indirect treatment comparison in favour of either treatment?	• As discussed in the company submission, the company acknowledge that heterogeneity exists in the trials included within the ITC, but welcome the conclusion from the NICE Technical Team that the patient populations are reasonably comparable overall and agree with the conclusion provided by the ERG (ERG report, page 78) and acknowledged by the NICE Technical Team that differences in secondary censoring rules between the dupilumab trials and baricitinib BREEZE-AD trials could bias the relative efficacy results in favour of dupilumab. • The company accept that some trial heterogeneity is an expected limitation of an ITC, as acknowledged by the ERG (ERG report, page 15) but consider that this heterogeneity is reasonable and does not reduce the validity of the data produced.
Issue 13: Impact of baricitinib on flare control	
Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable, in the absence of longer term data for baricitinib?	• The company acknowledge the uncertainty surrounding this assumption and provide updated analyses in which baricitinib flare rate is assumed to be equal to the placebo flare rate. • As outlined in the company submission, the flare annual flare rate of baricitinib was assumed to be equal to the dupilumab flare rates presented in the dupilumab submission (TA534) given the lack of long term data for baricitinib.[2] • As discussed in the company submission and in response to ERG clarification question A3, the company acknowledge that limitations are associated with the use of alternative sources of flare rate data derived from the dupilumab trials in the absence of this long term baricitinib data. The washout period for topical treatments prior to randomisation was longer in the BREEZE-AD trials than in the CHRONOS and SOLO1/2 dupilumab trials and the CAFÉ trial utilised a wash-in period in which patients could use TCS at investigator discretion during the initial 2 weeks of the screening period.[1, 4-7, 18-20] This shorter washout period, or lack thereof, in the dupilumab trials is likely to have resulted in patients being less likely to experience flares and more likely to have a better response

	than patients in the BREEZE-AD baricitinib trials. The company welcome acknowledgement of this by the ERG (ERG report, page 78). • The company acknowledge the preferences of the ERG and NICE technical engagement team for the use of long-term data on the flare rate of baricitinib derived from the JAIN trial. The company have updated the economic model to consider the flare rate of patients treated with baricitinib as equivalent to the flare rate of patients treated with placebo as presented in the dupilumab submission. This conservative assumption has been applied in the revised company base case as presented in the Company Technical Engagement Appendix, in which the additional input data are also presented.
Issue 14: Concomitant treatments	
Does the ERG's revised scenario in which the costs of bathing products are removed from the model best reflect current NHS clinical practice?	• The company acknowledge the preference of the ERG and NICE Technical Team to remove the cost of bathing products from the economic analysis. • The company provide an updated economic model in which the revised company base case models concomitant treatments in line with the preferred assumptions of the ERG. The additional input data associated with this update are presented in the Company Technical Engagement Appendix.
Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?	• The company acknowledge the preference of the ERG and NICE Technical Team to update how concomitant treatment is altered following treatment response in the economic analysis. • The company provide an updated economic model in which the revised company base case models concomitant treatments in line with the preferred assumptions of the ERG. The additional input data associated with this update are presented in the Company Technical Engagement Appendix.

References

1. I4V-MC-JAIN (BREEZE-AD4) Clinical Study Report (13th April 2020).
2. National Institute for Health and Care Excellence (NICE). Dupilumab for treating moderate to severe atopic dermatitis: Technology appraisal guidance [TA534]. Available at: https://www.nice.org.uk/guidance/ta534 [Last accessed: 28th January 2020].
3. EMC. Dupixent 300 mg Solution for Injection in Pre-filled Syringe: SmPC. Available at: https://www.medicines.org.uk/emc/product/8553/smpc [Last accessed: 28th January 2020].
4. I4V-MC-JAHL (BREEZE-AD1) Clinical Study Report (24th September 2019).
5. I4V-MC-JAHM (BREEZE-AD2) Clinical Study Report (8th October 2019).
6. I4V-MC-JAIY (BREEZE-AD7) Clinical Study Report (25th October 2019).
7. I4V-MC-JAHN (BREEZE-AD3) Clinical Study Report (7th November 2019).
8. Akinlade B, Guttman-Yassky E, de Bruin-Weller M, et al. Conjunctivitis in dupilumab clinical trials. British Journal of Dermatology 2019;181:459-473.
9. Taylor K, Swan DJ, Affleck A, et al. Treatment of moderate-to-severe atopic eczema in adults within the U.K.: results of a national survey of dermatologists. British Journal of Dermatology 2017;176:1617-1623.
10. Boguniewicz M, Alexis AF, Beck LA, et al. Expert Perspectives on Management of Moderate-to-Severe Atopic Dermatitis: A Multidisciplinary Consensus Addressing Current and Emerging Therapies. J Allergy Clin Immunol Pract 2017;5:1519-1531.
11. EMC. Olumiant 4 mg Film-Coated Tablets: SmPC. Available at: https://www.medicines.org.uk/emc/product/7486/smpc [Date accessed: 3rd November 2020].
12. Storm A, Andersen SE, Benfeldt E, et al. One in 3 prescriptions are never redeemed: primary nonadherence in an outpatient clinic. J Am Acad Dermatol 2008;59:27-33.
13. Svedbom A, Dalén J, Mamolo C, et al. Treatment patterns with topicals, traditional systemics and biologics in psoriasis - a Swedish database analysis. J Eur Acad Dermatol Venereol 2015;29:215-23.
14. Krejci-Manwaring J, Tusa MG, Carroll C, et al. Stealth monitoring of adherence to topical medication: adherence is very poor in children with atopic dermatitis. J Am Acad Dermatol 2007;56:211-6.
15. Hix E, Gustafson CJ, O'Neill JL, et al. Adherence to a five day treatment course of topical fluocinonide 0.1% cream in atopic dermatitis. Dermatol Online J 2013;19:20029.
16. Storm A, Benfeldt E, Andersen SE, et al. A prospective study of patient adherence to topical treatments: 95% of patients underdose. J Am Acad Dermatol 2008;59:975-80.
17. Retzler J, Smith A, Reaney M, et al. Process utilities for topical treatment in atopic dermatitis. Qual Life Res 2019;28:2373-2381.
18. Blauvelt A, de Bruin-Weller M, Gooderham M, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corticosteroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. The Lancet 2017;389:2287-2303.
19. de Bruin-Weller M, Thaci D, Smith CH, et al. Dupilumab with concomitant topical corticosteroid treatment in adults with atopic dermatitis with an inadequate response or intolerance to ciclosporin A or when this treatment is medically inadvisable: a placebo-controlled, randomized phase III clinical trial (LIBERTY AD CAFE). Br J Dermatol 2018;178:1083-1101.

20. Simpson EL, Bieber T, Guttman-Yassky E, et al. Two Phase 3 Trials of Dupilumab versus Placebo in Atopic Dermatitis. New England Journal of Medicine 2016;375:2335-2348.

Company Technical Engagement Appendix

1. Revised company base case

Following feedback from the ERG and the NICE Technical Team, the company have updated the economic model to produce a revised base case. This updated model is provided alongside this document and the Technical Engagement Response Form. A summary of the updates made to form the revised base case of the model, and which are therefore applied to all analyses, is presented in Table 1 before a full description of each is provided below.

Table 1: Summary of changes in the revised base case

Model input	Original company base case	Revised company base case	Section	New input data	TE Issue
Treatment response definition	EASI50 + ≥4-point improvement in DLQI	EASI75	1.1	Table 3	4
Maintenance of response to Week 52	Dupilumab and BSC rates sourced from CHRONOS; baricitinib assumed to be equivalent to dupilumab	Baricitinib rates derived from the JAIN trial; dupilumab rates sourced from TA534;[1] no discontinuation modelled for BSC	1.2	Table 4	8
Long-term discontinuation after Week 52			1.3	Table 5	8
Health-related quality of life data and utility values	Defined based on achievement of EASI50 + ≥4-point improvement in DLQI, non-responders assigned baseline utility	Defined based on achievement of EASI75; non-responders assigned a weighted average of the EASI<50 and EASI50 to EASI<75 utility values	1.4	Table 8	10
Utility waning	No utility waning	Utility waning applied across all treatment arms as per assumptions in TA534	1.4	Table 9	7 and 9
Flare rate and treatment	Dupilumab and BSC rates and treatment distribution sourced from TA534; baricitinib assumed to be equivalent to dupilumab	Dupilumab and BSC rates and treatment distribution sourced from TA534; baricitinib assumed to be equivalent to BSC	1.5	Table 10	13
BSC and concomitant treatment costs	Costs of BSC treatment and concomitant bathing products included	Costs of BSC treatment and concomitant bathing products excluded	1.5	Table 11	14
Dupilumab dosing	10 doses in induction period	9 doses in induction period	1.5	N/A	N/A
Monitoring costs	No full blood count tests for patients on baricitinib	4 full blood count tests per annum for patients on baricitinib	1.5	Table 12	N/A

Abbreviations: BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; N/A: not applicable.

1.1. Treatment response definition

Following feedback from the ERG and the NICE Technical Team, the company acknowledge the relevance of EASI75 as a determinant of treatment response to baricitinib. Therefore, in the revised company base case, the efficacy response criterion is defined as achievement of EASI75 which reflects an effective proxy for capturing sufficient clinical benefit to justify continuing treatment after the trial period. This is supported by the clinical expert opinion sought by the ERG and the NICE Technical Engagement Team as part of this appraisal and by the British Association of Dermatologists as noted on Pages 17 and 79 of the ERG report.

Treatment response for baricitinib and dupilumab at the completion of the 16 week induction period, which was considered to be the most appropriate time point as per the response to Issue 5, are taken from an indirect treatment comparison (ITC) informed by baricitinib treatment response data from the JAIN + JAIY JAIN-like population and dupilumab treatment response data available in the public domain. Data for baricitinib were derived from the analysis where the secondary censoring rule was applied, which may be considered a better reflection of clinical practice and be more in alignment with TA534, as noted on Page 110 of the ERG report. Treatment response for BSC is represented by the placebo response rate in the corresponding trial. All patients included in the ITC had moderate-to-severe AD, received treatment in combination with TCS and a history of prior ciclosporin treatment failure or contraindication. The methodology of the ITC was in line with that described in Section B.2.9.3 of the original Company Submission (CS). As per the original submission, a fixed-effect model was chosen as there were too few studies included to produce reliable between-study variations for random-effects models. The results of the ITC are presented in Table 2.

Table 2: Relative treatment effect of pairwise comparisons expressed as OR and RR (with 95% CI) in EASI75 at Week 16 in the JAIN-like JAIY and JAIN patients and CAFÉ-like CHRONOS and CAFÉ patients, fixed-effects model

Population	Comparison	OR (95% CI) [p value]	RR (95% CI) [p value]	Risk difference (95% CI) [p value]
JAIN-like JAIY and JAIN	4 mg BARI qd + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
CAFÉ-like CHRONOS and CAFÉ	300 mg Dupi q2w + TCS vs PBO + TCS	[redacted]	[redacted]	[redacted]
Indirect comparison	4 mg BARI + TCS vs 300 mg Dupi + TCS	[redacted]	[redacted]	[redacted]

All p-values were derived from fixed-effects model.
Abbreviations: BARI: baricitinib; CI: confidence interval; Dupi: dupilumab; EASI: Eczema Area and Severity Index; OR: odds ratio: PBO: placebo; q2w: once every two weeks; qd: once daily; RR: risk ratio; TCS: topical corticosteroids.

In the original submission, an additive method based on the risk difference observed in the ITC was chosen to derive the dupilumab response rates. However, based on feedback from the ERG and in line with NICE Decision Support Unit (DSU) Technical Support Document (TSD) 5,[2] the

response rate for dupilumab in the revised company base case was derived by taking the odds of the response rate for placebo from the baricitinib trials and the relative effects (odds ratio [OR]) for dupilumab versus placebo from the dupilumab trials, adding both on the log scale and finally transforming it back into proportions. A standard error for the dupilumab response rate was derived using 10,000 Monte-Carlo simulations.

The response rates used in the revised economic model are presented in Table 3. Within the base case analysis, the EASI75 response rate is used to define entry into the Maintenance health state. The EASI50 response rate is used in the derivation of the utility value associated with the Non-Response state, as discussed further in Section 1.4.

Table 3: Response for EASI75 and EASI50 at Week 16 in the revised economic model

	Probability of response at Week 16, % (SE%)	
	EASI50	EASI75
Baricitinib	66.44 (3.87)	42.28 (4.05)
Dupilumab	85.72 (4.55)	57.16 (7.31)
BSC	42.36 (4.12)	22.22 (3.46)

Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; SE: standard error.

1.2. Maintenance of response up to Week 52

After the end of the 16-week induction period, responders enter the Maintenance treatment phase and receive continuous treatment. Feedback from the ERG suggested that it would be most appropriate to use all-cause discontinuation from the trials to inform discontinuation rates between Week 16 and 52. In the original company base case, conditional probability of response at Week 52 based on response at Week 16 was used to derive the discontinuation rates between Week 16 and 52. The company consider conditional probability of response to be a more appropriate proxy of "sustained response" than discontinuation rates, since, unlike the trial setting, patients in clinical practice are unlikely to continue treatments that are not effective. This assumption is also in line with that taken in TA534, and thus the same approach was taken for baricitinib to ensure consistency with the discontinuation rate used for dupilumab.

In the revised company base case, the conditional probability of response to baricitinib at Week 52 based on response at Week 16 has been based on longer term data from the JAIN trial which have become available since submission of the CS. The conditional probability of response to dupilumab at Week 52 was sourced from the CHRONOS study as presented in the dupilumab NICE submission (TA534).[1] If response is not sustained, patients can discontinue treatment and move to a subsequent line of treatment. The conditional probabilities of sustained response at Week 52 are presented in Table 4. As per the eligibility criteria of the JAIN trial, all patients were post-ciclosporin, whereas the CHRONOS trial considered patients with a history of inadequate response to medium-to-high potency TCS with or without topical calcineurin inhibitors, and/or patients with a history of systemic treatment.[3] As a fully post-systemic treatment population, the JAIN trial represents a more severe AD population than the patient population of the CHRONOS trial. Therefore, the use of probabilities of response in the updated base case that are derived from these patient populations is likely to bias against baricitinib and can be considered a conservative approach.

