Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Baricitinib for treating severe alopecia areata [ID3979]

Contents:

The following documents are made available to stakeholders:

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

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

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. Alopecia UK
   b. British Association of Dermatology

4. **External Assessment Report** prepared by BMJ-TAG

5. **External Assessment Report – factual accuracy check**

6. **Technical engagement response from company**

7. **Technical engagement responses and statements from experts:**
   a. Dr Abby MacBeth, consultant dermatologist – clinical expert, nominated by British Association of Dermatology
   b. Dr Matthew Harries, consultant dermatologist – clinical expert nominated by British Association of Dermatology
   c. Lynn Wilks, trustee and volunteer – patient expert, nominated by Alopecia UK
   d. Sue Schilling, chief executive officer – patient expert, nominated Alopecia UK

8. **Technical engagement responses from stakeholders:**
   a. British Association of Dermatology

9. **External Assessment Report critique of company response to technical engagement** prepared by BMJ-TAG

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 treating severe alopecia areata [ID3979]

Document B

Company evidence submission

August 2022

File name	Version	Contains confidential information	Date
ID3979_Baricitinib in AA_Document B [ACIC REDACTED]_04August2022	FINAL	Yes	4th August 2022

Contents

Tables and figures

Abbreviations

AA	Alopecia areata
AD	Anxiety disorder
ADR	Adverse drug reactions
AE	Adverse event
AESI	Adverse events of special interest
AGA	Androgenic alopecia
ANCOVA	Analysis of covariance
AT	Alopecia totalis
ATE	Arterial thromboembolism
ATP	Adenosine triphosphate
AU	Alopecia universalis
BAD	British Association of Dermatologist
BARI	Baricitinib
BID	Twice per day
BMI	Body mass index
BSC	Best supportive care
CD8	Cluster of differentiation 8
ClinRO	Clinician reported outcome
CPK	Creatine phosphokinase
CRD	Centre for Reviews and Dissemination
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DE	Depressive episodes
DLQI	Dermatology Life Quality Index
DPCP	2,3-diphenylcyclopropenone
DSA	Deterministic sensitivity analysis
DSP	Disease Specific Programme
EB	Eyebrow
EMA	European Medicines Agency
EQ-5D	European Quality of Life-5 Dimensions
FAS	Full analysis set
FDA	Food and Drug Administration
GBP	Great British pound
GCP	Good clinical practice
GP	General practitioner
HADS	Hospital Anxiety Depression Scale
HDL-C	High-density lipoprotein-cholesterol
HI	High intensity
HMG CoA	3-hydroxy-3-methylglutaryl coenzyme A
HRQoL	Health-related quality of life
HSE	Health Survey for England

HSUV	Health state utility values
ICER	Incremental cost-effectiveness ratio
IFN	Interferon
IGA	Investigator global assessment
IL-15	Interleukin-15
IMT	Immunotherapy
IPD	Individual patient data
IS	Immunosuppressant
ITT	Intention-to-treat
IV	Intravenous
JAK	Janus kinase
KOL	Key opinion leader
LDL-C	Low-density lipoprotein-cholesterol
LSM	Least squares mean
LYG	Life-years gained
MACE	Major adverse cardiovascular events
MCS	Mental Component Summary
MHC	Major histocompatibility complex
MHRA	Medicines and Healthcare products Regulatory Agency
mLOCF	Modified last observation carried forward
NHB	Net health benefit
NHS	National Health Service
NICE	National Institute for Health and Care Excellent
NIHR	National Institute for Health Research
NKG2DL	Natural killer group 2D ligand
NMB	Net monetary benefit
NMSC	Nonmelanoma skin cancer
NRI	Non-responder imputation
NRS	Numeric rating scale
ONS	Office for National Statistics
OWSA	One-way sensitivity analysis
PAS	Patient Access Scheme
PBO	Placebo
PCFB	Percent change from Baseline
PCS	Physical Component Summary
PPS	Per-Protocol Set
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTFU	Post-trial follow up
PUVA	Psoralen plus ultraviolet A light therapy
PYE	Patient-year exposure

PYR	Patient-years at risk
QALY	Quality-adjusted life year
QD	Once daily
RCT	Randomised controlled trial
RDD	Recurrent depressive disorder
RWE	Real-world evidence
SADBE	Squaric acid dibutylester
SAE	Serious adverse event
SALT	Severity of Alopecia Tool
SD	Standard deviation
SE	Standard error
SF-36	Medical outcomes study 36-item short form health survey
SH	Scalp hair
SLR	Systematic literature review
SOC	System organ class
STAT	Signal transducers and activators of transcription
TCR	T cell receptor
TCS	Topical corticosteroids
TE	Treatment emergent
TEAE	Treatment emergent adverse event
TGF	Transforming growth factor
TM	Trademark
UK	United Kingdom
UK	United Kingdom
ULN	Upper limit of normal
US	United States
VAS	Visual analogue scale
VTE	Venous thromboembolism
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 evaluation. Any differences between the decision problem addressed with 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 adults with severe alopecia areata (AA). The indication of relevance for this submission covers the full marketing authorisation for baricitinib.

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 severe alopecia areata.	Adults with severe alopecia areata.	N/A- in line with the NICE final scope.
Intervention	Baricitinib	Baricitinib	N/A- in line with the NICE final scope.
Comparator(s)	Established clinical management without baricitinib	Established clinical management without baricitinib, which may include supportive care	The current clinical management for severe AA predominantly relies on using off-label treatments that are not supported by robust evidence and are associated with suboptimal efficacy and poor tolerability. These include topical, intralesional or oral corticosteroids, topical immunotherapy such as 2,3-diphenylcyclopropenone (DPCP), immunosuppressive drugs such as methotrexate, and psoralen plus ultraviolet A light therapy (PUVA). Some of these treatments can be burdensome for patients, or are limited to short-term use due to side effects.[1, 2] As such, leaving AA untreated and taking a 'watch and wait' approach is considered by experts as a legitimate option for many patients with AA. However, unlike patients with mild disease, patients with severe AA rarely experience spontaneous remission, at which point they must instead rely on best supportive care options such as psychological support and cosmetic concealment of hair loss in order to cope with the psychological burden of the disease.
Outcomes	The outcome measures to be considered include: • Disease severity e.g. Severity of Alopecia Tool (SALT)	The outcome measures to be considered include: • Measures of disease severity and improvement in hair loss (including	N/A- in line with the NICE final scope.

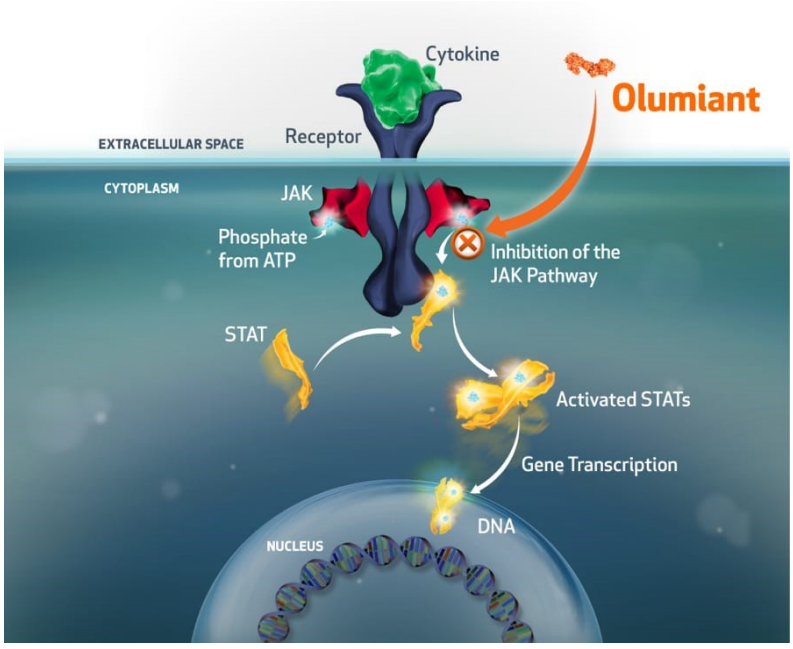
	• Improvement in hair loss e.g. Scalp Hair Assessment Score, Measure for Eyebrow Hair Loss, Measure for Eyelash Hair Loss • Adverse effects of treatment • Health-related quality of life	SALT, ClinRO for eyebrow hair loss and eyelash hair loss, PRO measures for scalp hair assessment, PRO measures for eyelashes and eyebrows) • Adverse effects of treatment (including AEs, SAEs, AESIs) • Health-related quality of life (including EQ-5D, Skindex-16 AA, HADS and SF-36)	
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 NICE final scope	N/A- in line with the NICE final scope.
Special considerations including issues related to equity or equality	None identified.	None identified.	N/A- in line with the NICE final scope.

Abbreviations: AA: alopecia areata; AE: adverse event; AESI: adverse event of special interest; ClinRO: clinician reported outcome; DPCP: 2,3-diphenylcyclopropenone; EQ-5D: the European Quality of Life-5 Dimensions; HADS: Hospital Anxiety Depression Scale; NRS: numeric rating scale; PRO: patient reported outcome; PUVA: psoralen plus ultraviolet A light therapy; SAE: serious adverse event; SALT: Severity of Alopecia Tool; SF-36: Short Form 36 Health Survey Questionnaire.

B.1.2 Description of the technology being evaluated

A summary of the mechanism of action, marketing authorisation status, costs and administration requirements of baricitinib in the treatment of severe AA is presented in Table 2. The draft Summary of Product Characteristics (SmPC) is located in Appendix C.

Table 2: Technology being evaluated

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 Janus kinase (JAK) 1 and JAK2. These enzymes mediate pathways involved in the underlying immunopathophysiology of AA.[3, 4] The JAK signaling pathway mediates cellular responses to numerous different cytokines via a cascade of activation. This process is initiated when a cytokine binds to its target cell surface receptor, which induces a conformational change in the cytoplasmic portion of the receptor. Downstream, this leads to phosphorylation of the receptor-associated JAK and the subsequent phosphorylation and activation of signal transducers and activators of transcription (STATs). STATs then translocate to the nucleus and mediate target gene regulation.[3] The proinflammatory cytokines interferon-γ and interleukin-15 (IL-15), which signal via the JAK-STAT pathway, are considered to be the key mediators of the hair loss observed in AA. Therefore, by inhibiting JAK1/JAK2, baricitinib interrupts the underlying immunopathogenesis of AA, thereby reversing the hair loss that is characteristic of the disease.[4] **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[5]

Marketing authorisation/CE mark status	Marketing authorisation for baricitinib in AA from the Medicines and Healthcare products Regulatory Agency (MHRA) is expected in [redacted].
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated marketing authorisation for baricitinib in this indication is for the treatment of adults with severe AA. Baricitinib is also currently indicated for the treatment of: • moderate to severe active rheumatoid arthritis in adult patients who have responded inadequately to, or who are intolerant to one or more disease-modifying anti-rheumatic drugs. Baricitinib may be used as monotherapy or in combination with methotrexate.[6] • moderate to severe atopic dermatitis in adult patients who are candidates for systemic therapy.[7] **Contraindications:** • 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. The recommended dose for AA patients is 4 mg once daily. A dose of 2 mg once daily may be appropriate for some patients, such as those aged 75 years or older or those patients with a history of chronic or recurrent infections. A dose of 2 mg once daily may also be considered for patients who have achieved sustained control of disease activity with 4 mg once daily and are eligible for dose tapering.
Additional tests or investigations	No additional tests or investigations are required to identify patients that are eligible for baricitinib. However, as is common with many immunomodulating medications prescribed on the NHS, some standard screening tests are recommended prior to initiation of baricitinib, such as tuberculosis and various laboratory measures, as described in the SmPC[8]
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.[9] Assuming that patients take one tablet per day for one year, the annual cost of a baricitinib treatment course is £10,508.24.
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 4mg baricitinib treatment course is [redacted]