The conditional probabilities of response at Week 52 are used to estimate of the probability of treatment discontinuation between the end of the induction period (Week 16) and Week 52. This

is performed by deducting the conditional probability of response for each treatment at Week 52, as presented in Table 4, from the probability of response at the time of response assessment. Thereafter, the probabilities are converted into a constant instantaneous rate, which is in turn converted to the desired length probability of four weeks. This approach assumes that patients who have lost response between the response assessment and 52 weeks have done so at a continuous and constant rate and that they discontinue treatment once they have lost response. As per the Technical Team and ERG's preferences, no discontinuation has been modelled for BSC.

Table 4: Response probabilities for EASI75 at Week 52 conditional upon response at Week 16 for baricitinib and comparators employed in the revised company base case analysis

	Response probability, % (SE%)
Baricitinib	[redacted]
Dupilumab	82.10 (5.30)

Based on the JAIN trial, the probability of achieving EASI75 at Week 52 conditional upon response at Week 16 for placebo was [redacted]. However, to align with the Technical Team and ERG's preferences, no discontinuation has been modelled for BSC.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; SE: standard error.

1.3. Long-term discontinuation after Week 52

After 52 weeks, the model includes an annual probability of discontinuation that represents the annual rate at which patients discontinue baricitinib or dupilumab each year due to lack of long-term efficacy, adverse event, patient preference, or physician preference. The annual probability of discontinuation is applied to patients in the Maintenance health state starting at the second year of the model. Patients who discontinue dupilumab or baricitinib during this time transition to another line of treatment or BSC.

In the original company model, it was assumed that dupilumab and baricitinib have the same annual probability of treatment continuation, reflecting the withdrawal probabilities observed in the CHRONOS trial: 5.1%.[3] The annual probability of study withdrawal or use of rescue medication was set at 57.0%, as sourced from the BSC arm in the CHRONOS trial. These assumptions were necessary given that reliable and valid estimation for long-term discontinuation rates (beyond 52 weeks) based on JAIN data was not available at the time of submission.

Longer-term (Week 52) data from the JAIN trial have since become available to inform the annual probability of discontinuation from baricitinib. Annual discontinuation rates from dupilumab have been sourced from the dupilumab NICE submission (TA534).[1] The annual probability of discontinuation with EASI75 as the response criterion for the second and subsequent years utilised in the revised company base case are presented in Table 5. As per the Technical Team and ERG's preferences, no discontinuation has been modelled for BSC.

Table 5: Annual probabilities of discontinuation after Week 52 employed in the revised company base case

	Annual probability of discontinuation in Years 2+, %
Baricitinib	[redacted]
Dupilumab	5.10

The annual probability of discontinuation for placebo based on the JAIN trial was [redacted]%. However, to align with the Technical Team and ERG's preferences, no discontinuation has been modelled for BSC.

Abbreviations: BSC: best supportive care; CS: company submission.

1.4. Health-related quality of life data and utility values

In the revised company base case, health states are defined based on achievement of EASI75. Health state utility values were derived using the methodology described in the CS (Document B, Section B.3.4.4), but including EASI response categories based on the secondary censoring rule (EASI <50, EASI50 to <75 and EASI ≥75 [EASI75 to <90 combined with EASI ≥90]) as fixed effects in the mixed model. The combined EASI ≥75 response category was used because EASI90 data were not available from the dupilumab trials.

The resulting EQ-5D-3L utility scores are presented in Table 6, including the number of observations included in the analysis. The change from baseline utility increased with increasing levels of response. This trend is in contrast to the model based on EASI50 with ΔDLQI ≥4 response categories, where the change from baseline utility for non-responders was greater than the change from baseline for responders, lacking face validity. Standard errors (SE) were generated from the output of the regression models and incorporated into the probabilistic sensitivity analysis. For the baseline utility value, the SE was calculated the observed data, based on the sample size and the standard deviation (as reported in Table 6). For the change from baseline in EQ-5D, the SE was estimated based on the 95% confidence intervals (CI) generated in the output of the MMRM analysis (i.e. by dividing the CI reported in Table 6 by 1.96, assuming a normal distribution).

Table 6: EQ-5D-3L utility score at baseline and Week 16 by EASI response category at Week 16 (secondary censoring)

EASI score	Baseline EQ-5D-3L[a]			Change in EQ-5D-3L (baseline to Week 16)[b]	
	Number of patients	Mean	Standard deviation	LS Mean	95% CIs
Overall	[redacted]	0.6182	0.2786	-	-
EASI <50	[redacted]	[redacted]	[redacted]	0.1312	0.1021, 0.1603
EASI 50 to <75	[redacted]	[redacted]	[redacted]	0.1827	0.1541, 0.2112
EASI ≥75[c]	[redacted]	[redacted]	[redacted]	0.2310	0.2052, 0.2567

[a] Observed values. [b] From mixed model. Number of observations used = [redacted]. [c] EASI 75 to <90 combined with EASI≥90
Abbreviations: CI: confidence interval; EASI: Eczema Area and Severity Index; LS: least squares; SD: standard deviation.

The health state utility value for the maintenance state was derived by applying the change from baseline utility increment for the EASI ≥75 response category to the baseline EQ-5D value. The health state utility value for non-response state was derived by applying a weighted average of the EASI<50 and EASI50 to EASI<75 change from baseline utility increments to the baseline utility value, according to the relative proportions of patients on BSC in each response category at Week 16, as presented in Table 7. As such, in contrast to the original company base case, the utility values applied are no longer based on a within-group analysis. This approach is consistent with recent appraisals in psoriasis (e.g. TA511).[4] The health state utility values implemented in the revised company base case are presented in Table 8.

Table 7: Derivation of the weighted utility value implemented in the revised company base case for non-responders on BSC

EASI score	Proportion of BSC non-responders at Week 16, %	Change in EQ-5D-3L (baseline to Week 16), LS mean	Weighted marginal utility value
EASI <50	74.11	0.1312	0.1445
EASI 50 to <75	25.89	0.1827	

Abbreviations: CI: confidence interval; EASI: Eczema Area and Severity Index; LS: least squares.

Table 8: Utility values implemented in the revised company base case by EASI response category

EASI level	Utility values (JAIN + JAIY JAIN-like)
Baseline values, mean (SD)	0.6182 (0.2786)
Change from baseline at Week 16: mean LS	
Non-response (EASI <75)	0.1445
Maintenance (EASI ≥75)	0.2310

Abbreviations: EASI: Eczema Area and Severity Index; LS: least squares; SD: standard deviation.

Utility waning

In the original submitted model, a discontinuation rate was applied to the BSC maintenance state from Year 2 onwards to reflect that patients who had initially achieved a response on BSC would experience a loss of efficacy in the long term. These patients discontinued to the non-response state where they were assigned the baseline utility value. Given the limited data on the loss of response to BSC outside the trial setting in the relevant patient population, the probability of study withdrawal or use of rescue medication from the BSC arm in the CHRONOS trial was applied (57.0%). The company acknowledge the view of the ERG and the NICE Technical Team that this does not accurately take into account the waxing and waning nature of AD, given the rapid rate of discontinuation modelled beyond Week 52 for BSC and the permanent loss of disease control. The company therefore understand the ERG's preference to remove discontinuation from BSC from Week 16–52 in Year 1 and in all subsequent years to reflect an approximately constant proportion of patients on BSC who may be responding at any given time. However, as outlined in the response to Issue 7, this assumption of sustained effectiveness beyond the trial period is highly implausible and represents a substantial departure from both the ICER and TA534 models. In essence, it assumes that there is no efficacy-effectiveness gap for topical treatment, when from a clinical standpoint, patients have previously failed the constituents of BSC and would therefore be expected to have only limited improvement in symptoms in a real-world setting. The Company therefore welcome the technical team's acknowledgement that the ERG scenario does not take into account a reduction in BSC efficacy outside of the clinical trial setting, where treatment adherence is likely to be reduced.

To address the ERG and Technical Team's concerns described above, no discontinuation has been modelled for BSC, and a weighted utility value based on data for EASI75 non-responders is applied in the non-response state. However, to reflect the efficacy-effectiveness gap expected for BSC in clinical practice (e.g. due to reduced adherence outside the clinical trial setting), a treatment waning assumption has also been applied to the BSC maintenance and non-response sates, in alignment with the sensitivity analyses presented in the dupilumab submission (TA534) and as suggested by the NICE Technical Team. For consistency, treatment waning assumptions

have also been applied to the baricitinib and dupilumab treatment arms in line with the TA534 sensitivity analyses, as per the response to Issue 9.

Across all health status, the utility value applied is adjusted down over time by assuming a proportion of patients maintain the utility benefit associated with the maintenance/non-response health state (as shown in Table 8), with the remaining proportion of patients assigned baseline utility. The proportion of the utility benefit maintained for each treatment in these scenarios is presented in Table 9.

Table 9: Proportion of utility benefit maintained (waning of utility over time)

Year of Treatment	Baricitinib	Dupilumab	BSC	
			NICE TA534 Sensitivity Analysis 1	NICE TA534 Sensitivity Analysis 2
Year 2	0.98	0.98	0.18	0.43
Year 3	0.95	0.95	0.10	0.18
Year 4	0.93	0.93	0.06	0.08
Year 5+	0.92	0.92	0.04	0.03

Abbreviations: BSC: best supportive care; TA: Technology Appraisal.

Waning is applied to the year of treatment while in the maintenance health state, for example the second year on baricitinib or second year on BSC. To implement waning within the cohort perspective, the model applied an assumption that the cohort of patients discontinuing from first line treatment will receive the same "waned" utility weight upon arrival in second line maintenance as the cohort that has been in that state since the 2nd line induction. Improving the accuracy of this assignment (and avoiding this simplifying assumption) would require an individual level simulation.

1.5. Cost and healthcare resource use

Flare rate and treatment

In the original model, the proportions of treatments of flares and annual flare rates for dupilumab and BSC were assumed to be the same as those presented in the dupilumab NICE submission, with baricitinib rates further assumed to be equal to those of dupilumab.[1]

The company acknowledge the uncertainty surrounding this assumption. In the revised company base case, it is assumed that the annual flare rate for baricitinib and the distribution of treatment per flare are equal to those of patients treated with placebo as presented in the dupilumab submission (TA534). The flare management inputs utilised in the revised company base case are presented in Table 10, which is an update of Table 97 in the CS.

Table 10: Proportion of treatment of flare and annual flare rate in the revised company base case

	Baricitinib	Dupilumab	BSC
Annual flare rate	0.78	0.18	0.78
Distribution of treatment per flare			
Proportion TCS (potent) at 52 weeks	0.54	0.42	0.54
Proportion TCS (very potent) at 52 weeks	0.27	0.23	0.27

Proportion systemic steroids at 52 weeks	0.13	0.29	0.13
Proportion TCI at 52 weeks	0.06	0.00	0.06

Abbreviations: BSC: best supportive care; TCI: topical calcineurin inhibitor; TCS: topical corticosteroids.
Source: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534).[1]

BSC and concomitant treatment costs

Following feedback from the ERG and NICE Technical Team, the company acknowledge the preference of the ERG and NICE Technical Team to remove the BSC treatment costs from the economic analysis in order to avoid double counting these costs. In the revised company base case, BSC costs have been removed in line with the preferred assumption of the ERG.

The company further acknowledge the preference of the ERG and NICE Technical Team to remove bathing products from consideration in the economic model. This update has also been made in the revised company base case in line with the preferred assumption of the ERG.

The costs of concomitant medication, consisting of emollient products, mid-potency background TCS (mometasone 0.1% ointment) and TCI (Protopic 0.1% ointment, tacrolimus), utilised in the revised company base case are presented in Table 11, which is an update of Table 96 in the CS.

Dupilumab dosing

The company acknowledge the incorrect dosing of dupilumab in the induction period of the original model. In the revised company base case, the number of dupilumab doses administered in the induction period has been corrected from 10 to 9.

Monitoring costs

The healthcare resource use and monitoring costs in the revised company base case have been changed in alignment with the preference of the ERG. Four full blood count tests per annum have been included for baricitinib additional monitoring as presented in Table 12, which is an update and combination of Tables 99 and 100 in the CS.

Table 11: Costs of bathing products and emollients used in the revised company base case

Medication	Pack costs	Pack size	Proportion of product prescribed	Amount per week (non-responders)	Weekly costs (non-responders)	Weekly costs (responders)[a]	Resource use (induction, 16 weeks)	Resource use (maintenance, annual)
Emollients					**£5.24**	**£2.22**	16.0	52.0
Aveeno cream	£6.47	500ml	-	1	£6.47	£3.24		
Cetraben ointment	£5.39	450g	-	1	£5.39	£2.70		
Dermol cream	£6.63	500g	-	1	£6.63	£3.32		
Diprobase ointment	£5.99	500g	-	1	£5.99	£3.00		
Epaderm ointment	£12.25	1000g	-	0.5	£6.13	£1.53		
Hydromol ointment	£8.20	1000g	-	0.5	£4.10	£1.03		
White soft paraffin 50%/ liquid paraffin 50% ointment	£4.57	500g	-	1	£4.57	£2.29		
Oilatum cream	£5.28	500ml	-	0.5	£2.64	£0.66		
TCS								
Mometasone 0.1% ointment	£9.50	100g	-	112.04	£10.64	£5.39[b]	16	52
TCI								
Protopic 0.1% ointment, tacrolimus	£47.28	60g	-	1.75	£1.38	£0.00[c]	16	52

[a] Assuming 50% reduction from non-responder. [b] Assuming usage of 56.70g per week in responders. [c] Assuming no usage of TCIs in responders
Sources: NICE: Dupilumab for treating moderate to severe atopic dermatitis (TA534);[1] MIMS.[5]

Table 12: Administration and monitoring health care resource use for responders used in the revised company base case

Health care resource	Unit cost	Baricitinib		Dupilumab		BSC	
		Induction (16 weeks)	Maintenance (annual)	Induction (16 weeks)	Maintenance (annual)	Induction (16 weeks)	Maintenance (annual)
Dermatologist outpatient consultation (consultant led)	£114.57	2.00	4.30	2.00	4.30	2.00	4.30
Dermatologist nurse visit	£10.50	0.11	0.35	0.11	0.35	0.11	0.35
GP consultation	£39.00	1.91	6.20	1.91	6.20	1.91	6.20
Accident & Emergency visit	£182.58	0.01	0.02	0.01	0.02	0.01	0.02
Hospitalisation	£1,854.72	0.01	0.02	0.01	0.02	0.01	0.02
Day case	£454.67	0.00	0.00	0.00	0.00	0.00	0.00
Full blood count (FBC)	£3.00	1.23	4.00	0.00	0.00	1.23	4.00
Phototherapy	£103.00	0.00	0.00	0.00	0.00	0.02	0.06
Psychological support	£289.46	0	0	0	0	0.02	0.07
Subcutaneous injection	£56.50	0.00	0.00	1.00	0.00	0.00	0.00

Source: PSSRU,[6] NHS Reference Costs,[7] Dupilumab for treating moderate to severe atopic dermatitis (NICE TA534).[1]

2. Scenario analyses

2.1. Summary

Further to the revised company base case, the company present several scenario analyses to explore the impact of alternative model input values, summarised in Table 13. The additional input data for these scenarios are presented below.

Table 13: Summary of scenarios

#	Model input	Revised company base case	Scenario input	Reference to new input data	TE Issue
1	Population	Narrower than full license population: moderate-to-severe AD patients who are candidates for systemic therapy who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control	Comparison versus ciclosporin in the full license population (adult patients who are candidates for systemic therapy)	Table 15	1
2	Treatment sequencing	No treatment sequencing	Baricitinib and dupilumab in sequence and vice versa	N/A	6
3	Population	JAIN + JAIY JAIN-like population	European-only (JAIN) population	Table 16, Table 18	11

Abbreviations: BSC: best supportive care; DLQI: Dermatology Life Quality Index; EASI: Eczema Area and Severity Index; N/A: not applicable.

2.2. Population scenario: Comparison versus ciclosporin in adult patients who are candidates for systemic therapy

As discussed further in response to Issue 1 in the Technical Engagement Response Form, baricitinib is positioned in UK clinical practice for a population for whom dupilumab and BSC are the only remaining treatment options. However, the company acknowledge that there is an unmet need in UK clinical practice for alternatives to current systemic immunosuppressants, which are associated with poor safety profiles, reflected in the feedback from the National Eczema Society which indicated that patients inadequately controlled on topical treatments would benefit from a new alternative treatment option. Therefore, at the request of the ERG and the NICE Technical Team, the company carried out a matching-adjusted indirect comparison (MAIC) for baricitinib versus ciclosporin and utilised methodology in line with the MAIC presented in the dupilumab submission (TA534). It should be acknowledged that the comparison attempted in the TA534 submission was considered by the ERG to not be robust.[1]

Methodology

The two papers on the efficacy of ciclosporin included in TA534 (Haeck *et al*. [2011] and Jin *et al*. [2015]) were considered for inclusion in the MAIC presented here.[8, 9] Of these, only one (Haeck *et al*. [2011]) reports ciclosporin efficacy data at the relevant timepoint of Week 16, so this paper

alone was selected for inclusion in the baricitinib versus ciclosporin MAIC. The eligibility criteria for the BREEZE-AD7 (JAIY) trial are in line with the full license for baricitinib in this indication: adult patients with moderate-to-severe AD who are candidates for systemic therapies. As such, these data are the most relevant to inform a comparison versus ciclosporin in the full population and were included in the MAIC.

The MAIC was conducted using individual patient-level data (ITC) from the JAIY trial and summary evidence from the ciclosporin trial, as presented in Haeck *et al*. (2011). Specifically, EASI75 response in the JAIN trial was predicted using a propensity score weighting approach in an approach in line with the TA534 MAIC which was adapted from the methodology proposed in the NICE Decision Support Unit (DSU) Technical Support Document (TSU) 18.[10]

The MAIC aimed to adjust for baseline characteristics with known or suspected associations with the efficacy outcomes that were reported in both the JAIY trial and Haeck *et al*. (2011). Due to the limited availability of aggregated baseline characteristics reported in Haeck *et al*. (2011), only three characteristics could be matched to the respective data of the JAIY trial: sex, SCORAD score and DLQI score (moderate and low only; high was not matched due to too few cases). Thymus and activation-regulated kinase (TARC) could not be matched as it was not collected in the JAIY study, and data categorisation for immunoglobulin E (IgE) was incompatible between the studies (two categories reported in the JAIY study, mean and SD reported in Haeck *et al*. (2011).

To balance the baseline characteristics between JAIY and EXAM, individual patient data (IPD) from JAIY were assigned weights such that:

- Weighted mean baseline characteristics in JAIY patients matched those reported for patients in Haeck *et al*. (2011)
- The weight for each individual patient was equal to the patient's estimated odds (propensity) of being in JAIY versus Haeck *et al*. (2011)

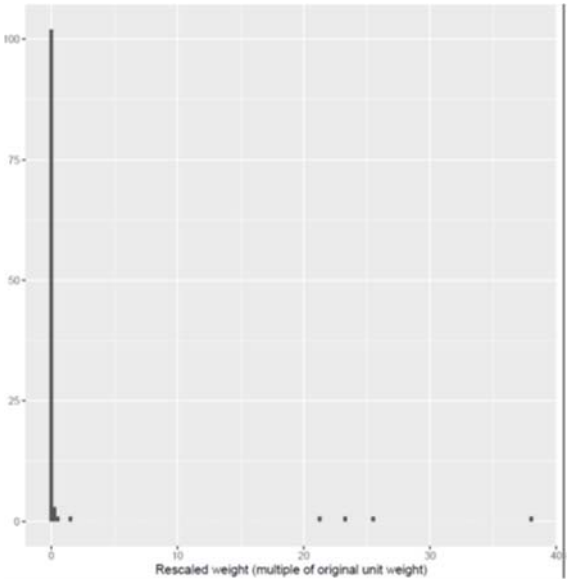
Data on ciclosporin efficacy as defined by achievement of EASI75 are not available from Haeck *et al*. (2011), which reports SCORAD results at Week 16. Therefore, the approach developed in the TA534 submission was followed, in which the Week 16 SCORAD values are mapped to Week 16 EASI75 values by applying a cut-off value of SCORAD 25 (TA534, CS Table 2.41). The same mapping approach was applied to the IPD SCORAD values from the JAIY trial. As this MAIC is exploratory in nature and the analysis is associated with many limitations, statistical testing of the results has not been applied.

Results

Distribution of weights

As shown in Figure 1, the distribution of weights derived from the JAIN trial IPD showed that very few patients could contribute to the calculation of the weighted EASI responses. Accordingly, the effective sample size (following application of weights) reduced from 111 patients to [redacted] patients. This is likely due to the baseline SCORAD in JAIY being substantially different from that in Haeck *et al*. (2011) (means of [redacted] and 42.2, respectively), which results in minimal overlap between the two studies, which prevents meaningful study comparisons.

Figure 1: Distribution of weights



Baseline characteristics matching

Despite this extreme distribution of weights, baseline characteristics matching was successfully performed, with the weighted baseline characteristics for 4 mg baricitinib exactly matched to those reported for ciclosporin, as presented in Table 14.

Table 14: Matching baseline characteristics between JAIY and Haeck *et al*. (2011)

Variable	Statistic	JAIY, before matching	JAIY, after matching	Ciclosporin, aggregated (Haeck *et al*. [2011])
Sample	n	[redacted]	[redacted]	26
Male	%	[redacted]	[redacted]	65.4
Female	%	[redacted]	[redacted]	34.6
SCORAD	Mean	[redacted]	[redacted]	42.2
SCORAD	SD	[redacted]	[redacted]	10.6
DLQI (medium)	%	[redacted]	[redacted]	42.3
DLQI (low)	%	[redacted]	[redacted]	57.7

Abbreviations: DLQI: Dermatology Life Quality Index; SCORAD: SCORing Atopic Dermatitis; SD: standard deviation.

Matched EASI responses

As discussed, valid weighted comparisons of responses between 4 mg baricitinib and ciclosporin could not be derived due to the substantial differences in the baseline characteristics of the JAIY and Haeck *et al*. (2011) studies. Despite this, the calculated Week 16 EASI75 results are presented in Table 15, based on the SCORAD mapping process previously described.

Table 15: Matching EASI75 response rates between JAIY and Haeck *et al*. (2011)

Variable	JAIY, before matching	JAIY, after matching	Ciclosporin, aggregated (Haeck *et al*. [2011])
EASI75 (%)	[redacted]	[redacted]	33

Abbreviations: EASI: Eczema Area and Severity Index.

Conclusion

Given the significant limitations of this analysis, these values are not suitable for clinical interpretation or use in cost-effectiveness analyses or decision-making. Therefore, a robust scenario versus ciclosporin in the full population could not be conducted.

2.3. Treatment sequencing scenarios

Treatment sequencing with dupilumab is not the intended positioning of baricitinib in UK clinical practice. The company consider that baricitinib should be recommended for use as a treatment option alongside dupilumab, positioned identically to dupilumab in the treatment pathway. As discussed further in response to Issue 6 in the Technical Engagement Response Form, the company note that any analyses considering treatment sequencing of baricitinib and dupilumab are subject to considerable uncertainty given the assumption of unchanged efficacy regardless of positioning, for which no supporting data are available. Therefore, the company highlight that these results should be interpreted with caution.

Despite this uncertainty, the company present two scenarios which consider baricitinib followed by dupilumab and dupilumab followed by baricitinib. This required no updates or modification to the model structure, and no new data inputs.

2.4. Population scenario: Europe-only JAIN patients

As discussed in response to NICE Technical Team Issue 11, the company do not conclude that convincing evidence for a specific effect of Japanese ethnicity on the treatment effect of baricitinib exists. However, for completeness, the company present a scenario analysis in which this population is considered. Response rates specific to this population are utilised in this analysis, as presented in Table 16, impacting the derivation of the non-response utility as shown in Table 17. All other inputs and assumptions are in line with the revised company base case described above. These results are less robust than the base case analysis, given the smaller sample sizes informing the inputs.

Table 16: Response rates for EASI75 and EASI50 in the JAIN Europe scenario

	Probability of response at Week 16, % (SE%)	
	EASI50	EASI75
Baricitinib	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]
BSC	[redacted]	[redacted]

Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; SE: standard error.

Table 17: Derivation of the weighted utility value implemented in the JAIN Europe scenario for non-responders on BSC

EASI score	Proportion of BSC non-responders at Week 16, %	Change in EQ-5D-3L (baseline to Week 16), LS mean	Weighted marginal utility value
EASI <50	[redacted]	0.1312	[redacted]
EASI 50 to <75	[redacted]	0.1827	

Abbreviations: CI: confidence interval; EASI: Eczema Area and Severity Index; LS: least squares.

Table 18: EQ-5D-3L utility score at baseline and Week 16 by EASI response category at Week 16 in the JAIN Europe scenario

EASI level	Utility values (JAIN Europe)
Baseline values, mean (SD)	0.6182 (0.2786)
Change from baseline at Week 16: mean LS	
Non-response (EASI <75)	[redacted]
Maintenance (EASI ≥75)	0.2310

Abbreviations: CI: confidence interval; EASI: Eczema Area and Severity Index; LS: least squares; SD: standard deviation.

3. Results

3.1. Revised company base case results

A summary of the results in the revised company base case are presented in Table 19.

When employing utility waning assumptions in line with TA534 Sensitivity Analysis 1, BSC, baricitinib 4 mg and dupilumab Q2W accumulated total costs of £[redacted], £[redacted] and £[redacted], respectively, and accumulated [redacted], [redacted], and [redacted] total QALYs, respectively. At the confidential PAS price, all ICERs in the base case population of patients who have failed at least one current systemic immunosuppressant due to intolerance, contraindication or inadequate disease control, were within the range considered cost-effective. In pairwise comparison of baricitinib versus BSC, the ICER was estimated at £27,037/QALY which falls below the NICE willingness-to-pay (WTP) threshold of £30,000. Versus dupilumab, baricitinib was cost-effective in the South-West quadrant, accruing considerably fewer costs and slightly fewer QALYs (ICER: £113,459/QALY foregone). The probability of cost-effectiveness at WTP thresholds of £20,000 and £30,000 is presented in Table 20 at which baricitinib had a cost-effectiveness probability of [redacted]% and [redacted]%, respectively. Net monetary benefit (NMB) as a function of willingness-to-pay is presented in Figure 2. Similar results were observed employing utility waning assumptions in line with TA534 Sensitivity Analysis 2. These results demonstrate baricitinib to be a cost-effective option for the treatment of moderate-to-severe AD in the target population versus the two comparators relevant to UK clinical practice.

Table 19: Economic results from the revised company base case (JAIN + JAIN-like JAIY)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)
TA534 Sensitivity Analysis 1						
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£27,037	£27,037
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£89,350	£113,459
TA534 Sensitivity Analysis 2						
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£28,396	£28,396
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£91,027	£114,262

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LY: life years; QALYs: quality-adjusted life years.