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

Overview and burden of the disease

- Alopecia areata (AA) is a chronic autoimmune disorder affecting the hair follicles, characterised by the sudden onset of non-scarring hair loss.[10, 11]
- AA evolution is unpredictable, and the prognosis varies considerably depending on disease severity and duration. Patients with extensive hair loss rarely experience spontaneous regrowth of hair and tend to have a poor prognosis even on treatment.[2, 12]
- AA affects males and females equally, and can occur at any age, though there is generally a higher incidence in younger adults.[10]
- The burden of AA on patients impacts many aspects of their lives and AA is associated with a significant impact on HRQoL. AA is further associated with a variety of comorbid diseases, including psychiatric disorders and other immune diseases.[2, 13]

Clinical pathway of care and unmet need

- There are currently few evidence-based treatments for AA. Current clinical management for AA therefore often involves using off-label treatments, of which few have been well-evaluated in clinical trials or have consensus on their efficacy. Many treatments are also associated with unpleasant or uncomfortable side effects.[1]
- Leaving AA untreated and taking a 'watch and wait' approach is considered by some experts as a legitimate option for many patients with AA. A substantial proportion of patients with limited patchy hair loss of short duration experience spontaneous remission, meaning that the use of off-label treatments, with their associated side effects, may not be justified.[1, 14, 15] A 'watch and wait' approach may also be taken initially for patients with severe disease; however, unlike those with a milder form of the disease, this subgroup of patients often has a poor prognosis and rarely experiences spontaneous remission.[1, 16]
- Aside from a 'watch and wait' approach, various off-label treatment options may be trialled in patients with AA in an attempt to restore hair regrowth. These treatments include topical, intralesional or oral corticosteroids, topical immunotherapy such as DPCP, immunosuppressive drugs such as methotrexate, psoralen plus ultraviolet A light therapy (PUVA), minoxidil and calcineurin inhibitors.[1] These treatments are not supported by robust evidence from clinical trials, and are associated with suboptimal efficacy and many have a safety profile that limits long-term use.
- There is a substantial unmet need for evidence-based, effective, and well-tolerated medications for the treatment of severe AA.
- A positive recommendation for baricitinib in patients with severe AA would allow these patients to benefit from improved outcomes compared with current management, and would provide a novel, evidence-based, therapeutic option for AA.

B.1.3.1 Overview of the disease

Alopecia areata is a chronic autoimmune disease that can lead to significant, and in some cases, total, hair loss on the scalp, face and or body

Alopecia areata (AA) is a chronic autoimmune disorder affecting the hair follicles, characterised by the sudden onset of non-scarring hair loss.[10, 11] Although AA can affect any hair-bearing skin including the beard, eyebrows, eyelashes, body and limbs, the scalp is most commonly affected, with hair loss on the scalp being observed in approximately 90% of cases.[17] The clinical presentation of AA is heterogenous, with the extent of hair loss ranging from well-defined patches on the scalp to extensive or total hair loss on the scalp, face, and/or body.[2]

Disease classification is usually by extent or pattern of hair loss.[17] Patchy AA is the most common presentation of AA, which is characterised by round or oval patches of hair loss, and

can encompass both patients with mild and severe AA. Other types of AA include alopecia totalis (AT), referring to total or near-total hair loss on the scalp, and alopecia universalis (AU), referring to total or near-total loss of body hair. Different patterns of hair loss have also been described, such as diffuse AA where sudden thinning of hair all over the scalp is observed, rather than in patches (Table 3).[12]

Table 3. Types of alopecia areata

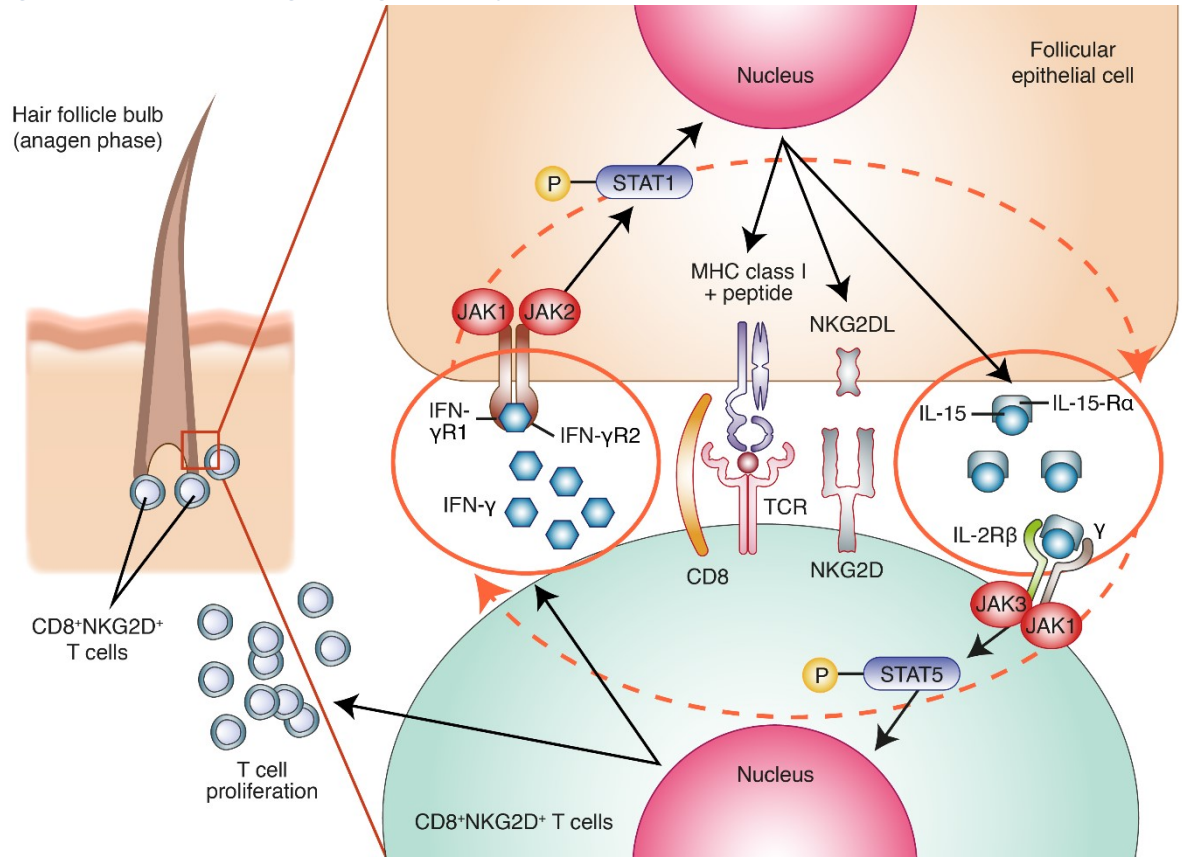
Classification	Presentation
Patchy AA	Single or multiple well-defined patches of scalp hair loss
Alopecia totalis	Total or near-total scalp hair loss
Alopecia universalis	Total or near-total loss of all body hair
Ophiasis	Hair loss on the occipital and temporal scalp site
Ophiasis inversus (sisapho)	Central hair loss, sparing lateral and posterior scalp sites
Diffuse/AA incognita	Diffuse hair loss and reduction of hair density
Alopecia barbae	Discrete circular or patchy hair loss in the moustache or beard

Abbreviations: AA: alopecia areata.
Source: Lintzeri, et al. 2022.[12]

Alopecia areata is an autoimmune disease directed against the hair follicle

While the exact cause of AA is unknown, it is understood to be a multifactorial disease driven by genetic, epigenetic and environmental factors that contribute to an immune-mediated attack on hair follicles.[18] Research suggests that the pathogenesis of AA involves an interaction between lymphocytes and hair follicular cells mediated by the JAK-STAT pathway. CD8+NKG2D+ T-cells produce IFN-γ which then binds to receptors on the surface of hair follicle cells and signals via JAK1 and JAK2 to stimulate the production of interleukin [IL]-15 in the hair follicle cell (Figure 2). IL-15 is released from the hair follicle and binds to its receptor on the surface of T-cells, further stimulating the production of IFN-γ via JAK1 and JAK3 signalling and creating a positive feedback loop between the follicular cell and the CD8+NKGD2+ T-cells.[12, 19, 20] As a result, hair follicles convert prematurely from the growth (anagen) phase into the loss (telogen) phase, resulting in the hair loss that is characteristic of AA.[21]

Figure 2. JAK-STAT signalling pathway in alopecia areata



Abbreviations: MHC: major histocompatibility complex; TGFβ1: transforming growth factor beta 1; TCR: T cell receptor; CD8: cluster of differentiation 8; NKG2DL: natural killer group 2D ligand; NKG2D: natural killer group 2D; IL-15: interleukin-15; JAK: Janus kinase; IL-15RA: interleukin-15 alpha subunit; STAT: signal transducer of activation; P: phosphate; IFNγ: interferon gamma; IFNγR: interferon gamma receptor.
Adapted from: Divito SJ, et al. 2014

Diagnosis of AA is typically determined by clinical history and physical examination and can be differentiated from other hair loss disorders

AA is typically diagnosed based on presenting features and once other causes of hair loss have been excluded.[1, 22] Aside from the patches of hair loss that are observed in AA, typical clinical features of AA that support a diagnosis include the presence of so-called “exclamation mark” hairs, that are short, broken and taper proximally.[21, 22] In addition, the pull test may be used to support an AA diagnosis, which involves the examiner grasping approximately 40–60 hairs between their thumb, index, and middle fingers and gently pulling them away from the scalp. A positive result is achieved if >10% of the hairs are pulled out, indicating hair shedding, though a negative pull test does not always rule out an AA diagnosis.[23]

While further testing is often not required beyond careful evaluation of the patients’ clinical history and thorough physical examination, additional investigations such as a trichoscopy or histopathology may sometimes be used to confirm the diagnosis. This is particularly useful for certain forms of AA, including diffuse AA.[2, 24]

The Severity of Alopecia Tool provides an objective and standardised method to estimate the extent of scalp hair loss and describe disease severity

Determination of the severity of AA is important for optimal disease management since it informs therapeutic decision-making and can aid evaluation of clinical response and prognosis.[12] Experts are generally in agreement that the definition of AA severity should be driven by the extent of scalp hair loss, as patients with AA often report scalp hair loss as being the most bothersome symptom of AA.[25] As such, it has been suggested that the Severity of Alopecia Tool (SALT) should be used to describe disease severity, as it can provide an objective estimation of scalp hair loss in patients with AA.[1, 21] The SALT score is calculated by measuring the hair loss in each of the four areas of the scalp (left side, right side, top and back) following a physical examination or using a photograph. Within each area, the percentage of hair loss is determined independently using a visual aid, before adding the total of each area to capture the overall percentage of hair loss.[26] An absolute score of 0% indicates no hair loss, while a score of 100% indicates complete hair loss on the scalp. Given that the SALT score measures the extent of hair loss irrespective of the underlying cause, the SALT score does not assess hair loss specific to AA. In addition, the SALT score does not consider other anatomical sites beyond the scalp (e.g., beard hair), and also does not consider elements beyond hair loss, such as the impact of AA on health-related quality of life (HRQoL).[12] While further scores have been proposed to address these limitations, there is currently no consensus on their use in clinical trials or practice.[12, 27] The SALT therefore represents the only validated and standardised tool for measuring the extent of hair loss in patients with AA for which there is consensus on its use in studies and clinically.[28]

In clinical practice in the UK, SALT scores may be used by dermatologists to determine the extent of hair loss and thus, the most appropriate management option. A SALT score of ≥50 (indicating 50% or more scalp hair loss) has been used consistently in the literature and clinical practice guidelines to define severe disease, encompassing patients with extensive patchy AA, alopecia totalis and alopecia universalis.[29] Some dermatologists may combine this with an assessment of psychological burden using a HRQoL measure to give a comprehensive picture of disease severity of an individual patient.[21] However, this is not currently incorporated into clinical practice guidelines as a way in which to determine disease severity.[1, 21]