Table 20: Probability of cost-effectiveness at a WTP threshold of £20,000 and £30,000

	WTP threshold £20,000	WTP threshold £30,000
BSC	[redacted]	[redacted]
Baricitinib	[redacted]	[redacted]
Dupilumab	[redacted]	[redacted]

Analysis run using utility waning in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; WTP: willingness-to-pay.

Figure 2: NMB as a function of WTP per incremental QALY

Abbreviations: BSC: best supportive care; NMB: net monetary benefit; QALY: quality-adjusted life year; WTP: willingness-to-pay.

The deterministic sensitivity analyses (DSA) and probabilistic sensitivity analyses (PSA) conducted to test the robustness of the model to the uncertainties within the model parameters are presented below.

Probabilistic sensitivity analysis

Probabilistic sensitivity analyses (PSAs) with 3,000 iterations were performed for each pairwise and fully incremental comparison as described in Section B.3.8.1 of the CS and using utility waning in line with TA534 Sensitivity Analysis 1.

The probabilistic base case results are presented in Table 21 and the cost-effectiveness plane scatterplot and cost-effectiveness acceptability curve are presented in Figure 3 and Figure 4, respectively. Baricitinib has a higher probability of being cost-effective than both dupilumab and BSC at a willingness to pay (WTP) threshold of £30,000/QALY gained over the range of values tested in the model.

Table 21: Probabilistic base case results

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)
BSC	[redacted]	[redacted]	[redacted]	[redacted]	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£27,268	£27,268
Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	£89,287	£113,147

Analysis run using utility waning in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LY: life years; QALYs: quality-adjusted life years.

Figure 3: Cost-effectiveness plane scatterplot

Generated using 3,000 iterations of the PSA in which utility waning was in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; QALY: quality-adjusted life year.

Figure 4: Cost-effectiveness acceptability curve

Generated using 3,000 iterations of the PSA in which utility waning was in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; QALY: quality-adjusted life year.

Deterministic sensitivity analyses

The ten most influential variables in the DSA for the analysis of baricitinib versus dupilumab and baricitinib versus BSC are presented as tornado plots in Figure 5 and Figure 6, respectively. For the comparison of baricitinib versus dupilumab, the baseline health state utility value had the largest impact on the ICER, with the discount rates for costs and utility values and the pack cost for dupilumab also proving influential. For the comparison of baricitinib versus BSC, the baseline health state utility value had the largest impact on the ICER, proving substantially more influential than all other factors.

Figure 5: Tornado plot (ICER) of baricitinib-BSC versus dupilumab-BSC

Analysis run using utility waning in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; HSUV: health state utility value; ICER: incremental cost-effectiveness ratio.

Figure 6: Tornado plot (ICER) of baricitinib-BSC versus BSC

Analysis run using utility waning in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; EASI: Eczema Area and Severity Index; HSUV: health state utility value; ICER: incremental cost-effectiveness ratio; TCS: topical corticosteroids.

3.2. Scenario analyses results

Table 22: Results of scenario analyses (incremental)

Scenario[a]	Treatment	Total costs (£)	Total QALYs	Total LYs	Incremental costs (£)	Incremental QALYs	ICER vs BSC (£/QALY)	Incremental ICER (£/QALY)
2	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£27,037	£27,037
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£89,350	Extendedly dominated
	Baricitinib + dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£74,468	£87,918
	Dupilumab + baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£77,097	Strong Dominance
3	BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-
	Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£27,077	£27,077
	Dupilumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£91,806	£128,407

[a] All scenario analyses were run using utility waning in line with TA534 Sensitivity Analysis 1.
Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LY: life years; QALYs: quality-adjusted life years.

References

1. National Institute for Health and Care Excellence (NICE). Dupilumab for treating moderate to severe atopic dermatitis: Technology appraisal guidance [TA534]. Available at: https://www.nice.org.uk/guidance/ta534 [Last accessed: 28th January 2020].
2. NICE Decision Support Unit. Technical Support Document 5: Evidence Synthesis in the Baseline Natural History Model. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD5-Baseline.final-report.08.05.12.pdf [Last accessed: 22nd May 2020].
3. Blauvelt A, de Bruin-Weller M, Gooderham M, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corticosteroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. The Lancet 2017;389:2287-2303.
4. National Institute for Health and Care Excellence. TA511: Brodalumab for treating moderate to severe plaque psoriasis. Available at: https://www.nice.org.uk/guidance/ta511/ [Last accesse: 12th November 2020].
5. Monthly Index of Medical Specialties (MIMS). 2020. Available at: https://www.mims.co.uk/. Date accessed: 9th June 2020.
6. PSSRU: Unit Costs of Health and Social Care (2019). Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/. Last accessed: 9th June 2020.
7. NHS Improvement. National Cost Collection data: 2018-2019. Available at: https://improvement.nhs.uk/resources/national-cost-collection/. Date accessed: 9th June 2020.
8. Jin SY, Lim WS, Sung NH, et al. Combination of glucosamine and low-dose cyclosporine for atopic dermatitis treatment: a randomized, placebo-controlled, double-blind, parallel clinical trial. Dermatol Ther 2015;28:44-51.
9. Haeck IM, Knol MJ, Ten Berge O, et al. Enteric-coated mycophenolate sodium versus cyclosporin A as long-term treatment in adult patients with severe atopic dermatitis: a randomized controlled trial. J Am Acad Dermatol 2011;64:1074-84.
10. Phillippo DM AA, Dias S, Palmer S, Abrams KR, Welton NJ,. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. 2016. Available at: http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf. Date accessed: 10 November 2020.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments **13 November 2020**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential

information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	**Dr Richard Weller**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**No tobacco industry funding**

Questions for engagement

Issue 1: Patient population	
Are there any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe atopic dermatitis (AD) who have not yet had a systemic immunosuppressant?	**The choice of patients to be treated with baricitinib seems appropriate. As it is an oral drug with an apparently low side effect profile it might in time find a use earlier in the treatment journey as an alternative to current systemic medications but the economic modelling and financial viability of this would be very different.**

Issue 2: Comparators

Approximately what proportion of people would be offered the following treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants in the tables below (due to intolerance, contraindication or inadequate disease control)? Please provide a percentage for each treatment.

Following first-line **ciclosporin**	
Dupilumab (%)	10
Azathioprine (%)	20
Methotrexate (%)	60
Mycophenolate mofetil (%)	
BSC (%)	10
Other (%)	

Following first-line **methotrexate**	
Dupilumab (%)	10
Azathioprine (%)	20
Ciclosporin (%)	60
Mycophenolate mofetil (%)	
BSC (%)	10
Other (%)	

Following first-line **azathioprine**	
Dupilumab (%)	10
Ciclosporin (%)	40
Methotrexate (%)	40
Mycophenolate mofetil (%)	
BSC (%)	10
Other (%)	

Issue 3: Disease severity of patient population

Do you consider the patient population in the baricitinib trials to be generalisable to people with moderate-to-severe AD in the NHS?	**Yes. Although the recruitment criteria for the BREEZE 1 and 2 studies were for moderate to severe eczema (defined as having an EASI > 16), the data show that in fact baseline EASIs in both of these studies we're in the low 30s. I suspect this is more in keeping with the patients who will be given baricitinib in the early years. The experience with dupilumab has been that patients with more severe eczema have been enrolled first because of the backlog of patients with severe eczema waiting for novel agents. I suspect this will also apply to patients receiving baricitinib**

Issue 4: Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice	
Is EASI50 plus a ≥ 4-point improvement in DLQI a relevant outcome for determining treatment response in NHS clinical practice? Are there any other outcomes (or composite outcomes) that should be considered?	This is obviously a lower bar than an EASI 75, but an EASI of 50 still represents a significant improvement. Both EASI and DLQI scores have the advantage for NHS clinical practice that they are widely used and understood and can be performed relatively rapidly. I think that as with dupilumab there should be a decision made based on weather EASI 75 or 50 has been achieved. 75 would be regarded as success allowing continued treatment. An EASI of 50 would possibly allow continued treatment dependent on clinicians decision
Issue 5: Time to assessment of response	
Do you anticipate that response to baricitinib would be assessed at 12 or 16 weeks in NHS clinical practice?	Yes.
Would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks?	No. From the trial data shown maximum benefit occurs before 12 weeks following which there is a plateau in response so I think it makes sense to assess response at the earlier time point of 12 weeks
Issue 6: Treatment sequencing	
Assuming both baricitinib and dupilumab were available, how do you anticipate that they would likely be used in NHS practice? Please provide a rationale for your response. • Baricitinib followed by dupilumab • Dupilumab followed by baricitinib	. Although the manufacturers assess baracitinib as having equivalent effectiveness to Dupilumab by triangulating against placebo controlled trials I think they are being generous in their interpretation. The Breeze 1 and 2 studies show an EASI 75 in 21 and 25% of patients. The SOLO trials of Dupilumab showed and EASI 75 of over 50%, and this has been borne out in a recent Dutch study of real life Dupilumab use (JAAD- in press). Thus on efficacy terms, Dupilumab seems better- this would however need to be tested in a head to head clinical trial for confirmation however. Baricitinib is an oral drug which might be considered an advantage overdue belimumab. However at this stage of that disease patients will have had months or even years of having had blood tests taken for monitoring of systemic agents and I have not yet encountered a patient with

• No treatment sequencing (i.e. baricitinib or dupilumab followed by best supportive care)	moderate or severe eczema who objected to having an injection every two weeks to control it, particularly as Dupilumab avoids the need for extensive blood tests that are required with systemic treatments. I thus think the advantages of an oral over injection treatment are marginal at best.
Do you anticipate that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment? If so, why?	I think unlikely as they act via different mechanisms.
Issue 7: Modelling of best supportive care	
Which of the approaches (or combination of approaches) do you consider most appropriate for modelling best supportive care?	
Issue 8: Long-term discontinuation rates for baricitinib	
Is the company's assumption of equivalence between baricitinib and dupilumab for long-term discontinuation rates reasonable, in the absence of longer-term data for baricitinib?	No. baricitinib appears ot be less effective than Dupilumab which has a very low discontinuation rate. I suspect baricitinib discontinuation will be higher.
The ERG used data from JAHN to derive alternative discontinuation rates. Should the discontinuation rate from week 16-52 be based on the probability of response at week 52 conditional on response at week 16, or on all-cause discontinuation rates?	All cause- but response after week 16 seems fairly constant
Issue 9: Loss of utility benefit (consistency with TA534)	

Should an assumption be incorporated into the model that a proportion of patients discontinue baricitinib each year due to a loss of utility benefit, in addition to the all-cause discontinuation rates already applied? If so, is it reasonable to assume equivalence with dupilumab?	From preliminary data, I suspect benefit is constant (as is dupilumab). I suspect less loss of utility after initial drop out period.
Issue 10: Utility values	
Do you anticipate that a patient achieving a response as defined by EASI50 plus a ≥4 point improvement in DLQI would experience an improvement in health-related quality of life?	Yes.
Do you consider the utility values preferred by the company or ERG (derived from TA534) to be more appropriate?	
Issue 11: Differences in clinical outcomes based on region and skin type	
Do you consider the data from the Japanese patients recruited in the baricitinib trials to be generalisable to clinical practice in the NHS?	I think trial may be comparable, but I believe that clinical practice (i.e. access to novel drugs easier) may differ.
Would data from European patients only be more clinically relevant than the full intent-to-treat population?	European patients more typical of….European patients. Eczema may differ between ethic groups- certainly Africans- maybe Japanese too?
Is it reasonable to assume comparable efficacy for baricitinib in black patients compared to white	Not sure. Pattern of eczema is dfferent in African patients. Erythema scores in EASI also need to be adjusted to take into account darker skin

patients in the absence of robust data? If not, what adjustments should be made when assessing the effectiveness of treatment in black patients?	
Issue 12: Indirect treatment comparison heterogeneity	
Do you consider that the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the indirect treatment comparison in favour of either treatment?	No.
Issue 13: Impact of baricitinib on flare control	
Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable, in the absence of longer term data for baricitinib?	
Issue 14: Concomitant treatments	
Does the ERG's revised scenario in which the costs of bathing products are removed from the model best reflect current NHS clinical practice?	Yes. Since trial of bath oils showing no benefit, this is no longer paid for by most NHS trusts. Families may buy them privately though.
Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?	Yes.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments **13 November 2020**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential

information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	**Dr Andrew Pink**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**St John's Institute of Dermatology**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Patient population	
Are there any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe atopic dermatitis (AD) who have not yet had a systemic immunosuppressant?	**Yes, moderate to severe AD patients requiring systemic therapy (but who have not yet had a systemic agent). This agent would represent a more targeted (and potentially more tolerable) first line therapy compared with methotrexate or ciclosporin.**

Issue 2: Comparators

Approximately what proportion of people would be offered the following treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants in the tables below (due to intolerance, contraindication or inadequate disease control)? Please provide a percentage for each treatment.