In clinical trials, SALT is also used for the assessment of scalp hair loss, and a SALT score ≥50 has also been used in this setting to define severe AA. SALT scores may also be used in clinical trials to define the efficacy of a treatment, with improvements from baseline SALT scores, as well as the proportion of patients reaching pre-defined absolute SALT scores, being used as outcomes to measure treatment success.[21, 28] For instance, in the two pivotal trials that make up the current baricitinib clinical programme for AA, BRAVE-AA1 (NCT03570749) and BRAVE-AA2 (NCT03899259), the primary efficacy endpoint is the proportion of patients achieving SALT ≤20 at week 36 (≤20% scalp hair loss, or ≥80% scalp coverage with hair).[30, 31]

The disease course of AA is variable, and patients with severe disease are more likely to have a poor prognosis

The disease course of AA is variable and often unpredictable.[1] Consistent with the autoimmune foundations of the disease, patients with AA typically experience a relapsing and remitting disease course, with as many as 85% of individuals experiencing multiple episodes of hair loss.[17] In some individuals, hair loss occurs at intervals separated by episodes of regrowth and prolonged periods of remission, while others experience areas of hair loss resolving at the same time as new patches are appearing. For other patients, there is a more persistent disease course

without any regrowth or further deterioration. It is estimated that between 14–25% of patients with patchy AA progress to more severe forms of AA, including alopecia totalis or alopecia universalis.[2]

Since hair follicles are not destroyed by the disease process, some patients experience spontaneous hair regrowth.[2] Early in the disease course, spontaneous hair regrowth is relatively common, with as many as 80% of patients with mild patchy AA experiencing spontaneous regrowth within 1 year.[14] However, relapses are frequent and the likelihood of recovery decreases with increasing AA severity.[2, 32] Clinical trial data indicate that when hair loss becomes extensive, it tends to be chronic and spontaneous regrowth is rare, with very few patients with severe AA experiencing full recovery (<10%).[1, 16] Moreover, patients with severe AA are more likely to have further hair loss over time.[15, 24] The main prognostic factor in patients with AA is therefore considered to be the extent of hair loss, especially at presentation.[12, 33] As such, patients with more severe AA, such as alopecia totalis or alopecia universalis, generally have a poor prognosis.[1, 16, 33] In a review of studies that evaluated recovery rates in patients with alopecia totalis or alopecia universalis with a mean follow-up of ≥5 years, it was reported that only 8.5% (32/375) of patients achieved complete recovery as an endpoint, either with or without treatment.[34] Similarly, an Italian study including 191 patients with AA reported that of those study participants with hair loss of <25% initially, 68% were disease free at follow-up compared with 32% and 8% of those who had 25–50% or >50% hair loss at presentation, respectively.[1, 15]

Aside from the extent of hair loss, other poor prognostic factors include those who have the ophiasis variant or have coexisting changes in finger- and toenails. Compared with adults, early onset of AA in children often results in both a greater degree of and progression of AA. A history of atopy and concomitant autoimmune diseases also confers poor outcomes.[1]

AA is more common in younger adults and affects men and women equally, with severe AA representing a smaller subset of all AA patients

Estimates of the general population prevalence of AA range from 0.1–0.58%.[35] Similarly, a recent UK-based population-based cohort including 4.16 million subjects in primary care reported an overall incidence rate of 0.26 per 1000 person-years and a point prevalence of 0.58% in adults in 2018.[36] Of these, clinical experts have estimated that 15–30% have a severe form of the disease, meaning that severe AA affects a smaller, but not insignificant, subset of the total number of AA patients.[37]

AA affects both males and females and can occur at any age; however, AA has been shown to be more prevalent in younger age groups. A recent UK study, which represents the largest population-based study of AA to date, found that the onset of AA peaked at age 25–29 years for both sexes.[36] Similarly, 88% of patients affected by AA are affected before the age of 40.[10] This is considerably younger than would be expected for normal age-related hair loss and is also during the age range that is arguably the most productive with respect to career, social life, and relationships, potentially exacerbating the impact of AA on this population.[10]

B.1.3.2 Burden of alopecia areata

The burden of AA on patients impacts many aspects of their lives

Patients with AA do not typically have physical symptoms that accompany their hair loss, although occasionally itching, tingling, burning or pain may occur prior to hair loss or with disease

activity.[22] However, hair has a range of important physiological functions, including acting as a physical barrier to UV rays and protecting the eyes from debris. Therefore, some patients with AA experience a range of physical challenges that may impact quality of life, including excessive sneezing and/or a runny nose due to nasal hair loss, as well as sunburn from the lack of bodily hair.[38] Loss of the protective function of eyebrows and eyelashes also presents physical challenges that are not easily compensated for, due to water, sweat, and other debris getting into the eyes.[38]

Beyond its physical functions, hair also has psychosocial importance in society, as it plays a role in both sexual and social communication. It is therefore well recognised that AA can have a profound psychological and psychosocial impact on individuals.[2, 39, 40] For instance, healthy hair is generally considered an indication of youth and vigour, and plays a critical role in identity and self-worth.[25, 41] In addition, certain cultures consider hair sacred, with some religions viewing uncut hair as a sign of devotion.[41] These strong cultural meanings behind hair have been highlighted by AA patients as a factor that contributes to the significant psychological impact of AA. For example, women expressed seeing hair as a mark of their femininity, and therefore likened losing their hair to having a mastectomy.[42] The hair loss in AA can therefore negatively impact self-esteem, body image and confidence, leading to feelings of trauma, shock and loss of identify.[11, 42] People with AA also often report feeling judged and may experience stigma due to their hair loss; one study reported that the general population associated those with severe hair loss as being sick, unattractive and even dirty.[43] Patients with AA therefore often report the highly visible manifestations of AA as the most troublesome aspect of AA and the primary cause of their distress.[38, 40]

Given the feeling of social judgment and stigma experienced by people with AA, attempts at concealment are common, including the use of wigs, fake eyelashes and eyebrows. However, using these methods of concealment can produce feelings of inauthenticity, shame, and anxiety.[42] Furthermore, some patients report negative experiences associated with wig-wearing, including itchiness, discomfort, the fear of it being dislodged, and worry about the reactions of others with respect to them wearing a wig.[11]

The unpredictable disease course is another aspect of AA that can be very stressful for patients, and can contribute to the psychological burden of the disease.[44] In a qualitative study, patients reported that the unpredictable and often rapidly alternating cycles of hair loss and regrowth are particularly disturbing aspects of the disease.[42] Some patients may find it particularly difficult to cope with relapse, or if new bald areas occur as others improve.[1, 45] In addition, patients have indicated that the pattern of hair loss, with randomly distributed visible patches of hair loss, is particularly challenging. The poor prognosis associated with the disease may also be particularly difficult for patients to cope with and can lead to feelings of hopelessness.

AA is associated with a significantly reduced HRQoL

Given the psychological and emotional distress caused by the sometimes extensive hair loss, many patients with AA experience a reduced health-related quality of life (HRQoL).[46-49] Various different instruments have been applied to AA to measure its impact on HRQoL, including generic measures such as SF-36.[2] For instance, a meta-analysis of studies employing the SF-36 instrument reported that patients with AA experience a significant impairment in HRQoL compared to age- and sex-matched controls. HRQoL impairment was most pronounced across the role-emotional, mental health and vitality domains ($P<0.001$), indicating poor social functioning, higher psychological distress, and diminished energy levels.[47]

Given the mono-symptomatic nature of the hair loss in AA and the fact many patients with AA are otherwise healthy, other more dermatology-specific measures are also widely used in order to better define the factors affected HRQoL, such as the Skindex-16 and the Dermatology Life Quality Index (DLQI).[2, 16] Using these instruments and others, studies consistently demonstrate the poor HRQoL experienced by patients with AA, which can be similar to those reported in patients with other chronic skin disease, including atopic dermatosis and psoriasis.[46-49] For instance, a systematic literature review (SLR) investigating the effects of AA on HRQoL showed that people with AA had consistently low HRQoL scores. These scores, measured using a range of different instruments (DLQI, Skindex, SF-36), were comparable to other chronic skin diseases, despite the lack of physical symptoms beyond hair loss that are known to reduce HRQoL, such as itching and sleep disturbance.[48] Similarly, an observational cross-sectional study conducted in a large cohort of patients with a range of dermatological conditions (n=4010) across 13 European countries demonstrated that EQ-5D VAS scores were comparable in people with AA (n=33, mean: 69.7%; SD:18.1) and those with AD (n=177, mean: 66.0%; SD: 19.0]) or psoriasis (n=682, mean: 65.6%; SD: 20.0) and lower than healthy controls (n=1359, mean: 82.2%; SD: 15.5). The authors also noted that the EQ-5D VAS scores of <70 associated with several dermatological conditions, including AA, were comparable to published health state estimates for chronic diseases such as diabetes mellitus, cardiovascular disease, anxiety, cancers, and liver disease.[50]

Several studies also demonstrate the greater quality of life impairment in patients with more severe disease.[47-49, 51] An SLR of 21 studies conducted by Rencz *et al.* (2016) identified that greater scalp involvement, disease recurrence and longer treatment duration all negatively impacted HRQoL.[47] Similarly, in a recent Europe-wide study, patients with AA had worse HRQoL (measured by DLQI) compared to age-, sex- and comorbidity-matched controls and patients with androgenic alopecia (AGA) (p=0.022), with greater DLQI score impairment in those with the greatest AA severity.[39]

AA is associated with psychological burden and a variety of comorbid diseases

In addition to the quality of life impairment reported in patients with AA, psychological stress levels, frequency of psychiatric disease and levels of psychiatric symptoms are typically higher in adult patients with AA than controls.[2, 11, 13] As such, it is estimated that there is a 66%–74% lifetime prevalence of psychiatric disorders in AA patients, with a 39% lifetime prevalence of depression and a 39%–62% prevalence of generalised anxiety disease.[10] In the UK, a study including 338 patients with alopecia (mixed-severity AA, n=279) reported clinically significant levels of social anxiety (37.5%), anxiety (35.5%) and depression (29%).[11] Similarly, in a population-based study based in UK primary care, adults with newly diagnosed mixed-severity AA (n=5,435) had significantly higher background prevalence of depressive episodes (DE), recurrent depressive disorder (RDD) and anxiety disorder (AD) (DE 19.4%, RDD 12.3%, AD 16.6%) compared to matched controls (n=21,740; DE 14.7%, RDD 8.6%, AD 12.9%). This study also found that those with AA were more likely to go on to develop new-onset DE and AD (adjusted hazard ratio [95% CI] 1.38 [1.13, 1.69]), RDD (1.30 [1.04, 1.62]), and AD (1.33 [1.09, 1.63]). Higher rates of antidepressant prescribing were also seen in people with AA.[13] These conditions may be primary disorders that manifest themselves as medical problems associated with AA, or could result from the chronic, relapsing nature of AA and its negative effect on a person's appearance.[39] Similarly, during an observational study among patients with a range of dermatological conditions (n=4010), AA was associated with significantly worse anxiety and depression compared with controls (OR [95% CI] 4.19 [2.0, 8.9]; p<0.05), prompting the authors to suggest there may be a need for psychiatric support in such individuals.[50] Another observational cross-sectional study in 17 European countries found that patients with

dermatological conditions, including alopecia, had a more than eleven-fold increased chance of having symptoms of body dysmorphic disorder compared with controls, a common psychiatric disorder associated with high costs for healthcare systems.[51] In severe AA patients with a longer disease duration, rates of depression and anxiety may be even greater than estimates in mixed-severity populations, given that greater disease severity is associated with greater impairment in quality of life.[47-49]

There have also been reports of suicidal ideation in people with AA, translating into an elevated mortality risk associated with intentional self-harm/psychiatric disease among patients with AA compared with control subjects, with the risk being particularly elevated among individuals with alopecia totalis (AT) or alopecia universalis (AU).[52, 53]

Besides psychiatric disorders, it has been reported that AA is associated with a variety of other comorbidities, including autoimmune or atopic diseases, which can also contribute to the overall burden of the disease (Table 4).[12, 17, 27] Other comorbidities that appear to be more prevalent among AA patients include vitamin D deficiency, iron-deficiency anaemia, metabolic syndrome, and infection with *Helicobacter pylori.*[12, 27]

Table 4. Comorbid diseases commonly reported in patients with AA.