Following first-line **ciclosporin**	
Dupilumab (%)	60
Azathioprine (%)	5
Methotrexate (%)	35
Mycophenolate mofetil (%)	0
BSC (%)	
Other (%)	

Following first-line **methotrexate**	
Dupilumab (%)	75
Azathioprine (%)	5
Ciclosporin (%)	20
Mycophenolate mofetil (%)	0
BSC (%)	
Other (%)	

Following first-line **azathioprine**	
Dupilumab (%)	50
Ciclosporin (%)	15
Methotrexate (%)	30
Mycophenolate mofetil (%)	0
BSC (%)	
Other (%)	

Issue 3: Disease severity of patient population

Do you consider the patient population in the baricitinib trials to be generalisable to people with moderate-to-severe AD in the NHS?	**Overall yes. It is notable however that the populations in the trials were more towards the severe end compared with average NHS population. Geographically, it is also noteworthy that there was significant Japanese enrolment, and there are differences in practice between countries.**

Issue 4: Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice

Is EASI50 plus a ≥ 4-point improvement in DLQI a relevant outcome for determining treatment response in NHS clinical practice? Are there any other outcomes (or composite outcomes) that should be considered?	Yes, I believe that a 50% improvement in eczema severity represents an important and meaningful response. It is very important to emphasise that patients can experience a dramatic improvement/ significant normalisation in their QoL with a >50% improvement (threshold effect). The DLQI remains the best measure of QoL in this context, recommended by the Harmonising Outcomes in Eczema Initiative (HOME), The key symptom that affects patients with eczema is itch, which is not adequately addressed with either EASI or DLQI. This could either be considered by requesting a >4 point difference in POEM score (which incorporates more symptoms + itch) or a > 4 point difference on the 11 Peak Pruritis NRS (both MCID, both recommended by HOME).
Issue 5: Time to assessment of response	
Do you anticipate that response to baricitinib would be assessed at 12 or 16 weeks in NHS clinical practice?	Baricitinib reaches peak effectiveness by 8 weeks in majority of patients, 12 weeks would seem like a reasonable assessment point to ensure chances are maximised. This also ties in with other systemic follow up in routine practice (e.g. methotrexate, ciclosporin) and bloods required for baricitinib as stated by SmPC (lipids)
Would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks?	It is possible that there will be some late responders, however the risk of leaving patients on an ineffective therapy for an extra 4 weeks possibly outweighs that very small %.
Issue 6: Treatment sequencing	
Assuming both baricitinib and dupilumab were available, how do you anticipate that they would likely be used in NHS practice? Please provide a rationale for your response. • Baricitinib followed by dupilumab • Dupilumab followed by baricitinib	Baricitinib may be used ahead of dupi in the following situations: 1/ Flares – it works very quickly 2/ Need dose control – dose flexibility not offered by dupi 3/ Co-morbidities – e.g.eye problems, alopecia areata 4/ Facial prominence – dupi less effective for this cohort

• No treatment sequencing (i.e. baricitinib or dupilumab followed by best supportive care)	5/ needle phobics Baritinib likely to be used after dupi if: 1/ side effects e.g. eye symptoms/joint symptoms/ facial flare 2/ treatment failures
Do you anticipate that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment? If so, why?	JAKi cover a wider array of cytokine pathways so whilst there is some possible overlap, they are different enough in both target and potency for there to be likely sequential efficacy.
Issue 7: Modelling of best supportive care	
Which of the approaches (or combination of approaches) do you consider most appropriate for modelling best supportive care?	Best supportive care in AD is emollients and TCS (+/- TCI). Drug + TCS studies are therefore more pragmatic and real world e.g. BREEZE AD7
Issue 8: Long-term discontinuation rates for baricitinib	
Is the company's assumption of equivalence between baricitinib and dupilumab for long-term discontinuation rates reasonable, in the absence of longer-term data for baricitinib?	Baricitinib is potentially associated with more tolerability issues than dupilumab from available data, hence longterm discontinuation may be slightly higher, but hard to say as safety profile still on the whole reassuring. Dosing flexibility with 2/4mg options (if an option), may in part negate that effect too.
The ERG used data from JAHN to derive alternative discontinuation rates. Should the discontinuation rate from week 16-52 be based on the probability of response at week 52 conditional on response at week 16, or on all-cause discontinuation rates?	I am not sure I am qualified to answer this.

Issue 9: Loss of utility benefit (consistency with TA534)	
Should an assumption be incorporated into the model that a proportion of patients discontinue baricitinib each year due to a loss of utility benefit, in addition to the all-cause discontinuation rates already applied? If so, is it reasonable to assume equivalence with dupilumab?	Yes Hard to say re: dupilumab as one would have predicted that an oral small molecule inhibitor is likely to have a better drug survival than a monoclonal antibody (where secondary failure can be a significant problem). That said, interstingly we have not observed secondary failure with dupi so I would predict that ther may be equilolence.
Issue 10: Utility values	
Do you anticipate that a patient achieving a response as defined by EASI50 plus a ≥4 point improvement in DLQI would experience an improvement in health-related quality of life?	Yes, very much so. As mentioned above, eczema patients often don't need to get 90-100% better to get back to almost normal function, there is a threshold above which they do very well and that is often around 50%.
Do you consider the utility values preferred by the company or ERG (derived from TA534) to be more appropriate?	Unable to comment.
Issue 11: Differences in clinical outcomes based on region and skin type	
Do you consider the data from the Japanese patients recruited in the baricitinib trials to be generalisable to clinical practice in the NHS?	Relevant, and obviously important to review the full dataset (for which the trials were powered). Also important however to see if key outcome data differences exist between European/ Japanese trial populations as clinical practice and patient behaviours vary significantly between cultures. Optimally, European datasets would be more reliably generalisable.
Would data from European patients only be more clinically relevant than the full intent-to-treat population?	See above

Is it reasonable to assume comparable efficacy for baricitinib in black patients compared to white patients in the absence of robust data? If not, what adjustments should be made when assessing the effectiveness of treatment in black patients?	Yes, pending more data. There are emerging immunological studies indicating subtle differences in the immune phenotype between ethnicities. That said Th2 responses are key across all skin types. Furthermore, JAKi affect multiple cytokine pathways which should impact all skin types.
Issue 12: Indirect treatment comparison heterogeneity	
Do you consider that the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the indirect treatment comparison in favour of either treatment?	Key differences in trials. One being the wash out period for topical steroids -1 week for dupi, 2 weeks for bari. The extra week makes a huge difference (many patients would drop out before reaching baseline, which with a primary censoring rule stating that they are then non-reponders would then significantly affect the results. It is much more real world to look at data including patients that required topical rescue in the monotherapy trials (secondary censoring, as that is like real practice – patients do require topical top up at times with these treatments).
Issue 13: Impact of baricitinib on flare control	
Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable, in the absence of longer term data for baricitinib?	Yes in the absence of any data on flares.
Issue 14: Concomitant treatments	
Does the ERG's revised scenario in which the costs of bathing products are removed from the model best reflect current NHS clinical practice?	Yes, no evidence for benefit of bath additives

Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?	Yes

NICE National Institute for Health and Care Excellence

Technical engagement response form

Baricitinib for treating moderate to severe atopic dermatitis [ID1622]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments **13 November 2020**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential

information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Sanofi**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

Questions for engagement

Issue 1: Patient population	
Are there any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe atopic dermatitis (AD) who have not yet had a systemic immunosuppressant?	

Issue 2: Comparators

Approximately what proportion of people would be offered the following treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants in the tables below (due to intolerance, contraindication or inadequate disease control)? Please provide a percentage for each treatment.

Following first-line **ciclosporin**	
Dupilumab (%)	
Azathioprine (%)	
Methotrexate (%)	
Mycophenolate mofetil (%)	
BSC (%)	
Other (%)	

Following first-line **methotrexate**	
Dupilumab (%)	
Azathioprine (%)	
Ciclosporin (%)	
Mycophenolate mofetil (%)	
BSC (%)	
Other (%)	

Following first-line **azathioprine**	
Dupilumab (%)	
Ciclosporin (%)	
Methotrexate (%)	
Mycophenolate mofetil (%)	
BSC (%)	
Other (%)	

Issue 3: Disease severity of patient population	
Do you consider the patient population in the baricitinib trials to be generalisable to people with moderate-to-severe AD in the NHS?	

Issue 4: Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice	
Is EASI50 plus a ≥ 4-point improvement in DLQI a relevant outcome for determining treatment	

response in NHS clinical practice? Are there any other outcomes (or composite outcomes) that should be considered?	
Issue 5: Time to assessment of response	
Do you anticipate that response to baricitinib would be assessed at 12 or 16 weeks in NHS clinical practice?	
Would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks?	
Issue 6: Treatment sequencing	
Assuming both baricitinib and dupilumab were available, how do you anticipate that they would likely be used in NHS practice? Please provide a rationale for your response. • Baricitinib followed by dupilumab • Dupilumab followed by baricitinib • No treatment sequencing (i.e. baricitinib or dupilumab followed by best supportive care)	We do not currently have evidence to support the sequential use dupilumab and baricitinib in clinical practice.
Do you anticipate that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment? If so, why?	We do not currently have evidence to inform to effectiveness of dupilumab and baricitinib in sequential use.

Issue 7: Modelling of best supportive care	
Which of the approaches (or combination of approaches) do you consider most appropriate for modelling best supportive care?	
Issue 8: Long-term discontinuation rates for baricitinib	
Is the company's assumption of equivalence between baricitinib and dupilumab for long-term discontinuation rates reasonable, in the absence of longer-term data for baricitinib?	We agree with the ERG that baricitinib and dupilumab have different mechanisms of actions and routes of administration and are therefore likely to differ in long-term efficacy and adherence.
The ERG used data from JAHN to derive alternative discontinuation rates. Should the discontinuation rate from week 16-52 be based on the probability of response at week 52 conditional on response at week 16, or on all-cause discontinuation rates?	
Issue 9: Loss of utility benefit (consistency with TA534)	
Should an assumption be incorporated into the model that a proportion of patients discontinue baricitinib each year due to a loss of utility benefit, in addition to the all-cause discontinuation rates already applied? If so, is it reasonable to assume equivalence with dupilumab?	
Issue 10: Utility values	

Do you anticipate that a patient achieving a response as defined by EASI50 plus a ≥4 point improvement in DLQI would experience an improvement in health-related quality of life?	
Do you consider the utility values preferred by the company or ERG (derived from TA534) to be more appropriate?	It is recognised that small sample sizes can pose issues when estimating robust utility weights and so the use of utility weights from TA534 are likely to be appropriate to support aspects of the baricitinib cost-utility analysis. It is believed that applying the utility weight of a dupilumab responder to that of a baricitinib responder may be an overestimation of the true health-related quality of life experienced by these patients for the following reasons: i) Mixed model regression used to derive utility weights in TA534 included a number of clinical outcome covariates using data observed in dupilumab pivotal trials. It has not been shown that these clinical outcomes are comparable in the baricitinib trials. Most importantly, the regression model includes a statistically significant treatment covariate. This indicates that dupilumab treatment increases the health-related quality of life beyond the benefit attributable to clinical disease improvement. ii) Utility weights take into consideration utility benefits as well as disutilities of treatments. Utility weights derived from dupilumab studies have shown that health-related quality of life improves with treatment despite the incidence of adverse effects. Applying dupilumab-responder utility weights to baricitinib-responders does not take into consideration the impact of the baricitinib adverse effect profile on health-related quality of life. iii) Evidence presented in the TA534 submission demonstrated a reduction in the use of rescue medication including topical corticosteroids, systemic corticosteroids and immunosuppressants. It also provided evidence for the patient burden of regular and repeated use of emollients and topical therapies. The burden of maintenance therapies on baricitinib responders should be considered when assuming HRQoL equivalence between therapies.

Issue 11: Differences in clinical outcomes based on region and skin type	
Do you consider the data from the Japanese patients recruited in the baricitinib trials to be generalisable to clinical practice in the NHS?	
Would data from European patients only be more clinically relevant than the full intent-to-treat population?	
Is it reasonable to assume comparable efficacy for baricitinib in black patients compared to white patients in the absence of robust data? If not, what adjustments should be made when assessing the effectiveness of treatment in black patients?	
Issue 12: Indirect treatment comparison heterogeneity	
Do you consider that the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the indirect treatment comparison in favour of either treatment?	
Issue 13: Impact of baricitinib on flare control	
Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable, in the absence of longer term data for baricitinib?	

Issue 14: Concomitant treatments	
Does the ERG's revised scenario in which the costs of bathing products are removed from the model best reflect current NHS clinical practice?	
Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?	

Single Technology Appraisal (STA)

Baricitinib for the Treatment of Moderate-to-Severe Atopic Dermatitis

ERG addendum: review of company's response to technical engagement

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

Date completed 26th November 2020

Source of funding

NIHR Systematic Reviews Programme (project number 131739)

Declared competing interests of the authors

None.

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.

Note on the text

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

Table of Contents

Table of Tables

List of abbreviations

AD	Atopic dermatitis
BSC	Best supportive care
CMU	Commercial medicine unit
CS	Company submission
DLQI	Dermatology life quality index
ERG	Evidence review group
HRQoL	Health related quality of Life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ITC	Indirect treatment comparison
MAIC	Matching-adjusted indirect comparison
NES	National Eczema Society
NICE	National Institute for Health and Care Excellence
QALY	Quality adjusted life-year
SCORAD	Scoring atopic dermatitis
SmPC	Summary of Product Characteristics

1 OVERVIEW

This addendum to the Evidence Review Group (ERG) report provides the ERG critique of the additional evidence provided by Eli Lilly (the company) in their response to the draft Technical Report for the appraisal of Baricitinib for the treatment of moderate to severe atopic dermatitis.

The draft Technical Report outlined 14 key issues for consideration and provides the technical team's preliminary scientific judgement on each issue. The company's response to the draft Technical Report indicated that they accepted the technical team's preliminary judgement on some issues, which the ERG now considers resolved (Table 1). The company's response to all issues, along with relevant stakeholder responses, are discussed in Section 2.