Autoimmune diseases	Atopic diseases	Psychiatric disorders	Other
• Autoimmune thyroid disease • Vitiligo • Lupus erythematosus • Rheumatoid arthritis • Psoriasis	• Atopic dermatitis • Allergic rhinitis • Allergic conjunctivitis • Asthma	• Depression • Anxiety	• Vitamin D deficiency • Iron-deficiency anaemia • Metabolic syndrome • Helicobacter pylori infection

Abbreviations: AA: alopecia areata.
Source: Lintzeri, et al. 2022; Lee, et al. 2019a.[12, 27]

AA negatively impacts employment and relationships

The impact of AA on working life has been explored in several studies, which have shown that people with AA are treated differently by others because of their disease, or that they have limited their professional lives due to their condition.[40, 43] In a UK population-based study by Macbeth *et al.* (2022) that included 5435 people with AA and 21,740 matched controls, certificates for time off work were issued more frequently significantly ($p<0.001$) to people with AA (13.0% within a year of diagnosis) compared to matched controls (7.9%). Similarly, people with AA were more likely to have a record of unemployment in the year after diagnosis (1.3% of AA cases vs 0.6% of matched controls).[13]

The negative effect of AA on relationships with friends, family, or a romantic partner has also been well documented. Patients with AA may feel that they are unable to have a romantic relationship due to their hair loss, or may experience the end of a romantic relationship because their partner was unable to cope with the hair loss.[42] Further to this, people with AA may also withdraw from social situations frequently; a survey of individuals with mixed-severity AA found that avoiding social activities (62%) and reducing interactions with friends (54%) were both

common after a first episode of AA, highlighting the impact of AA on all aspects of an individual's social life.[54]

B.1.3.3 Clinical pathway of care

Current clinical management of AA usually relies on the use of off-label treatments that are not supported by robust evidence

Following on from recent approvals from the European Medicines Agency (EMA) and Food and Drug Administration (FDA),[55, 56] baricitinib will be the first MHRA-approved therapy specific to adult patients with severe AA, and the only licensed therapy for any form of AA other than intralesional corticosteroids. The current clinical management for AA in the UK therefore mostly relies on using off-label treatments; however, these options are not supported by robust evidence and their reported efficacies are low. Clinical practice guidelines specific to the UK include the NICE Clinical Knowledge Summary for AA, and the British Association of Dermatologists' (BAD) guidelines for the management of alopecia areata, published in 2012.[1, 29] However, these are limited in their treatment recommendations and highlight a lack of high-quality evidence for treatments in AA, with the majority of evidence based on studies with small sample size and short follow-up times.

Initial management of AA in the UK is typically based on the severity of hair loss on the scalp and the priorities of AA treatment generally include to:[29]

- Arrest the progression of hair loss and induce hair growth
- Improve patients' HRQoL
- Limit the adverse events related to therapy

Treatment of mild AA

In the UK, patients with mild AA (<50% scalp hair loss) commonly receive no treatment immediately after diagnosis and instead undergo a period of 'watch and wait', where management involves the provision of reassurance and advice on the nature and course of AA alone.[1, 29] This is based on the fact that up to 80% of patients with limited patchy hair loss experience spontaneous regrowth within a relatively short period of time, meaning that the use of off-label treatments, with their associated side effects, may not be justified.[1, 14, 15] While waiting for their hair to regrow, these patients can benefit from advice on cosmetic options to conceal hair loss, such as wigs or protheses, as this can help the patient cope with the psychological impact of the disease.[1, 29] Patients may be eligible for a free or reduced cost wig on the NHS.[29, 57] Psychological support may also be valuable for some patients, which can include contact with patient support organisations, such as the National Alopecia Areata Foundation and Alopecia UK, as well as professional support from a clinical psychologist or other practitioner skilled in helping patients cope with their mental health.[1, 2, 44]

Although many patients with mild AA are likely to spontaneously remit without intervention, it may be preferrable for some patients to initiate treatment immediately after diagnosis. In these patients, treatment typically involves the use of either potent topical steroids or intralesional corticosteroids. While topical corticosteroids may advance regrowth in some patients with mild AA, evidence supporting the effectiveness of topical steroids is generally limited and often conflicting.[1, 45] Similarly, while intralesional corticosteroids (the only other licensed treatment for AA) are generally considered the most effective treatment for mild patchy AA, evidence is limited, especially in the long term. [1, 2] This treatment involves injecting corticosteroids directly into areas

of hair loss (~1 injection per square centimetre [cm²]), resulting in high drug concentrations at the lesion site.[58] Both of these treatments may also be trialled in patients with severe AA, though the benefits appear to be particularly low in this subgroup, as discussed below.[1, 16]

There is a substantial unmet need for evidence-based, effective, and well-tolerated medications for the treatment of severe AA

Given the lack of evidence-based treatment options for AA and the often uncomfortable and unpleasant side effects of treatment, leaving AA untreated and taking a 'watch and wait' approach may also be used initially for patients with severe AA. However, unlike those with a milder form of the disease, this subgroup of patients often has a poor prognosis and rarely experiences spontaneous remission.[1, 16]

Aside from a 'watch and wait' approach, various other mostly off-label treatment options may be trialled in patients with severe AA including topical, intralesional or oral corticosteroids, topical immunotherapy such as DPCP, immunosuppressive drugs such as methotrexate, psoralen plus ultraviolet A light therapy (PUVA), or minoxidil and calcineurin inhibitors.[1] Patients may also benefit from supportive care options such as psychological support and cosmetic concealment of hair loss through the use of wigs may be used to help these patients cope with the potentially significant physiological and psychosocial impact of the condition, which may be particularly pronounced among patients with severe AA.[1, 47-49, 51] Unfortunately, the quality of wigs can vary and some patients may also feel self-conscious about wearing a wig for fear of being discovered.[1, 57]

While topical corticosteroids and minoxidil may advance regrowth in some patients with mild AA, evidence supporting the effectiveness of the treatments is generally limited and often conflicting, especially in patients with severe disease.[1, 45] Therefore, these patients typically fail to achieve a sustained and clinically meaningful response, if any, with these treatments.[1, 2] Evidence for intralesional corticosteroids is similarly limited.[1] Prospective studies have shown that the intralesional injections of corticosteroids, in the form of triamcinolone acetonide (5–10 mg/ml) can stimulate tufts of hair regrowth at 60–67% of injection sites.[44] However, this effect is temporary, meaning that further injections are required every 4 to 6 weeks in order to achieve a more sustained response.[1] Frequent injections are resource intensive and can also lead to common complications such as skin atrophy and hypopigmentation, and may also be uncomfortable and unpleasant for patients during administration.[1, 2] In addition, due to the number of injections and drug volumes that would be required for large surfaces, intralesional steroids are not considered feasible beyond 20% hair loss.[33, 59]

PUVA, calcineurin inhibitors and continuous and pulsed systemic oral corticosteroids have also been used to treat AA, though due to their potentially serious side effects and inadequate evidence of efficacy, the risks associated with these treatments may not be justified in many patients.[1] For instance, in the only placebo-controlled study of systemic corticosteroid use in AA, 8 of 23 eight patients with >40% hair loss receiving a weekly single dose of prednisolone had substantial (>31%) hair regrowth compared with none in the placebo group, though this was not statistically significant. However, within three months, 25% of the responders had relapsed.[60] Therefore, while a response may be observed in some patients, continued treatment is usually needed to maintain hair growth and avoid relapse. Furthermore, the observed response is in many cases insufficient to justify the known side effects of systemic corticosteroid use, especially those observed with long-term treatment, such as glaucoma, hypertension, osteoporosis and Cushing's syndrome.[1, 44, 61] Similarly, the use of PUVA is characterised by high rates of relapse

and unacceptably high cumulative UVA doses, and its use is therefore not recommended in the BAD clinical practice guidelines.[1]

Contact immunotherapy is the best-documented treatment in severe AA, though there are no RCTs comparing this intervention with placebo. The aim of treatment is to induce low grade allergic contact dermatitis by initially sensitising the patient, and then applying very weak concentrations of a contact allergen directly to the scalp once a week.[1, 44] Commonly used contact allergens include 2,3-diphenylcyclopropenone (DPCP) and squaric acid dibutyl ester (SADBE), though the former is usually the agent of choice.[1] A review of all the published studies of contact immunotherapy concluded that although the response rates vary widely (9–87%), in general, 20–30% patients achieve a worthwhile response, such as sufficient regrowth to enable patients to manage without a wig.[62] However, relapse rates on maintenance regimens or following discontinuation are high.[1, 12, 63] Topical immunotherapy is generally not well tolerated, and can induce severe contact dermatitis in patients and, sometimes, in the provider. It also involves multiple visits to hospital over several months.[1] Contact immunotherapy is also limited to selected specialist centres in the UK, which limits its role in AA management.[1]

Unmet need and positioning of baricitinib in the clinical pathway of care

The expected eligible patient population for baricitinib in UK clinical practice is adult patients with severe AA. This population is in line with the anticipated license indication for baricitinib in the UK and the eligibility criteria for the BRAVE-AA1 and BRAVE-AA2 trials.[30, 31]

Despite the significant burden associated with AA, there are currently few evidence-based treatment options for patients with AA. The current management of AA therefore chiefly relies on 'watch and wait' or off-label treatments that are not supported by robust evidence from clinical trials.[1] While some patients respond to existing therapies initially, most experience relapse after stopping treatment and most current modalities cannot be used in the long term due to adverse events. In addition, current treatments can be uncomfortable for the patient, are time-consuming and have limited availability. The efficacy of these treatments is also reported to be particularly low in patients with severe AA.[1, 21, 33] Existing treatments for AA are therefore insufficient for patients with severe disease due to the lack of robust efficacy evidence that often cannot outweigh the poor long-term safety and tolerability.[1] As such, many patients with severe AA must rely on management options such as concealment of hair loss and psychological support to cope with the potentially significant physiological and psychosocial impact of the condition, even if they still want to be treated.[1]

The results of the pivotal Phase III clinical trials for baricitinib, BRAVE-AA1 and BRAVE-AA2, demonstrate that baricitinib is an efficacious new treatment for AA with an acceptable safety profile, including with longer-term use up to 52 weeks. Among patients with severe AA, baricitinib significantly improved hair regrowth on the scalp, eyebrows and eyelashes, with some responders experiencing complete hair regrowth within 36 weeks. In addition, Week 52 data indicate that hair regrowth continues to increase among responders beyond 36 weeks, demonstrating that some patients may experience further clinical benefit from baricitinib beyond this timepoint. Furthermore, Week 76 data suggest that efficacy is maintained over time in most of the patients who have reached SALT≤20 by Week 52. Patients who received baricitinib also experienced greater reductions in Skindex-16 AA and HADS scores compared with placebo, indicating an improvement of the HRQoL and psychological burden associated with AA.[30, 31]

A positive recommendation for baricitinib in this setting would therefore allow patients to benefit from improved outcomes compared with current management, reducing the significant and

negative burden of the disease on patients' lives. This would also provide the first evidence-based therapeutic option to address the significant unmet need for an effective and tolerable treatment option for patients with severe AA, by representing the first evidence-based treatment specific to AA.