Table 1 Questions for engagement and current status regarding issue resolution

Issue 1: Patient population	
1. Are there any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe atopic dermatitis (AD) who have not yet had a systemic immunosuppressant?	Resolved but uncertainty remains
Issue 2: Comparators	
2. Approximately what proportion of people would be offered the following treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants in the tables below (due to intolerance, contraindication or inadequate disease control)? Please provide a percentage for each treatment.	Resolved but uncertainty remains
Issue 3: Disease severity of patient population	
3. Do you consider the patient population in the baricitinib trials to be generalisable to people with moderate-to-severe atopic dermatitis (AD) in the NHS?	Unresolved
Issue 4: Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice	
4. Is EASI50 plus a ≥ 4-point improvement in DLQI a relevant outcome for determining treatment response in NHS clinical practice? Are there any other outcomes (or composite outcomes) that should be considered?	Unresolved
Issue 5: Time to assessment of response	
5. Do you anticipate that response to baricitinib would be assessed at 12 or 16 weeks in NHS clinical practice?	Resolved
6. Would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks?	Resolved
Issue 6: Treatment sequencing	
7. Assuming both baricitinib and dupilumab were available, how do you anticipate that they would likely be used in NHS practice? Please provide a rationale for your response. • Baricitinib followed by dupilumab • Dupilumab followed by baricitinib No treatment sequencing (i.e. baricitinib or dupilumab followed by best supportive care)	Unresolved
8. Do you anticipate that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment? If so, why?	Unresolved

Issue 7: Modelling of best supportive care	
9. Which of the approaches (or combination of approaches) do you consider most appropriate for modelling best supportive care?	Unresolved
Issue 8: Long-term discontinuation rates for baricitinib	
10. Is the company's assumption of equivalence between baricitinib and dupilumab for long-term discontinuation rates reasonable, in the absence of longer-term data for baricitinib?	Unresolved
12. The ERG used data from JAHN to derive alternative discontinuation rates. Should the discontinuation rate from week 16-52 be based on the probability of response at week 52 conditional on response at week 16, or on all-cause discontinuation rates?	Unresolved
Issue 9: Loss of utility benefit (consistency with TA534)	
13. Should an assumption be incorporated into the model that a proportion of patients discontinue baricitinib each year due to a loss of utility benefit, in addition to the all-cause discontinuation rates already applied? If so, is it reasonable to assume equivalence with dupilumab?	Unresolved
Issue 10: Utility values	
14. Do you anticipate that a patient achieving a response as defined by EASI50 plus a ≥4-point improvement in DLQI would experience an improvement in health-related quality of life?	Unresolved
15. Do you consider the utility values preferred by the company or ERG (derived from TA534) to be more appropriate?	Unresolved
Issue 11: Differences in clinical outcomes based on region and skin type	
16. Do you consider the data from the Japanese patients recruited in the baricitinib trials to be generalisable to clinical practice in the NHS?	Unresolved
17. Would data from European patients only be more clinically relevant than the full intent-to-treat population?	Unresolved
18. Is it reasonable to assume comparable efficacy for baricitinib in Black patients compared to White patients in the absence of robust data? If not, what adjustments should be made when assessing the effectiveness of treatment in black patients?	Resolved but uncertainty remains
Issue 12: Indirect treatment comparison heterogeneity	
19. Do you consider that the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the indirect treatment comparison in favour of either treatment?	Resolved but uncertainty remains
Issue 13: Impact of baricitinib on flare control	
20. Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable, in the absence of longer-term data for baricitinib?	Resolved
Issue 14: Concomitant treatments	
21. Does the ERG's revised scenario in which the costs of bathing products are removed from the model best reflect current NHS clinical practice?	Resolved
22. Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?	Resolved

2 DESCRIPTION AND CRITIQUE OF ADDITIONAL EVIDENCE

2.1 Issues 1 & 2: Patient population/comparators

Two questions were raised by the technical team: whether there are any other relevant populations that should be considered in this appraisal, for example adults with moderate-to-severe atopic dermatitis (AD) who have not yet had a systemic immunosuppressant; and approximately what proportion of people would be offered certain treatments in routine NHS practice, following failure on the first-line systemic immunosuppressants, due to intolerance, contraindication or inadequate disease control.

The company do not consider any further populations to be relevant for consideration in this appraisal. The company acknowledge that there is an unmet need in UK clinical practice for alternatives to current systemic immunosuppressants and therefore, that adults with moderate-to-severe AD who have not yet had systemic immunosuppressants may be a relevant population for baricitinib. However, ciclosporin is the only treatment with marketing authorisation in this population.

2.1.1 Indirect comparison to ciclosporin

The company conducted a matching-adjusted indirect comparison (MAIC) of ciclosporin versus baricitinib due to the absence of direct evidence. One study was deemed eligible for inclusion, Haeck et al 2011,[1] which compared ciclosporin to mycophenolate sodium. These data were compared to the BREEZE-AD JAIY trial data, which resulted in a disconnected network. "Unanchored" population adjustment may be considered in the absence of a connected network of randomised studies, or where single-arm studies are involved.[2] The company conducted an unanchored MAIC, which effectively assumes that all effect modifiers and prognostic factors are accounted for. This assumption is very strong, and largely considered impossible to meet. In this case, the MAIC only adjusted for a few variables: sex, scoring atopic dermatitis (SCORAD) score and dermatology life quality index (DLQI) score, however there are other potential effect modifiers and prognostic factors which have not been accounted for. Therefore, there may be an unknown amount of bias in the unanchored estimate. The study by Haeck et al.[1] only reported efficacy SCORAD data at week 16, whereas, the JAIY trial reported EASI 75 data. Therefore, SCORAD values were mapped to EASI 75 values, as was done in TA534,[3] which resulted in an extreme distribution of weights and an effective sample size of [redacted] patients. Additionally, the MAIC adjusts to the population characteristics of the aggregate data, i.e. the study of ciclosporin, which may not be representative of UK. Thus, the ERG agrees with the company that the results from the MAIC are not reliable and should not be used to inform decision making.

2.1.2 Comparators in UK clinical practice

In response to estimating the proportion of people who would be offered various treatments in UK clinical practice following failure of three first-line treatments, the company stated that the majority of patients would be given dupilumab after first-line treatment with ciclosporin and methotrexate. However, this contradicted the clinical expert, who stated that only 10% of patients would be given dupilumab after first-line ciclosporin and methotrexate. This suggests that at least two systemic immunosuppressants are given before moving onto dupilumab. Clinical advice to the ERG is that baricitinib acts in a similar manner to other systemic immunosuppressants such as methotrexate and ciclosporin in targeting a broader range of cellular processes and mediators than dupilumab.[4] Therefore, in practice, it is expected that baricitinib would be given after topical treatment fails, when systemic immunosuppressants are considered. The National Eczema Society (NES) also states that patients with moderate-to-severe eczema for whom topical treatments are insufficiently effective and who must progress to second-line treatments would benefit from the introduction of a new second-line treatment option. Thus, suggesting that systemic immunosuppressants are relevant comparators to baricitinib.

The company states that whilst dupilumab is effective in controlling disease, there are considerable limitations to its use, such as injection-site reactions and eye disorders. However, the clinical expert comments to the technical engagement suggest that the advantages of an oral drug compared with an injection treatment are marginal at best. At this stage of the disease, patients have had months or even years of blood tests to monitor systemic agents. The clinical expert noted at technical engagement that none of their patients with moderate or severe eczema has objected to having an injection every two weeks to control the disease, particularly as dupilumab avoids the need for extensive blood tests that are required with systemic treatments. The ERG notes that baricitinib also has a poor safety profile, with a higher proportion of patients in the 4 mg baricitinib group of the JAIN trial (75.0%) experiencing at least one Treatment Emergent Adverse Event compared to the placebo group (53.8%).

The ERG acknowledges that there is limited evidence comparing systemic immunosuppressants and baricitinib and it is therefore difficult to assess comparative efficacy in this patient population. However, the clinical expert at technical engagement stated that, as baricitinib is an oral drug with an apparently low side-effect profile, it might in time find a use earlier in the treatment journey as an alternative to current systemic medications, but the economic modelling and financial viability of this would be very different. This supports the ERG's view that there is an unmet need for alternatives to systemic immunosuppressants and in practice baricitinib may be used at this point in the treatment pathway.

2.2 Issue 3: Disease severity of patient population

The issue of whether the patient population in the baricitinib trials is generalisable to people with moderate-to-severe AD in the NHS was raised by the technical team.

The company state that the patient population in the BREEZE-AD trials is generalisable to people with moderate-to-severe AD in the NHS. However, the ERG maintains that in terms of disease severity, the patients in the trial populations may not represent all moderate-to-severe patients in the NHS population. The trial inclusion criteria may exclude patients on the lower end of the moderate scale and mean EASI scores indicate that the patients included are more likely to have severe AD, as discussed in Section 3.2.2 of the ERG report.

The company states that no definitive and widely accepted cut-off exists between moderate and severe AD, with classification based on a variety of clinical factors in UK clinical practice. As per the conclusions of a steering committee consisting of a multidisciplinary group of AD experts, the factors for diagnosis of moderate-to-severe AD include consideration of body surface area affected, individual lesion severity, lesion location and/or quality of life impairment. Standardised scales, such as EASI score, can provide one classification system, but do not capture all of these criteria and therefore may not be reflective of all aspects of disease, particularly given that this scoring could provide an inconsistent classification of a flaring disease. However, the ERG maintains that there are published strata that allow classification by EASI score. Clinical advice to the ERG is that EASI is widely accepted and considers all relevant aspects of the clinical signs of AD. Therefore, it would have been appropriate for the company to consider classification of moderate and severe AD by EASI score.

Additionally, the company state analysis of moderate versus severe subgroups are prevented by a lack of available data for these populations for dupilumab. Together, these clinical and data limitations mean that it is not feasible to conduct any efficacy comparisons with the key comparator in these populations. However, the ERG considers that subgroup analyses based on AD severity in the BREEZE-AD trials would still be feasible and beneficial.

2.3 Issues 4 and 10: Response criteria and utilities used in the model

Three questions were asked by the Technical team in relation to these: i) is EASI50 plus a ≥ 4 DLQI a relevant outcome for assessing response; ii) would a response (as defined above) be related to improvements in health related quality of life (HRQoL); iii) which set of utility values are more appropriate those preferred by the company or ERG (derived from TA534)?

2.3.1 Issue 4: Relevance of EASI 50 plus a ≥ 4-point improvement in DLQI outcome in clinical practice

Aligning with assumptions made in TA534 the original company base-case analysis used a composite endpoint based on EASI50 plus ≥ 4 DLQI to categorise a patient's response to treatment. In the ERG report, this consistency with TA534 was acknowledged, but several concerns were raised questioning whether the EASI/DLQI was recognised or treated as clinically meaningful in practice. It was also noted that EASI75 was the primary trial outcome in the JAIN study and that this measure was also considered a clinically significant improvement by the British Association of Dermatologists. Importantly, it was also noted that there is no correlation between response and HRQoL in the company's regression analysis of JAIN and JAIN-like JAIY patients, which suggests that these criteria do not reflect the benefits of treatment.

Acknowledging the concerns raised by the ERG, the company outlines in their response that they now consider EASI75 to represent a more relevant outcome measure than the composite EASI50/DLQI outcome. The company justified this new position noting that EASI75 formed the primary or key secondary endpoint in all of the BREEZE-AD trials and that it is widely recognised as a clinically relevant outcome, as supported by clinical expert opinion sought by the ERG, NICE Technical Team, and the company. The company further suggests that there is also a move towards using EASI75 as an assessment criterion in the NHS.

In principle, the ERG agrees that EASI75 may represent a better and more clinically relevant response criterion than EASI50 plus ≥ 4 DLQI. The revised base-case presented by the company using EASI75 is therefore potentially reasonable, particularly given the limitations of the utility data available. The ERG does have some concerns; one regarding the implementation and another more conceptual criticism.

With regards to the implementation, the ERG notes that the estimates of response were generated using JAIN plus JAIY JAIN-like patients to model response in the best supportive care (BSC) arm of the model. The ERG's base-case used, both JAIN plus JAIY JAIN-like and CAFÉ plus CHRONOS CAFÉ-like patients and we considered this a more reasonable approach given the relevance of all the contributing trials to the decision problem and the use of data from all these studies in the meta-analysis.

There are also important limitations to using EASI75 as the response criterion because the analysis presented by the company imposes the EASI75 criterion not just on baricitinib but also on BSC and dupilumab. The application of EASI75 to dupilumab is particularly problematic given that current NICE guidance specifies that the composite EASI 50/DLQI criteria should be used to assess response.

The application of the EASI75 criterion to dupilumab is therefore inconsistent with current guidance and, insofar as it is adhered to, current practice.

To address this issue, one could consider an approach where the EASI 50/DLQI criteria are retained for patients receiving dupilumab, and EASI75 is used to evaluate response in patients receiving baricitinib. This approach would, however, have important practical implications and it is questionable whether applying two different criteria would be acceptable to clinicians and patients, particularly if, as the company desires, baricitinib is to be evaluated as a mutually exclusive alternative to dupilumab. There are also considerable challenges with implementing this approach in terms of generating a reliable comparison and applying this within the economic model.

We would need to assume that the different definitions of response in the control (placebo) arm, which acts a common comparator for the indirect treatment comparison (ITC), yield comparable results, and that 'response' in patients is comparable regardless of the measure used, which may not be a clinically meaningful assumption to make. However, the ERG notes that the OR of response for the pooled baricitinib trials (using secondary censoring) are very similar for both measures: [redacted] for EASI75 and [redacted] for EASI50/DLQI (ERG report Tables 68 and 34, respectively).With respect to the application of the model, these challenges are largely a product of how we define the health states and link them to utilities and costs. The application of two different response rates would necessarily require modifications to the current model structure, such that there would need to be two responder health states and two non-responder health states, with each set defined by the alternative response criteria. While theoretically, this is not a problem, pragmatically the available data could not support such a model structure. There is, for example, no utility set that could be used to generate a consistent and coherent analysis, without resorting to strong assumptions about the impact of response on HRQoL. Given these challenges and the questionable clinical validity of using two different criteria, the ERG does not attempt to produce a scenario that addresses this specific issue and retains the composite EASI50/ DLQI response criteria in our base-case. Given the potential advantages of the EASI75 outcome measure, scenario analysis is, however, also presented using the EASI75 criterion for BSC, baricitinib and dupilumab.

2.3.2 Issue 10: Utility values

As noted above, part of the justification for using EASI75 to assess response was that no quality of life benefits could be demonstrated using the EASI50/DLQI endpoint. In response to this issue, the revised company base-case therefore not only revises the response criteria applied, but also revises the utility set applied so that they align with the new criteria. These revised utilities are based on a regression analysis of data from JAIN and JAIN-like JAIY patients which are used to estimate coefficients for the response categories EASI<50, EASI50 to 75, and EASI75. These coefficients are then applied in the model to estimate the utilities applied to responders and non-responders.

The ERG is satisfied with the approach adopted by the company and notes that the predicted utilities suggest a utility gain from response. The ERG, however, is disappointed that the company chose to only use data from JAIN and JAIN-like JAIY patients and not all available observations on JAIN-like patients across the pivotal trials. As noted in the ERG report, the JAHL and JAHM trials are potentially relevant sources of data for this input and there is no clear rationale to exclude these observations.

2.3.3 Modelling EASI50/DLQI

Given the outlined difficulties of applying the EASI 75 to assess response, it may be necessary to accept that baricitinib should be evaluated using the composite EASI50/DLQI endpoint so as to align with previous guidance. The use of the EASI50/DLQI, however, poses a problem as the revised company base-case and new utility set do not directly address the issues raised in the ERG report regarding the face-validity of the utility values reported for the EASI50/DLQI response criteria. We are therefore still left with the issue of what is the most appropriate utility set to apply. In the ERG's revised base-case, we opted to use data from TA534 to populate the model. This choice was, however, largely directed by the absence of any (face) valid alternatives and the ERG notes that these assumptions were criticised by Sanofi in their submission. Specifically, the Sanofi submission highlighted the following:

- That dupilumab increases the HRQoL beyond the benefit attributable to clinical disease improvement. And that no such equivalent evidence is available for baricitinib.
- That the applied utility set does not take into account the adverse event profiles of the two treatments.
- That the use of dupilumab is associated with significant reductions in rescue medications the use of which places a burden on the patient. And that the burden of maintenance therapies on baricitinib responders should be considered when assuming HRQoL equivalence between therapies.

The issues raised by Sanofi represent a valid criticism of the approach adopted by the ERG and we agree that the listed factors may contribute to differences in HRQoL benefits between baricitinib and dupilumab. Generating an appropriate comparison of utility benefits is, however, very difficult given the lack of a common data set, and it is not clear that any such comparison would necessarily favour dupilumab, despite Sanofi's contentions otherwise. Furthermore, the lack of an alternative utility set means that pragmatic assumptions are necessary, and the ERG does not consider it unreasonable to use the values reported in TA534.

2.4 Issue 5: Time to assessment of response

Two questions were raised by the technical team: would the response to baricitinib be assessed at 12 or 16 weeks in NHS clinical practice; and would there be any disadvantages to assessing response to baricitinib at 12 weeks rather than 16 weeks.

The company consider Week 16 to be the most suitable timepoint for assessment of response to baricitinib in the economic model because assessment of response at Week 16 was selected in alignment with the primary timepoint of response assessment performed in the pivotal BREEZE-AD trials. Furthermore, the dupilumab trials used a 16-week timepoint for response assessment, and thus the ITC performed using data from the Week 16 timepoint provides the most robust comparative efficacy estimates for baricitinib versus dupilumab.

The wording in the recently published summary of product characteristics (SmPC) for baricitinib has been updated since the draft version shared at the time of submission. The SmPC now states that consideration should be given to discontinuing treatment in patients who show no evidence of therapeutic benefit after 8 weeks of treatment. The ERG is satisfied that 16 weeks is a reasonable timepoint at which to assess response given that a 16-week assessment period is already in use for dupilumab, and that this was the time point at which all primary and key secondary outcomes were assessed in the BREEZE-AD trial programme. However, the ERG remains concerned that 16 weeks is not mentioned in the SmPC and that current wording could lead some patients to discontinue treatment before 16 weeks.

2.5 Issue 6: Treatment sequences

The Technical Team raised two questions with respect to this issue. The first asked whether baricitinib would be used as part of a treatment sequence that included dupilumab or whether it would be used as a standalone alternative to dupilumab. The second asked whether there was an expectation that the efficacy of baricitinib or dupilumab would be impacted by the prior use of the other treatment?

In their response, the company states that it is not their desire to position baricitinib as part of a treatment sequence that includes dupilumab. Instead, the company states that baricitinib should be considered as an alternative to dupilumab and positioned identically to dupilumab in the treatment pathway. The company further highlights that there are no data to inform the effectiveness of baricitinib and dupilumab when used as part of a sequence and that the modelling of baricitinib and dupilumab in a sequence requires the simplifying assumption that efficacy is unchanged, regardless of positioning.

The ERG acknowledges the desire of the company to position baricitinib as an alternative to dupilumab, but as outlined in the ERG report, we consider there to be a substantial clinical appetite to use these treatments sequentially. This is indeed alluded to in the company's response where they highlight unmet needs in patients who have failed dupilumab (Issue 1 response). Further, the ERG feels it is important to note that the treatment sequence analysis does not supplant the one presented by the company, but instead simply extends it to allow the committee to consider how the cost-effectiveness of baricitinib is impacted by its placement either as a standalone alternative to dupilumab or as part of a sequence that includes both treatments. Thus, it is important to reiterate that the comparison being presented by the company is assessing each treatment as a mutually exclusive alternative. Consequently, a literal interpretation of this analysis would minimally require that a recommendation for baricitinib prohibit the sequential use of alternatives. As highlighted in our report this may be considered undesirable given the clear treatment benefits of dupilumab, the high unmet need in patients who have failed dupilumab, and the lack of alternatives for this patient group. Importantly, it may also not represent the most cost-effective positioning of baricitinib.

With regards to the availability of clinical data to inform this analysis, the ERG acknowledges the lack of data and the need to make simplifying assumptions. We also note that similar comments were made by Sanofi regarding the availability of data to inform these scenarios. Nonetheless, treatment sequence is an important issue and the clinical reality is such that these treatments will be used sequentially unless explicitly prohibited. Furthermore, the implied assumption that efficacy is unchanged, regardless of positioning, is one commonly required in the modelling of chronic diseases and there is considerable precedent for similar assumptions being accepted. The ERG additionally highlights expert opinion received as part of technical engagement which stated that the efficacy of baricitinib and dupilumab is unlikely to be impacted by the prior use of the other treatment due to the different mechanism of action.

2.6 Issues 7 and 9: Health related quality of life for best supportive care patients

Issues 7 and 9 are considered together as they both concern how BSC is modelled and the quality of life of these patients. The Technical Team asked two questions in relation to these issues. The first asked what is the best approach to modelling BSC? The second asked whether an assumption should be incorporated into the model, whereby patients discontinue baricitinib each year due to a loss of utility benefit, in addition to the all-cause discontinuation rates already applied?

These issues reflected concerns raised by the ERG that the model does not account for the waxing and waning nature of AD, nor how treatment is currently used to address this pattern of disease. These omissions are important and mean that patients who don't respond, or lose their response to treatment, remain in a state of chronic and severe AD until death. This is inconsistent with clinical reality as the

concept of discontinuation cannot be applied to BSC and misrepresents the effectiveness of BSC as patients will have periods of both good and poor disease control. This is reflected in longer-term data from the CHRONOS study where we observe relative stability in response rates for the Placebo arm across the week 16 to 52 period. The ERG's base-case analysis therefore did not model discontinuation in BSC patients and instead assumed that a constant proportion of patients receiving BSC would remain in the response health state.

In response to these issues, the company acknowledges the issues raised by the ERG but are concerned about the implications of the assumptions applied in the ERG base-case. The company explains that using placebo to proxy BSC may be problematic given the non-trivial response to placebo observed in the BREEZE-AD trials and contend that from a clinical perspective, patients who have previously failed the constituents of BSC cannot be expected to achieve any appreciable benefits from the continued use of BSC. The company, therefore, considers the assumption of sustained effectiveness beyond the trial period to be highly implausible and notes that the model assumptions represent a substantial departure from the TA534 model. The company further explains that these assumptions represent a failure to acknowledge the presence of an efficacy gap for topical treatments in dermatological clinical practice and highlight evidence suggesting that topical medicines are poorly adhered to (suggesting lack of efficacy) and that use of topical medication is associated with substantial disutility.

Reflecting the company's concerns with the ERG base-case assumptions, it proposes alternative assumptions based on those applied in TA534. In these scenarios, the assumption of no discontinuation for BSC is retained as per the ERG base-case, but utility waning is applied such that only a proportion of patients maintain the utility benefit associated with response. Details of the specific assumptions applied are included in the appendix to the company's response.

To properly consider the appropriateness of the proposed alternative methods of modelling BSC and the arguments put forward by the company, it is important to interpret the ERG base-case assumptions correctly. The ERG is not suggesting that patients receiving BSC will enjoy a sustained response from topical medicines, nor are we suggesting that the observed placebo effects will be infinitely durable. Instead, the ERG is suggesting that the waxing and waning nature of AD will imply that an individual patient will experience periods of both good and poor disease control and that the observed placebo response rates are a reasonable estimate of the proportion of BSC patients who achieve good disease control at any given point in time. Importantly, the assumption that patients receiving BSC may achieve temporary disease control stands in stark contrast to the model assumptions for biologic therapy which imply durable and stable disease control while on treatment. In the view of the ERG, these contrasting assumptions, therefore, imply a clear benefit of biological treatment and do not, as the company asserts, imply a lack of an efficacy gap. Furthermore, the suggestion that the observed

placebo effects are unrepresentative of clinical reality is a highly selective interpretation of the evidence. It implies that the observed outcomes for placebo patients are not generalisable to practice, while simultaneously suggesting that the trial is generalisable in all other respects including the observed response rates for patients receiving biologic therapy.

The noted lack of adherence to topical corticosteroids and other elements of BSC also serves to misrepresent the ERG's position which is founded not on the efficacy-effectiveness of topical medicines, but rather the underlying nature of AD i.e. that AD is characterised by waxing and waning of symptoms. The efficacy or indeed lack of efficacy of topical medicines would therefore not necessarily imply a different pattern of response amongst BSC patients. Furthermore, the observed correlation between HRQoL and the use of topical medicines does not necessarily imply a lack of efficacy but may instead simply reflect the fact that patients with poor disease control (and lower HRQoL) are likely to make greater use of these medicines.

The ERG also has concerns regarding the appropriateness of the utility waning assumptions. Firstly, the assumptions are not supported by any evidence. The assumed rates of decline in utility benefits are therefore purely speculative and are based solely upon clinical opinion; this contrasts with the assumption in the ERG base-case which was informed by evidence from CHRONOS. In this regard, it is notable that while the committee agreed with the concept of utility waning, there was no agreement on the rates of utility waning that should be applied. Secondly, while the ERG acknowledges that these assumptions were accepted in TA534, there is no evidence that a similar approach has been adopted in the appraisal of biologics for other inflammatory conditions such as psoriasis and psoriatic arthritis. [5-12] The balance of precedent therefore does not support the use of utility waning. Thirdly, the application of utility waning is methodologically flawed as it serves to undermine the assumption that the health states define the symptom burden of patients. This creates the contradictory scenario whereby the loss of efficacy implied by utility waning impacts only on quality of life but has no impact on care costs despite assumptions to the contrary elsewhere in the model.

2.7 Issue 8: Long-term discontinuation rates for baricitinib

With regards to this issue two questions were raised by the Technical Team. The first asked whether the company's assumption of equivalent discontinuation rates for baricitinib and dupilumab was reasonable? The second asked whether the discontinuation rate applied in Weeks 16 to 52 should be based on conditional response rates (the proportion of patients who continue to respond) or all-cause discontinuation rates?

In the company's response, it is acknowledged that there is uncertainty regarding the appropriate discontinuation rates to apply and that the assumption of equivalence adopted in the company's base-case reflected the lack of long-term data to support alternative assumptions for baricitinib. The

availability of a new (52-week) data cut from the JAIN study means that these assumptions are no longer necessary. The revised company base-case, therefore, presents analysis using these new data to inform the discontinuation rates for baricitinib. In implementing this new analysis, the company adopted an approach consistent with TA534, in which conditional response rates were used to model discontinuation between Weeks 16 and 52, with all-cause discontinuation rates applied in the post-52-week period. The company justified this decision on the grounds that the conditional response probability is a more appropriate proxy of "sustained response" than all-cause discontinuation rates, adding that patients in clinical practice are unlikely to continue on a treatment that is not effective.

The ERG welcomes the use of the new data, and considers the JAIN trial a suitable source of discontinuation data consistent with other assumptions and data sources used in the model. Importantly, the ERG notes that the JAIN study evaluated baricitinib combination therapy and in this regard is superior to the data used in the original ERG base-case where discontinuation rates were based upon the JAHN monotherapy trial. The ERG, however, does not agree with the company's use of conditional response to model discontinuation in the 16- to 52-week period. As outlined in our report, we acknowledge that this approach was accepted in TA534 and that loss of efficacy may be a primary driver of discontinuation in many patients. Loss of efficacy, is, however, not the only factor that will lead to discontinuation, with adverse events and other factors also contributing. Furthermore, in clinical practice, assessment of continued benefit is likely to be based on less formal criteria than continued response, informed by a combination of clinician judgement and patient experience, which may not fully align with formal response criteria. It is also important to note that the application of separate discontinuation rates for the Week 16 to 52 and post-52-week periods cannot be estimated properly when treatment sequences are evaluated. A simplified approach based upon applying a single discontinuation rate for both periods resolves this issue.