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

Summary of clinical effectiveness of baricitinib

- The efficacy and safety of baricitinib for the treatment of severe alopecia areata has been evaluated in two multicentre, randomised, double-blind, placebo-controlled trials: an adaptive phase II/III study (BRAVE-AA1) and a phase III study (BRAVE-AA2)
- Treatment with baricitinib was associated with a statistically significant hair regrowth on the scalp in comparison with placebo. The proportion of patients achieving SALT≤20 (≥80% scalp hair coverage) score at Week 36 in BRAVE-AA1 were [redacted]%, [redacted]%, and [redacted]% for baricitinib 4 mg, 2 mg, or placebo, respectively and [redacted]%, [redacted]%, and [redacted]% in BRAVE-AA2
    - Scalp hair regrowth following treatment with baricitinib 4 mg increased over 36 weeks of treatment, and was associated with a statistically significant higher proportion of patients achieving SALT≤20 at Week 16 and Week 24 as compared with placebo
    - A higher proportion of patients receiving baricitinib 4 mg also achieved SALT⩽10 (⩾ 90% scalp hair coverage) compared with placebo at Week 36 in BRAVE-AA1 ([redacted]% vs [redacted]%; p[redacted]), and in BRAVE-AA2 ([redacted]% vs [redacted]%; p[redacted]), representing almost complete hair regrowth in these patients
- Treatment with baricitinib 4 mg led to a statistically significant eyelash and eyebrow regrowth at Week 36 when compared with placebo treatment, with the proportion of patients achieving an improvement in hair growth increasing over the treatment period
- The SALT≤20 response rate for baricitinib continued to increase from Week 36 through Week 52. Improvements in other secondary endpoints including SALT≤10, $SALT_{50}$ (50% improvement in SALT score from baseline) and the ClinRO Measures for eyebrow or eyelash hair loss also continued to increase from Week 36 through Week 52.
- Among patients who achieved SALT≤20 at Week 52, [redacted] of those who were randomised to remain on baricitinib 4 mg in a down-titration sub-study, and [redacted] of those who were assigned to remain on baricitinib 4 mg in a withdrawal sub-study achieved SALT≤20 at week 76, suggesting that efficacy is maintained over time in most of the patients who have reached SALT≤20. Among the patients who had reached SALT≤20 at week 52, the proportion who had also achieved SALT≤10 increased up to week 76.
- Baricitinib was associated with a notable improvement in HRQoL when compared with placebo. Statistically significant improvements were associated with baricitinib 4 mg as measured by the Skindex-16 AA tool across both trials and by HADS-Depression and HADS-Anxiety scores in BRAVE-AA2

Summary of safety evidence of baricitinib

- The BRAVE-AA trials found baricitinib to have a tolerable safety profile, with nasopharyngitis and headaches representing the most common AE in BRAVE-AA1 and BRAVE-AA2, respectively. No new safety findings were identified in the BRAVE-AA trials compared with the known safety profile of baricitinib established in other indications.
- A numerically higher proportion of baricitinib-treated patients reported TEAEs and SAEs as compared with placebo
- No deaths occurred in the placebo or baricitinib treatment groups across both trials

Conclusions

- Severe AA leads to a significant burden on patients' health-related quality of life, with limited treatment options available
- Baricitinib is a clinically effective treatment for patients with severe AA and offers to fulfil the substantial unmet need for an evidence-based, effective and well-tolerated medication in this indication
- As the first evidence-based treatment specific to AA, baricitinib is an innovative therapy that represents an important milestone in the treatment of severe AA

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 comparators for the treatment of AA. The original SLR was conducted in July 2021 and was subsequently updated in February 2022. In total, the updated SLR identified 47 unique studies, with evidence generated from 13 RCTs and 34 observational studies. One randomised clinical trial investigating the safety and efficacy of baricitinib, the phase II portion of BRAVE-AA1, was identified during the SLR; the Phase III portion of the BRAVE-AA1 study will be presented in detail throughout Section B2. The remaining identified studies investigated various baricitinib comparators, the details of which will be discussed in greater detail in Section B.2.9. Full details of the SLR search strategy, study selection process and results can be found in Appendix D. A further study, investigating the safety and efficacy of baricitinib, BRAVE-AA2, was not captured in the SLR, however this is presented in detail throughout Section B2.

B.2.2 List of relevant clinical effectiveness evidence

One publication was identified in the SLR that provided clinical evidence for the efficacy and safety of baricitinib for the treatment of severe AA. *King et al.* (2021) reports the Phase II portion of the randomised, double-blind, placebo-controlled trial BRAVE-AA1 [NCT03570749].[64] Evidence for the efficacy and safety of baricitinib in severe AA is further provided by King *et al.* (2022) reporting the Phase III portion of the BRAVE-AA1, and the identically-designed Phase III study BRAVE-AA2 [NCT03899259].[30, 31] Given the availability of these Phase III data, the Phase II portion of the BRAVE-AA1 trial has not been considered further within this submission. Overviews of these Phase III RCTs are provided in Table 5.

The patient populations for BRAVE-AA1 and BRAVE-AA2 were adult patients with severe, with limited permitted concomitant medications allowed. This population is in line with the population of relevance for this evaluation; adult patients with severe AA. As such, a pooled population of the patients from BRAVE-AA1 and BRAVE-AA2 informs the base case economic analysis.

Table 5. Clinical effectiveness evidence

Study	BRAVE-AA1	BRAVE-AA2
Study design	An adaptive, Phase 2/3, multi-centre, randomised, double-blind, placebo-controlled, parallel-group, outpatient study.	A Phase 3, multi-centre, randomised, double-blind, placebo-controlled, parallel-group study.
Population	N=654	N=546
	Adult patients (≤60 years for males, ≤70 years for females) with severe AA, defined as: • current AA episode of more than 6 months' duration and hair loss encompassing ≥50% of the scalp, as measured by SALT at Visit 1 and Visit 2 • no spontaneous improvement over the past 6 months • current episode of severe or very severe AA of less than 8 years (patients with episode of ≥8 years may be enrolled if episodes of regrowth, spontaneous or under treatment, had been observed)	
Intervention(s)	Baricitinib once daily (4 mg, 2 mg)	
Comparator(s)	Placebo	
Indicate if trial supports application for marketing authorisation	Yes	
Indicate if study used in the economic model	Yes	
Rationale for use in the model	The objective of these trials was to demonstrate efficacy, safety and tolerability of baricitinib in patients with severe AA, defined as those with SALT score higher or equal to 50 at baseline. This patient population is considered to be most relevant to UK clinical practice as it is expected that clinicians will use baricitinib as a first-line therapy in patients with severe AA. Patients from these trials were therefore pooled to inform the base case of the economic model.	
Reported outcomes specified in the decision problem	Measures of disease severity and symptom control: • **SALT** • PRO for Scalp Hair Assessment • ClinRO Measure for Eyebrow Hair Loss • ClinRO Measure for Eyelash Hair Loss • PRO Measure for Eyebrows • PRO Measure for Eyelashes HRQoL • **HADS** • Skindex-16 adapted for AA • SF-36 • **EQ-5D-5L** **Safety outcomes**	

Abbreviations: AA: alopecia areata; ClinRO: clinician reported; HRQoL: health related quality of life; EQ-5L-5L: 5-level EuroQol 5 Dimensions; HADS: hospital anxiety and depression score; PRO: patient reported outcome; SALT: severity of alopecia tool; SF-36: medical outcomes study 36-item short form health survey.

Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

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

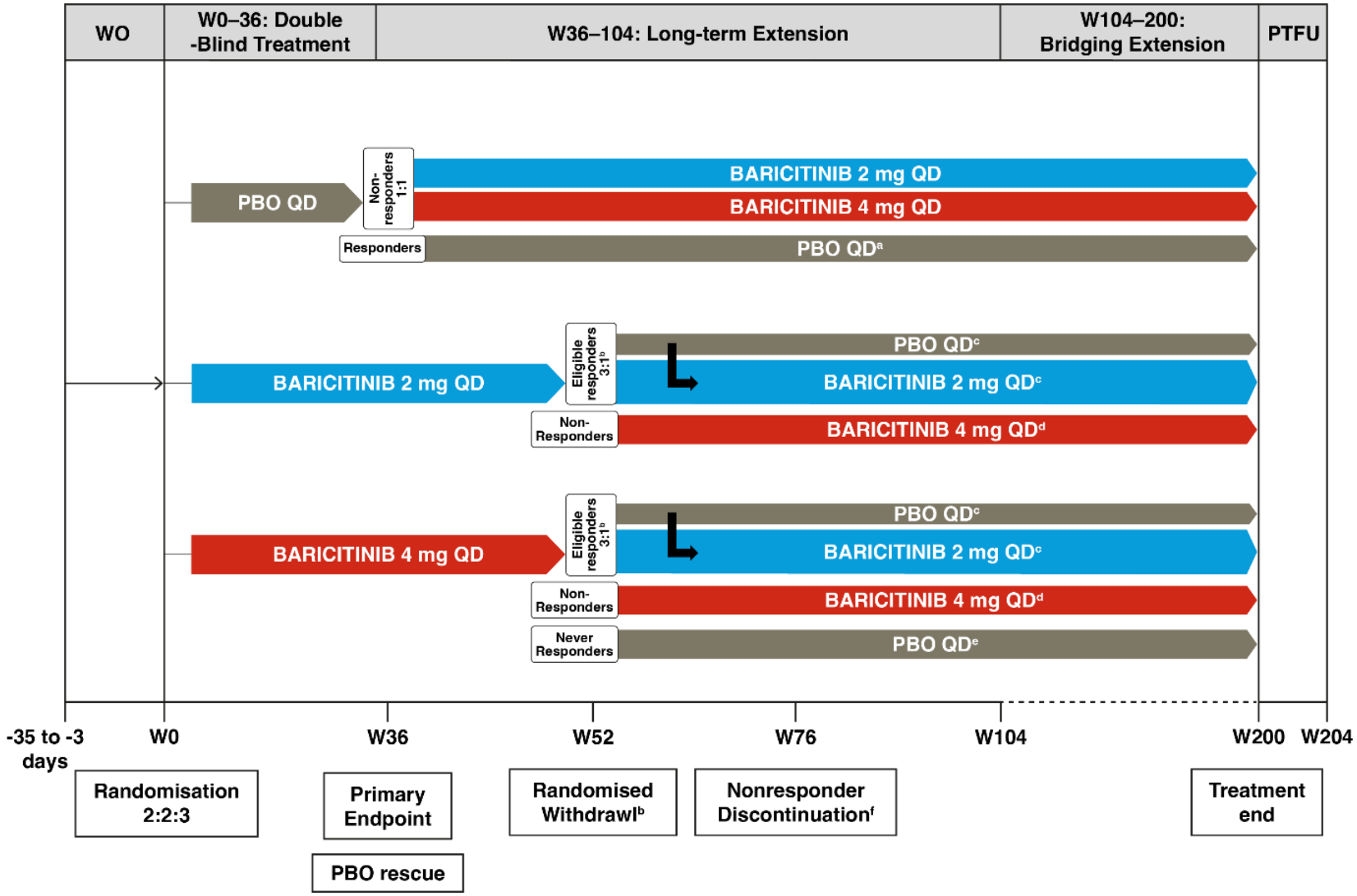
B.2.3.1 Summary of trial methodology

BRAVE-AA1 and BRAVE-AA2 are ongoing, multicentre, double-blind, randomised, placebo-controlled studies to determine the safety and efficacy of baricitinib in adult patients with severe AA. BRAVE-AA1 was an adaptive phase II/III study. The Phase II portion of the study was designed to determine the doses of baricitinib for use in the phase III trials; doses of 1 mg, 2 mg and 4 mg were evaluated. Based on the findings from the phase II portion of BRAVE-AA1, the 2 mg and 4 mg doses were selected for further investigation in the Phase III portion of BRAVE-AA1 and the Phase III BRAVE-AA2 study. The current submission will focus on the Phase III results only as the data from patients enrolled in the phase II portion of BRAVE-AA1 are not included in the efficacy analyses of the phase III part of the trial.

The patient populations for BRAVE-AA1 and BRAVE-AA2 were adult patients with severe AA. This was defined as patients with a current AA episode of more than 6 months duration prior to screening (Visit 1), with at least 50% scalp hair loss at screening and baseline (Visits 1 and 2) and no spontaneous improvement (no more than a 10-point spontaneous reduction in Severity of Alopecia Tool [SALT] score) over the past 6 months. The use of concomitant medications for the management of AA was generally prohibited throughout the trial, with some exceptions detailed within the study protocols.

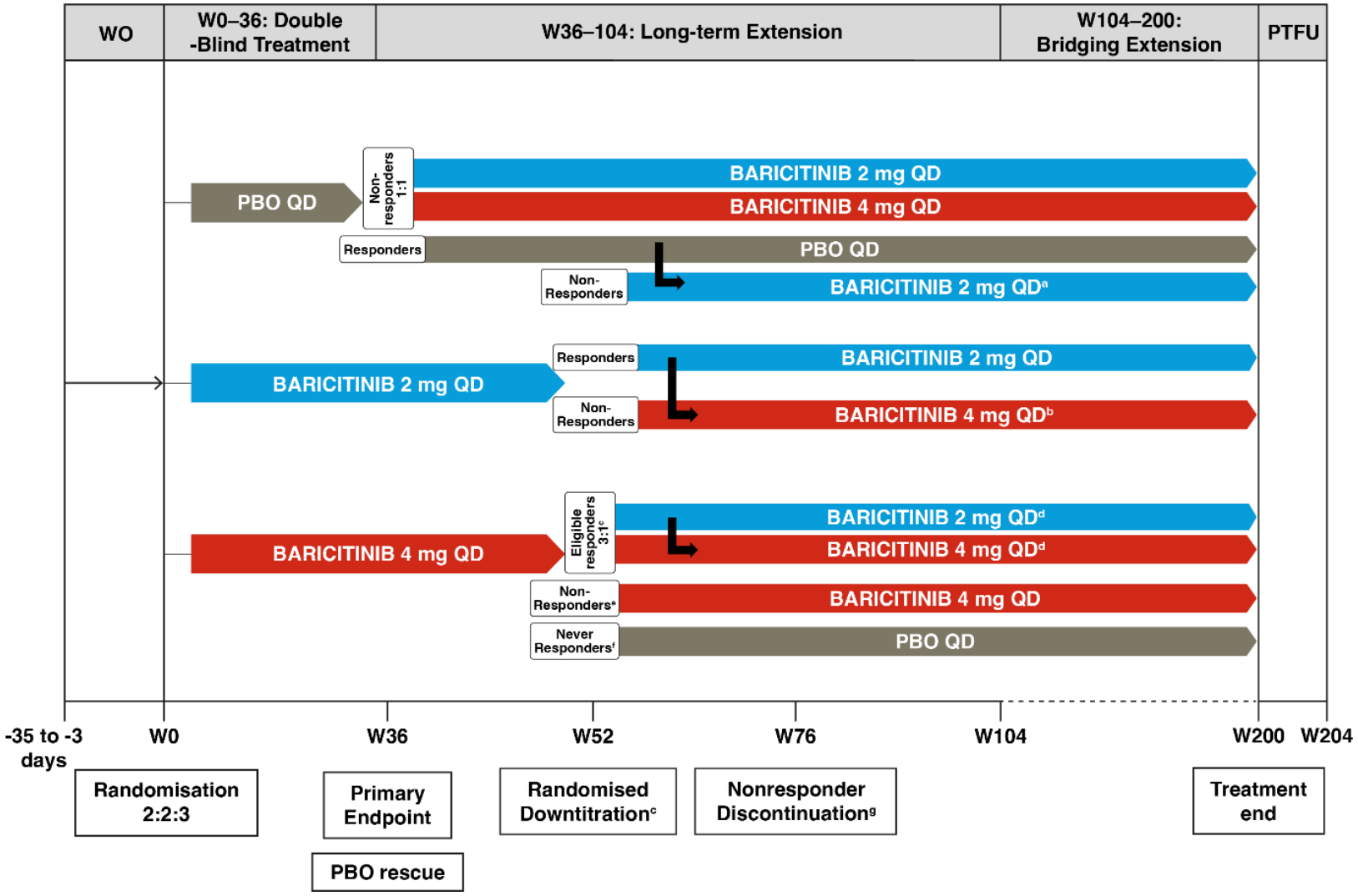
The study design of BRAVE-AA1 and BRAVE-AA2 is shown in Figure 3 and Figure 4, respectively. Before Week 52, the study designs of BRAVE-AA1 and BRAVE-AA2 were largely identical. Following a screening period between 3 and 35 days prior to Visit 2 (Week 0) patients were randomised in a double-blinded fashion to once-daily treatment with placebo, 2 mg baricitinib or 4 mg baricitinib (2:2:3) Once patients completed Week 36, they entered the long-term extension period and, through to Week 52, either remained on baricitinib regardless of response, remained on placebo if they had achieved SALT≤20 at Week 36, or in the case of non-response in placebo-treated patients were rescued to baricitinib. From Week 52 onward, the study designs for BRAVE-AA1 and BRAVE-AA2 differed. In BRAVE-AA1, baricitinib-treated (2 mg or 4 mg) patients who achieved SALT≤20 at Week 52 were eligible to participate in a randomised withdrawal sub-study if they had remained on the same dose of baricitinib from initial randomisation. In BRAVE-AA2, patients treated with 4 mg baricitinib who achieved SALT≤20 at Week 52 were eligible to participate in a randomised down-titration sub-study if they had remained on baricitinib 4 mg from initial randomisation. Further information on the predefined criteria for continuation into the long-term extension and bridging extension is provided in the study protocols.

Figure 3. Study design of BRAVE-AA1



Footnotes: [a] Placebo responders stayed on placebo for remainder of the trial, even if relapse was observed later. [b] Patients with SALT ≤20 who stayed on the same dose of baricitinib from week 0 were randomised to stay on current baricitinib dose, or transitioned to placebo.[c] Responders participating in randomised withdrawal who experienced >20-point absolute worsening in total SALT score after week 52 were retreated with baricitinib dose to which they were originally randomised if they were randomised to placebo at week 52, OR continued to receive same dose of baricitinib if they were randomised to remain on baricitinib at week 52. [d] Non-responders at week 52 were rescued to baricitinib 4 mg if receiving baricitinib 2 mg from baseline, OR remained on baricitinib 4 mg if they were in the 4-mg group and achieved SALT ≤20 before week 52. [e] Never responders (never achieved SALT ≤20 by week 52 despite being in the baricitinib 4-mg group from baseline and had not experienced a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss) were automatically transitioned to placebo.[f] Non-responders at week 52 AND week 76 were automatically discontinued at week 76 unless they had a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss.
Abbreviations: PBO: placebo; PTFU: post-trial follow up; QD: once daily; W: week.
Source: BRAVE-AA1 Clinical Study Report.[65]

Figure 4. Study design of BRAVE-AA2



Footnotes: [a] Placebo-treated patients not eligible for rescue to baricitinib at week 36 (due to spontaneous remission) were rescued to baricitinib if they were non-responders at week 52, OR if they experienced loss of treatment benefit after week 52. [b] Patients randomised to baricitinib 2 mg at week 0 were rescued to the 4-mg dose if they were non-responders at week 52, OR were responders at week 52 but experienced a >20-point worsening in SALT score after week 52. [c] Responders in the baricitinib 4-mg group (SALT ≤20 who stayed on 4 mg from week 0) were randomised to either stay on 4 mg OR transition to 2 mg.[d] Responders participating in the randomised down-titration who experienced a loss of treatment benefit after week 52 were re-treated with baricitinib 4 mg if they were randomised to the 2-mg dose at week 52, OR continued to receive baricitinib 4 mg if they randomised to remain on the 4-mg dose at week 52. [e] At week 52, non-responders (SALT >20) in the baricitinib 4-mg group since baseline who achieved SALT ≤20 before week 52 remained on 4 mg. [f] Never responders (never achieved SALT ≤20 by week 52 despite being in the baricitinib 4-mg group from baseline and had not experienced a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss) were automatically transitioned to placebo. [g] Non-responders at week 52 AND week 76 were automatically discontinued at week 76 unless they had a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss.

Abbreviations: PBO: placebo; PTFU: post-trial follow up; QD: once daily; W: week.

Source: BRAVE-AA2 Clinical Study Report.[66]

The primary outcome of both trials was to evaluate whether 2 mg or 4 mg baricitinib is superior to placebo, as measured by the proportion of patients in each treatment group achieving SALT≤20 after Week 36 of treatment. Key secondary endpoints measured scalp hair regrowth at Week 12, 16, 24 and 36 using various other SALT score thresholds. Hair regrowth was also measured at Week 36 using ClinRO for eyebrow and eyelash hair loss and patients reported outcomes (PRO scalp hair assessment score). Other secondary outcomes included HRQoL outcomes, including Skindex-16 AA domain scores and Hospital Anxiety and Depression Scale (HADS). Safety outcomes included AEs, SAEs and TEAEs by Week 36.

A summary of the methodology of BRAVE-AA1 and BRAVE-AA2 is presented in Table 6.

Table 6: Summary of methodology for BRAVE-AA1 and BRAVE-AA2

Trial name	**BRAVE-AA1 (Phase III portion)**	**BRAVE-AA2**
Location	International: patients recruited from 70 sites across 3 countries (United States, South Korea, and Mexico). United States sites enrolled 54.7% of patients, South Korea sites enrolled 37.8% of patients, and Mexico sites enrolled 7.5% of patients	International: patients recruited from 98 sites across 9 countries. United States sites enrolled 34.8% of patients, Asia sites recruited 26.9% of patients, and the remaining patients (38.3%) were recruited from sites including Argentina, Australia, Brazil and Israel
Trial design	An adaptive phase II/III, multi-centre, randomised, double-blind, placebo-controlled, study	A phase III, multi-centre, randomised, double-blind, placebo-controlled, parallel-group study
Eligibility criteria for participants	**Key inclusion criteria:** • Aged 18 years or older and ≤60 years of age for males and ≤70 years of age for females at the time of informed consent • Severe AA, as determined by all of the following: Current AA episode of more than 6 months' duration and hair loss encompassing ≥50% of the scalp, as measured by SALT at Visit 1 AND Visit 2 No spontaneous improvement (that is, no more than 10-point spontaneous reduction in SALT) over the past 6 months Current episode of severe AA of less than 8 years Patients who have severe AA for ≥8 years may be enrolled if episodes of regrowth, spontaneous or under treatment, have been observed on the affected areas of the scalp over the past 8 years **Key exclusion criteria:** • Primarily "diffuse" type of AA (characterised by diffuse hair shedding) • Currently experiencing other forms of alopecia, or any other concomitant conditions that would interfere with evaluations of the effect of study medication on AA • Had inadequate washout with the following therapies including but not limited to: ○ Corticosteroids ○ JAK inhibitors	