2.8 Issue 11: Differences in clinical outcomes based on region and skin type

Three questions were raised by the technical team: whether data from the Japanese patients recruited in the baricitinib trials are generalisable to clinical practice in the NHS; whether data from European patients only would be more clinically relevant than the full intent-to-treat population; and whether it reasonable to assume comparable efficacy for baricitinib in Black patients compared to White patients in the absence of robust data? If not, what adjustments should be made when assessing the effectiveness of treatment in Black patients.

2.8.1 Are data from the Japanese patients generalisable to clinical practice in the NHS

The company considers that patients recruited to the BREEZE-AD trials are representative of patients in UK clinical practice. However, they note that response variation in Japanese and non-Japanese patients should be interpreted with caution given the limited available data and potential differences in

clinical practice. The ERG agrees with the company that differences in baseline characteristics, such as disease severity and differences in treatment practices related to rescue topical corticosteroids (TCS) lead to differences in treatment efficacy in Japanese patients. The clinical expert at technical engagement also stated that eczema and clinical practice may differ in Japanese patients. Therefore, this is a source of uncertainty and the ERG considers that a scenario analysis based on European patients only may be more clinically relevant than the full intent-to-treat population.

The company conducted a scenario analysis using only European patients in the JAIN trial, which had a minimal impact on the cost-effectiveness results, compared with the revised company base case. Response rates specific to this population were utilised in this analysis. All other inputs and assumptions were in line with the revised company base case. However, these should be specific to the European only population to produce coherent and accurate results. Therefore, the cost-effectiveness results produced by the company are unreliable. Additionally, the results are less robust than the base-case analysis, given the smaller sample sizes informing the inputs.

2.8.2 Efficacy of baricitinib in Black patients

The company acknowledge the lack of data for Black patients from the BREEZE-AD trials, but note that the trial program was not designed to compare efficacy between different racial or ethnic groupings. The company further acknowledge that the pathology of AD and its underlying cytokine pathways may be distinct in Black patients, but note that the Appraisal Committee in the dupilumab appraisal (TA534) concluded there to be insufficient evidence to determine the extent to which differences in cytokine pathways could modify the treatment effect. The clinical expert at technical engagement also stated that the pattern of eczema is different in African patients. The ERG agrees that due to the lack of data, it is not possible to establish baricitinib efficacy in this population. However, the ERG maintains that this is a potential equalities issue.

2.9 Issue 12: Indirect treatment comparison heterogeneity

The issue of whether the patient population and censoring differences between the baricitinib and dupilumab trials are likely to bias the results of the ITC in favour of either treatment was raised by the technical team.

The company acknowledges that heterogeneity exists in the trials included within the ITC and that differences in secondary censoring rules between the dupilumab trials and baricitinib BREEZE-AD trials could bias the relative efficacy results in favour of dupilumab. The ERG agrees with the company that some trial heterogeneity is an expected limitation of an ITC and although this heterogeneity does not significantly reduce the validity of the data produced, this should be considered when interpreting the results.

2.10 Issues 13 and 14 Resource use

These two issues addressed concerns raised by the ERG regarding resource utilisation and costs. Specifically, the Technical team asked the following: i) Is the company's assumption of equivalence between baricitinib and dupilumab for flare control reasonable? ii) Are bathing products used in current NHS clinical practice? iii) Does a 50% reduction in concomitant treatments for responders to treatment reasonably reflect NHS clinical practice?

In response to Issue 13 (flare control), the company acknowledged the uncertainty surrounding the appropriate flare rates, but accepted the ERG and Technical Team's preference for setting baricitinib as equivalent to BSC, to better reflect the observed data.

Similarly, the company acknowledged the preference of the ERG and Technical Team to remove the cost of bathing products from the economic analysis and reflected this in their revised base-case. Stakeholder comments from Richard Weller reiterated this, stating costs of bathing products should be removed as most NHS trusts no longer pay for this, following trial evidence suggesting no benefit.

Concerning the resource utilisation rates, the company continued to assume a 50% reduction in their revised base case, aligning with assumptions made in TA534, as well as ERG and Technical Team preferences. A comment from Richard Weller suggests that this reflects NHS practice and that these assumptions are reasonable. The company also made several minor updates to resource utilisation rates and costs applied, to align with the ERG's preferences.

3 ERG ADDITIONAL ANALYSES

Due to significant delays in the ERG receiving an updated model and the limited time available, the additional work undertaken by the ERG does not constitute a formal critique of the company's resubmission and hence does not accord with the procedures and templates applied to the original submission. The ERG was not able to check the validity of the implementation of any proposed changes, but has ensured replication of the results presented by the company. Due to significant issues with the model and the above mentioned delays, the ERG has also been forced to implement the presented scenarios by manually updating input parameters. This approach substantially increases the risk of a user error.

3.1 Updated company base case

The company model submitted during the Technical Engagement step included a number of changes from the initial executable model provided by the company. A description of the changes can be seen below. The results of the updated base-case cost-effectiveness analysis are summarised in Table 2

along with an exploration of the impact of revised assumptions on the results of the economic analysis. The assumptions explored are detailed below, with results presented in Table 4.

i) Utility waning

As described in Issue 9, the company presented the updated results to include the assumption of utility waning. This scenario assumes that only a proportion of patients maintain the utility benefit associated with response, with differential rates applied to biologic treatments and BSC. The ERG does not consider these revised assumptions to be appropriate and therefore presents scenario analysis in which utility waning is removed.

ii) Discontinuation rates

As described in Issue 11, conditional response rates were used in the updated base case to model discontinuation between Week 16 and 52, with all-cause discontinuation used to model rates applied in the post 52-week period. The ERG prefers to use all-cause discontinuation rates in both periods and therefore explores a scenario applying all-cause discontinuation rates to both periods.

iii) Use of JAIN /JAIN-like and CAFE/CAFÉ-like patients to model BSC

Issue 4 describes how the revisions to the company base-case to use the EASI75 endpoint have been implemented using JAIN patients to model response rates to Placebo (BSC). The ERG prefers to use the placebo arm of both JAIN/JAIN-like and CAFÉ/CAFÉ-like patients to model placebo. In this analysis, response rates are updated using relative effects estimated from the ITC and applied to placebo patients in both JAIN/JAIN-like and CAFÉ/CAFÉ-like as described in the ERG report (Section 6.1).

iv) EASI50 plus ≥ 4 DLQI

As described in issue 4, the ERG considers there to be limitations to using the EASI75 endpoint and therefore a scenario is presented using the company's base-case assumption, but applying the EASI50 plus ≥ 4 DLQI endpoint. This scenario also updates the utilities to those used in TA534 as no face valid values are available from the BREEZE AD trials.

Table 2 Company base case results and impact of removal of the company's new assumptions

Analysis	Intervention	Discounted Costs	Discounted QALYS	Fully incremental ICER	Change from Base Case	Net Monetary Benefit	
						£20,000 WTP threshold	£30,000 WTP threshold
Company Base Case (revised at the TE stage)	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£27,037	-	-£1,972	£830
	Dupilumab	[redacted]	[redacted]	£113,459	-	-£69,676	-£59,629
ii) Removal of utility waning	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£123,090	+£96,053	-£6,347	-£5,731
	Dupilumab	[redacted]	[redacted]	£326,796	+£213,337	-£83,509	-£80,378
iii) Consistent Discontinuation rates, 16-52 weeks and 52+ weeks	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£26,285	-£752	-£2,016	£1,191
	Dupilumab	[redacted]	[redacted]	£112,107	-£1,351	-£79,132	-£67,553
iv) JAIN-like and CAFÉ-like used to model BSC	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£26,985	-£52	-£2,214	£956
	Dupilumab	[redacted]	[redacted]	£114,771	+£1,313	-£76,004	-£65,048
v) EASI50/DLQI endpoint*	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£29,262	+£2,225	-£2,002	£159
	Dupilumab	[redacted]	[redacted]	£90,083	-£23,376	-£119,339	-£100,435

BSC, best supportive care; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year; WTP, willingness to pay

*For this scenario utility values are also updated to those used in TA534.

3.2 ERG base case

Table 3 presents updated ERG base-case along with several exploratory analyses. The revised ERG base-case analysis retains the majority of the assumptions made in the original ERG base-case, with the only exception being that discontinuation rates are revised to reflect the newly available JAIN data. The ERG's base-case therefore rejects several revised assumptions made in the company's revised base-case including i) utility waning, ii) use of conditional response probabilities in the Week 16 to 52 period, and iii) use of the EASI75 endpoint. These alternative assumptions are however, explored in scenario analyses. Note, the ERG base case also includes treatment sequencing and therefore all scenarios are presented including treatment sequencing.

Table 3 ERG base case results and impact of inclusion of the company's new assumptions

Analysis	Intervention	Discounted Costs	Discounted QALYS	Fully incremental ICER	Change from Base Case	Net Monetary Benefit	
						£20,000 WTP threshold	£30,000 WTP threshold
Revised ERG base-case	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£70,825	-	-£4,680	-£3,759
	Baricitinib + Dupilumab	[redacted]	[redacted]	£173,912	-	-£133,604	-£124,306
	Dupilumab	[redacted]	[redacted]	Dominated	-	-£137,591	-£128,570
	Dupilumab + Baricitinib	[redacted]	[redacted]	£286,008	-	-£140,472	-£130,916
ii) Addition of utility waning	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£28,442	-£42,383	-£1,936	£357
	Baricitinib + Dupilumab	[redacted]	[redacted]	£82,186	-£91,725	-£112,162	-£92,144
	Dupilumab	[redacted]	[redacted]	Dominated	NA	-£118,094	-£99,324
	Dupilumab + Baricitinib	[redacted]	[redacted]	£308,453	+£22,445	-£119,068	-£98,810
iii) Changing 16-52 weeks discontinuation rates to conditional probabilities	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£77,298	+£6,472	-£4,383	-£3,618
	Baricitinib + Dupilumab	[redacted]	[redacted]	£174,072	+£161	-£131,833	-£122,796
	Dupilumab	[redacted]	[redacted]	Dominated	NA	-£133,473	-£124,748
	Dupilumab + Baricitinib	[redacted]	[redacted]	£253,037	-£32,970	-£136,330	-£127,100
iv) EASI75*	BSC	[redacted]	[redacted]	-	-	-	-
	Baricitinib	[redacted]	[redacted]	£112,724	£41,899	-£7,843	-£6,997
	Baricitinib + Dupilumab	[redacted]	[redacted]	£297,498	£123,586	-£102,236	-£97,989
	Dupilumab	[redacted]	[redacted]	Dominated	NA	-£104,046	-£100,248
	Dupilumab + Baricitinib	[redacted]	[redacted]	£551,663	+£265,655	-£109,745	-£105,356

BSC, best supportive care; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life year; WTP, willingness to pay

***** For this scenario utility values are also updated to those used in the company's revised base-case

4 REFERENCES

1. Haeck IM, Knol MJ, Ten Berge O, van Velsen SG, de Bruin-Weller MS, Bruijnzeel-Koomen CA. Enteric-coated mycophenolate sodium versus cyclosporin A as long-term treatment in adult patients with severe atopic dermatitis: a randomized controlled trial. *J Am Acad Dermatol* 2011;**64**:1074-84.

2. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. *NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE* Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2016.

3. National Institute for Health and Care Excellence (NICE). *Dupilumab for treating moderate to severe atopic dermatitis: Technology appraisal guidance [TA534]*. 2018. Available from: https://www.nice.org.uk/guidance/ta534 [accessed 28th January 2020].

4. Newsom M, Bashyam AM, Balogh EA, Feldman SR, Strowd LC. New and emerging systemic treatments for atopic dermatitis. *Drugs* 2020;**80**:1041-52.

5. National Institute for Health and Care Excellence. *Single Technology Appraisal. Tofacitinib for treating active psoriatic arthritis after inadequate response to DMARDs [ID1220]*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta543/documents/committee-papers [accessed 23rd November 2020].

6. National Institute for Health and Care Excellence. *Single Technology Appraisal. Tildrakizumab for treating chronic plaque psoriasis after systemic therapy [ID1060]. Committee Papers* NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta575/documents/committee-papers [accessed 23rd November 2020].

7. National Institute for Health and Care Excellence. *Single Technology Appraisal. Risankizumab for treating moderate to severe plaque psoriasis [ID1398]. Committee Papers*. NICE; 2019. Available from: https://www.nice.org.uk/guidance/ta596/documents/committee-papers [accessed 23rd November 2020].

8. National Institute for Health and Care Excellence. *Single Technology Appraisal. Ixekizumab for treating active psoriatic arthritis after DMARDs [ID1194]. Committee Papers*. NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta537/documents/committee-papers [accessed 23rd November 2020].

9. National Institute for Health and Care Excellence. *Single Technology Appraisal. Dimethyl fumarate for treating moderate to severe plaque psoriasis [ID776]. Committee Papers*. NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta475/documents/committee-papers [accessed 23rd November 2020].

10. National Institute for Health and Care Excellence. *Single Technology Appraisal. Certolizumab pegol for treating chronic plaque psoriasis [ID1232]. Committee papers*. NICE; 2019. Available from: https://www.nice.org.uk/guidance/ta574/evidence [accessed 23rd November 2020 2020].

11. National Institute for Health and Care Excellence. *Single Technology Appraisal. Brodalumab for treating moderate to severe plaque psoriasis [ID878]. Committee Papers*. NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta511/history [accessed 23rd November 2020].

12. National Institute for Health and Care Excellence. *Single Technology Appraisal. Apremilast for treating active psoriatic arthritis (rapid review TA372) [ID1017]. Committee Papers*. NICE; 2016. Available from: https://www.nice.org.uk/guidance/ta433/documents/committee-papers [accessed 23rd November 2020].