	○ Monoclonal antibodies ○ Phototherapy ○ Immunosuppressants • Patients who previously had an inadequate response to oral JAK inhibitors after ≥12 weeks of treatment The eligibility criteria for the long-term extension and bridging periods are outlined in the study protocol.
Duration of study	The total study duration is 200 weeks, with a screening period ranging from 3 to 35 days, a 36-week treatment period, a 68-week long-term extension period, a 104-week bridging extension period and a 28-day post-treatment follow-up
Method of randomisation	Patients were randomised 2:2:3 to once daily treatment with placebo, 2 mg baricitinib or 4 mg baricitinib
Method of blinding	Double-blinding
Trial drugs and method of administration	Administered orally once daily as 2 tablets: 1 placebo tablet with 1 treatment tablet for treatment groups, or 2 placebo tablets for the placebo group.
Permitted and disallowed concomitant medication	A summary of the key concomitant medications permitted during the study period is provided below. Full details are provided in the study protocol. • Topical corticosteroids except on the scalp, eyebrows and eyelashes • Topical calcineurin inhibitors except on the scalp, eyebrows, and eyelashes • Intranasal, ophthalmic, or inhaled steroid use • A maximum of 2 intra-articular or soft tissue (bursa, tendon, and/or ligament) corticosteroid injections are allowed up until the 36-week primary endpoint. After 36 weeks, such injections are permitted • Mon-live vaccinations such as seasonal vaccination, non-live herpes zoster (for subjects who become eligible during the trial), and/or all emergency vaccinations such as rabies or tetanus vaccinations • Bimatoprost ophthalmic solution (if on stable dose for 8 weeks prior to randomization) • Finasteride (or other 5 alpha reductase inhibitors) or oral or topical minoxidil, if on a stable dose for 12 months prior to randomization, and • HMG CoA reductase inhibitors or “statins” (for example, simvastatin, simvastatin + ezetimibe) for treatment of hypercholesterolemia and the prevention of cardiovascular disease
Primary outcomes (including scoring methods and timings of assessments)	Proportion of patients achieving SALT≤20 at Week 36.
Secondary outcomes (including scoring methods and timings of assessments)	A summary of the key secondary outcomes is provided below. • Proportion of patients achieving SALT≤20 (scalp hair loss of ≤20% or ≥80% scalp coverage with hair) at Weeks 16 and 24 • Percent change from baseline in SALT score at Week 36

	• Proportion of patients achieving a $SALT_{50}$ (at least 50% improvement from baseline) at Week 12. • Proportion of patients achieving a $SALT_{90}$ (at least 90% improvement from baseline) at Week 36 • Proportion of patients achieving an absolute SALT≤10 (scalp hair loss of ≤10% or ≥90% scalp coverage with hair) at Weeks 24 and 36 Other secondary endpoints of interest are provided below. Full details of all the secondary outcomes can be found in Appendix M. • Proportion of patients achieving $SALT_{50}$ at Week 36 • Proportion of patients achieving $SALT_{75}$ at Week 36 • Proportion of patients achieving ClinRO Measure for eyebrow Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with ClinRO Measure for eyebrow Hair Loss ≥2 at baseline) • Proportion of patients achieving ClinRO Measure for eyelash Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with ClinRO Measure for eyelash Hair Loss ≥2 at baseline) • Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement from baseline at Week 36 (among patients with a score of ≥3 at baseline) Key HRQoL endpoints are provided below: • Mean change from baseline in HADS-A and HADS-D total scores at Weeks 24 and 36 • Mean change from baseline to Weeks 24 and 36 in Skindex-16 AA domain scores (Symptoms, Emotions, Functioning) • Mean change from baseline in SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) scores to Weeks 24 and 36 • Mean change from baseline in EQ-5D-5L index and VAS scores to Weeks 24 and 36 • Mean change from baseline to weeks 24 and 36 in Skindex-16AA domain scores (Symptoms, Emotions, Functioning)
Pre-specified subgroup analyses	• Gender • Age group (<40, ≥40, <65, ≥65 years) • 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) • Duration of current episode of AA category (<4, ≥4 years) • Baseline SALT category (severe [SALT 50–94], very severe [SALT 95–100])

Footnotes: [a] Copyright issues prevented Skindex-16 from being available at the start of the phase III portion of BRAVE-AA1, therefore, it could not be given to the entire intention-to-treat set. For this reason, it was designated as an exploratory outcome. For BRAVE-AA2, it was available from the start of the study, so it was included as a secondary non-gated outcome measure.
Abbreviations: AA: alopecia areata; BMI: body mass index; AA-IGA: alopecia areata investigator global assessment; ClinRO: clinician reported; HMG CoA: 3-hydroxy-3-methylglutaryl coenzyme A; HRQoL: health related quality of life; EQ-5L-5L: 5-level EuroQol 5 Dimensions; HADS: hospital anxiety and depression score; JAK: Janus kinase; PRO: patient reported outcome; SALT: severity of alopecia tool; SF-36: medical outcomes study 36-item short form health survey.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

B.2.3.2 Outcome definitions

All outcome definitions were consistent across BRAVE-AA1 and BRAVE-AA2 trials and are summarised in Table 7.

Table 7. Definitions of outcomes used in BRAVE-AA1 and BRAVE-AA2

Outcome	Definition
SALT	The SALT uses a visual aid showing the division of the scalp hair into 4 areas, with the sides, top and back of the head constituting 18%, 40% and 24%, respectively. The percentage of hair loss in each area is determined and is multiplied by the percentage of scalp covered by that area. The total sum of the 4 products of each area will give the SALT score. Details on the interpretation of SALT scores are provided in Table 8.
PRO for Scalp Hair Assessment	Lilly developed a novel PRO assessment of the patient's current extent of scalp involvement. The PRO for Scalp Hair Assessment is comprised of 5 categorical response options, with higher score indicating greater disease severity: *0 = No missing hair (0% of my scalp is missing hair; I have a full head of hair)* *1 = A limited area (1% to 20% of my scalp is missing hair)* *2 = A moderate area (21% to 49% of my scalp is missing hair)* *3 = A large area (50% to 94% of my scalp is missing hair)* *4 = Nearly all or all (95% to 100% of my scalp is missing hair)*
Clinician-Reported Outcomes for eyebrow Hair Loss™	Lilly developed a novel ClinRO assessments to measure eyebrow hair loss. The ClinRO Measure for eyebrow hair loss is comprised of a 4-point response scale, with higher score indicating greater disease severity: *0 = The eyebrows have full coverage and no areas of hair loss* *1 = There are minimal gaps in eyebrow hair and distribution is even* *2 = There are significant gaps in eyebrow hair or distribution is not even* *3 = No notable eyebrow hair*
Clinician-Reported Outcomes for eyelash Hair Loss™	Lilly developed a novel ClinRO assessments to measure eyelash loss. The ClinRO Measure for eyelash hair loss is comprised of a 4-point response scale, with higher score indicating greater disease severity: *0 = The eyelashes form a continuous line along the eyelids on both eyes* *1 = There are minimal gaps, and the eyelashes are evenly spaced along the eyelids on both eyes* *2 = There are significant gaps along the eyelids, or the eyelashes are not evenly spaced along the eyelids* *3 = No notable eyelashes*
PRO Measure for Eyebrows	Lilly developed a novel PRO assessment measuring the extent of eyebrow hair loss, using a single item that uses a 4-point response scale, ranging from 0 = I have full eyebrows on each eye to 3 = I have no or barely any eyebrow hairs, with higher score indicating greater disease severity
PRO Measure for Eyelashes	Lilly has developed a novel PRO assessment measuring the extent of eyelash hair loss, using a single item that uses a 4-point response scale, ranging from a score of 0 = I have full eyelashes on each eyelid, to a score of 3 = I have no or barely any eyelash hair, with higher score indicating greater disease severity
HADS	The Hospital Anxiety and Depression Scale (HADS) is a 14 item self-assessment scale that determines the levels of anxiety and depression that a patient is experiencing over the past week. The HADS utilises a 4-point Likert scale (for example, 0 to 3) for each question and is intended for ages 12 to 65 years. Scores for each domain (anxiety and depression) can range from 0 to 21, with higher scores indicating greater anxiety or depression

Skindex-16 Adapted for Alopecia Areata	The Skindex-16 measure is a self-reported questionnaire that assesses the degree to which patients are bothered by their skin condition. The Skindex-16 items' wordings were adapted for use among adults with AA to specifically examine the degree to which the subjects are bothered by AA and its associated symptoms. It is composed of 16 items grouped under 3 domains: Symptoms (4 items), Emotions (7 items), and Functioning (5 items). The score of each item ranges from 0 (never bothered) to 6 (always bothered). Scores are transformed to a linear scale ranging from 0 (no effect) to 100 (effect experienced all the time), with higher scores indicating greater impact on quality of life.
SF-36	The SF-36 is a 36-item, patient-completed measure designed to be a short, multipurpose assessment of health. Items are answered on Likert scales of varying lengths. The SF-36 comprises 8 domain scores and 2 overarching component scores. SF-36 domain scores are: (1) Physical Functioning; (2) Role-Physical; (3) Role-Emotional; (4) Bodily Pain; (5) Vitality; (6) Social Functioning; (7) Mental Health; and (8) General Health. Higher scores indicate better levels of function and/or better health.
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 visual analogue scale (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: AA: alopecia areata; ClinRO: clinician reported outcomes; HRQoL: health related quality of life; EQ-5L-5L: 5-level EuroQol 5 Dimensions; HADS: hospital anxiety and depression score; PRO: patient reported outcome; SALT: severity of alopecia tool; SF-36: medical outcomes study 36-item short form health survey; VAS: visual analogue scale.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

SALT Score Interpretation

BRAVE-AA1 and BRAVE-AA2 employed both absolute and subscript SALT scores as outcomes to measure hair regrowth following treatment with baricitinib. The interpretation of these types of SALT scores is explained in Table 8. An absolute SALT score represents a percentage of scalp hair loss, which can also be expressed as a percentage of scalp coverage with hair, whereas a subscript of the SALT score represents a percentage change from baseline.

Table 8. Interpretation of the SALT score

Absolute scores with interpretation	Subscript score with interpretation
SALT 0: no hair loss (or 100% scalp coverage with hair)	**$SALT_{50}$:** at least 50% improvement from baseline
SALT 50: 50% hair loss (or 50% scalp coverage with hair)	**$SALT_{90}$:** at least 90% improvement from baseline
SALT 100: no hair (or 0% scalp coverage with hair)	

Abbreviations: SALT: Severity of Alopecia Tool.

B.2.3.3 Baseline characteristics of study participants

Demographic and Disease Characteristics

Patient demographics were similar between the BRAVE-AA1 and BRAVE-AA2 trials and were generally well balanced across treatment groups within each study. In addition, baseline disease characteristics and geographic region distribution did not differ substantially between the two studies and were well balanced across treatment groups (Table 9, Table 10).

Patients enrolled in BRAVE-AA1 and BRAVE-AA2 had a mean age of 37.5 years, with women representing a slightly higher proportion of the patient population across the studies compared with males (61% vs 39%). The median SALT score was 96 across the study populations, and the mean disease episode duration was 3.9 years. The mean duration from the first onset of AA diagnosis was 12.2 years, and over half of all patients enrolled in the studies had very severe AA at baseline (defined as SALT 95–100), with [redacted] of patients having investigator-reported alopecia universalis (AU).[67]

Across the two studies, 69% of patients had significant or complete eyebrow hair loss at baseline, and 58% had significant or complete eyelash hair loss at baseline, as measured by ClinRO measures for eyebrow and eyelash hair loss scores of 2 or 3. Approximately 38% of patients reported an atopic background, defined as a medical history of, or ongoing atopic dermatitis, allergic rhinitis, allergic conjunctivitis, or allergic asthma.[67]

Table 9. Baseline and disease characteristics of patients in BRAVE-AA1

Characteristic	BRAVE–AA1		
	PBO (N=189)[a]	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)
Mean (SD) age, years	37 (12.91)	38 (12.78)	36 (13.27)
Female, n (%)	109 (57.7)	109 (59.2)	165 (58.7)
Geographic region, n (%)			
North America	103 (54.5)	102 (55.4)	153 (54.4)

Characteristic	BRAVE–AA1		
	PBO (N=189)[a]	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)
Asia[b]	70 (37.0)	70 (38.0)	107 (38.1)
Rest of the world[c]	16 (8.5)	12 (6.5)	21 (7.5)
Race, n (%)			
White	83 (44.1)	93 (50.8)	123 (43.9)
Asian	78 (41.5)	76 (41.5)	114 (40.7)
Black or African American	17 (9.0)	7 (3.8)	28 (10.0)
American Indian or Alaska Native	8 (4.3)	5 (2.7)	8 (2.9)
Native Hawaiian or Other Pacific Islander	1 (0.5)	1 (0.5)	1 (0.4)
Mean (SD) BMI, kg/m2	[redacted]	[redacted]	[redacted]
Atopic background, n (%)	73 (38.6)	67 (36.4)	97 (34.5)
Mean (SD) duration since onset of AA, years	12.6 (11.2)	12.10 (9.8)	11.8 (11.1)
Mean (SD) duration of current AA episode, years	3.53 (3.65)	3.86 (4.69)	3.46 (3.37)
Age of AA onset, n (%)			
<18 years	[redacted]	[redacted]	[redacted]
≥18 years	[redacted]	[redacted]	[redacted]
Duration of current AA episode, n (%)			
<4 years	134 (70.9)	127 (69.0)	189 (67.3)
≥4 years	55 (29.1)	57 (31.0)	92 (32.7)
Mean (SD) SALT score	84.7 (17.82)	86.8 (18.01)	85.3 (18.18)
SALT category, n (%)			
Severe (SALT 50–94)	92 (48.7)	77 (41.8)	133 (47.3)
Very severe (SALT 95–100)	97 (51.3)	107 (58.2)	148 (52.7)
Patients with AU, n (%)	74 (39.2)	83 (45.1)	127 (45.2)
ClinRO for eyebrow hair loss, n (%)			
2	53 (28.3)	46 (25.0)	73 (26.3)
3	71 (38.0)	90 (48.9)	115 (41.4)
ClinRO for eyelash hair loss, n (%)			
2	38 (20.3)	35 (19.0)	74 (26.6)
3	58 (31.0)	76 (41.3)	93 (33.5)
PRO for Scalp Hair Assessment			
3 (50–94% hair loss)	72 (38.1)	57 (31.0)	102 (36.4)

Characteristic	BRAVE–AA1		
	PBO (N=189)[a]	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)
4 (95–100% hair loss)	109 (57.7)	118 (64.1)	173 (61.8)
Mean (SD) Skindex–16 AA baseline domain scores[f]			
Emotions	[redacted]	[redacted]	[redacted]
Functioning	[redacted]	[redacted]	[redacted]
Symptoms	[redacted]	[redacted]	[redacted]
Mean (SD) HADS total score			
HADS-Anxiety	6.7 (3.92)	6.2 (3.74)	6.1 (3.80)
HADS-Depression	4.0 (3.15)	4.2 (3.66)	4.0 (3.39)

Footnotes: [a] Disease characteristics data based on observations in n=188 patients.[b] Includes South Korea. [c] Includes Mexico. [d] Atopic background is defined as having an ongoing or a medical history of atopic dermatitis, allergic rhinitis, allergic conjunctivitis, or allergic asthma. [f] Copyright issues prevented Skindex-16 from being available at the start of the phase III portion of BRAVE-AA1, so it could not be administered to the entire intention-to-treat set (PBO N=119; 2 mg baricitinib N=111; 4 mg baricitinib N=175).
Abbreviations: AA: alopecia areata; AU: alopecia universalis; BMI: body mass index; ClinRO: clinician reported outcomes; HADS: hospital anxiety and depression score; PBO: placebo; PRO: patient reported outcome; SALT: severity of alopecia tool; SD: standard deviation.
Source: BRAVE-AA1 Clinical Study Report; King et al. 2022; EPAR.[65, 67, 68]

Table 10. Baseline and disease characteristics of patients in BRAVE-AA2

Characteristic	BRAVE-AA2		
	PBO (N=156)[a]	2 mg baricitinib (N=156)	4 mg baricitinib (N=234)
Mean (SD) age, years	37 (12.35)[a]	39 (12.99)	38 (12.65)
Female, n (%)	98 (62.8)	103 (66.0)	144 (61.5)
Geographic region, (%)			
North America	54 (34.6)	54 (34.6)	82 (35.0)
Asia[b]	42 (26.9)	42 (26.9)	63 (26.9)
Rest of the world[c]	60 (38.5)	60 (38.5)	89 (38.0)
Race, n (%)			
White	85 (54.5)	92 (59.0)	144 (61.5)
Asian	51 (32.7)	49 (31.4)	67 (28.6)
Black or African American	16 (10.3)	12 (7.7)	18 (7.7)
American Indian or Alaska Native	0 (0.0)	0 (0.0)	0 (0.0)
Native Hawaiian or Other Pacific Islander	0 (0.0)	1 (0.6)	0 (0.0)
Mean (SD) BMI, kg/m^2	[redacted]	[redacted]	[redacted]
Atopic background, n (%)[d]	67 (42.9)	63 (40.4)	87 (37.2)

Characteristic	BRAVE-AA2		
	PBO (N=156)[a]	2 mg baricitinib (N=156)	4 mg baricitinib (N=234)
Mean (SD) duration since onset of AA, years	11.79 (10.190)	13.08 (11.795)	11.89 (11.122)
Mean (SD) duration of current AA episode, years	4.68 (5.490)	4.39 (6.088)	3.94 (3.353)
Age of AA onset, n (%)			
<18 years	57 (36.5)	55 (35.3)	74 (31.6)
≥18 years	99 (63.5)	101 (64.7)	160 (68.4)
Duration of current AA episode, n (%)			
<4 years	94 (60.3)	103 (66.0)	140 (59.8)
≥4 years	62 (39.7)	53 (34.0)	94 (40.2)
Mean (SD) SALT score	85.0 (17.79)	85.6 (18.08)	84.8 (18.08)
SALT category, n (%)			
Severe (SALT 50–94)	74 (47.7)	70 (44.9)	115 (49.1)
Very severe (SALT 95–100)	81 (52.3)	86 (55.1)	119 (50.9)
Patients with AU, n (%)	66 (42.3)	70 (44.9)	111 (47.4)
ClinRO for eyebrow hair loss, n (%)			
2	46 (30.1)	35 (22.4)	49 (21.0)
3	66 (43.1)	69 (44.2)	112 (48.1)
ClinRO for eyelash hair loss, n (%)			
2	31 (20.3)	26 (16.7)	43 (18.5)
3	59 (38.6)	63 (40.4)	97 (41.6)
PRO for Scalp Hair Assessment			
3 (50–94% hair loss)	60 (38.5)	56 (35.9)	78 (33.3)
4 (95–100% hair loss)	91 (58.3)	93 (59.6)	137 (58.5)
Mean (SD) Skindex–16 AA baseline domain scores			
Emotions	[redacted]	[redacted]	[redacted]
Functioning	[redacted]	[redacted]	[redacted]
Symptoms	[redacted]	[redacted]	[redacted]
Mean (SD) HADS total score			
HADS-Anxiety	5.9 (4.01)	6.2 (3.88)	6.4 (3.95)
HADS-Depression	3.7 (3.46)	3.8 (3.27)	3.8 (3.49)

Footnotes: [a] One patient in the placebo group in BRAVE-AA2 was inadvertently enrolled with a SALT score of 32; therefore, this patient did not fall within the severe or very severe SALT categories. [b] Includes Japan, China, Taiwan and South Korea. [c] Includes Australia, Brazil, Argentina and Israel. [d] Atopic background is defined as having an ongoing or a medical history of atopic dermatitis, allergic rhinitis, allergic conjunctivitis, or allergic asthma
Abbreviations: AA: alopecia areata; AU: alopecia universalis; BMI: body mass index; ClinRO: clinician reported outcomes; HADS: hospital anxiety and depression score; PBO: placebo; PRO: patient reported outcome; SALT: severity of alopecia tool; SD: standard deviation.
Source: BRAVE-AA2 Clinical Study Report; King et al. 2022; EPAR.[66-68]

Prior and concomitant therapy

Approximately [redacted] of the patients in the BRAVE-AA1 (Table 11) and BRAVE-AA2 (Table 12) trials reported using prior medications to treat their AA. Overall, more than [redacted] of patients had used topical immunotherapy and [redacted] had used systemic immunosuppressant or immunomodulator therapy, the most common of which were corticosteroids. JAK inhibitors had been used by almost [redacted] of patients, however it should be noted that prior inadequate response to JAK inhibitors was an exclusion criterion for the trials. The high proportion of patients who had received prior medications highlights the inadequacy of current treatment options in maintaining a longer-term hair regrowth and reflects the unmet need for novel treatment options in patients with severe AA.

A washout of systemic and topical treatments for AA was incorporated before randomisation to minimise the confounding effects of prior treatment. Durations of required washouts took into consideration that a delay of several weeks may be observed between treatments and regrowth of hair in patients with AA.

Only very few concomitant medications and procedures for the treatment of AA were permitted during BRAVE-AA1 and BRAVE-AA2, and only [redacted] of the enrolled patients were reported to have used such therapies.

Table 11. Prior treatments for AA used among participants in BRAVE-AA1 (FAS population)

	BRAVE-AA1		
	PBO (N=189)	**2 mg baricitinib (N=184)**	**4 mg baricitinib (N= 281)**
Prior therapy, n (%)[a]	173 (91.5)	[redacted]	247 (87.9)
Topical therapy, n (%)[b]	108 (57.1)	102 (55.4)	173 (61.6)
Topical IMT, n (%)	45 (23.8)	57 (31.0)	84 (29.9)
Intralesional therapy, n (%)	101 (53.4)	92 (50.0)	152 (54.1)
Systemic agents, n (%)[c]			
Immunosuppressant/ immunomodulator	101 (53.4)	84 (45.7)	138 (49.1)
Corticosteroids	68 (36.0)	51 (27.7)	103 (36.7)
JAK inhibitor	12 (6.3)	7 (3.8)	15 (5.3)
Others	57 (30.2)	55 (29.9)	88 (31.3)
Cyclosporin	46 (24.3)	45 (24.5)	69 (24.6)
Methotrexate	15 (7.9)	17 (9.2)	28 (10.0)

Other systemic (non-immunosuppressant), n (%)	17 (9.0)	20 (10.9)	28 (10.0)
Phototherapy, n (%)	23 (12.2)	34 (18.5)	54 (19.2)
Procedures, n (%)[d]	30 (15.9)	41 (22.3)	65 (23.1)

Footnotes: [a] Patients may have had >1 prior therapy; therefore, the n in each group may not add up to the total N for the column. [b] Topical therapies excluding IMT. [c] All immunosuppressant and immunomodulator agents including CS, JAK inhibitors, and others. [d] Procedures include cryotherapy, micro-needling, and platelet-rich plasma injections.
Abbreviations: IMT: immunomodulatory therapy; JAK: Janus kinase; PBO: placebo.
Source: BRAVE-AA1 Clinical Study Report; EPAR.[65, 68]

Table 12. Prior treatments for AA used among participants in BRAVE-AA2 (FAS population)

	BRAVE-AA2		
	PBO (N=156)	**2 mg baricitinib (N=156)**	**4 mg baricitinib (N= 234)**
Prior therapy, n (%)[a]	149 (95.5)	144 (92.3)	211 (90.2)
Topical therapy, n (%)[b]	98 (62.8)	97 (62.2)	148 (63.2)
Topical IMT, n (%)	41 (26.3)	31 (19.9)	63 (26.9)
Intralesional therapy, n (%)	88 (56.4)	82 (52.6)	104 (44.4)
Systemic agents, n (%)[c]			
Immunosuppressant/ immunomodulator	97 (62.2)	89 (57.1)	124 (53.0)
Corticosteroids	77 (49.4)	77 (49.4)	102 (43.6)
JAK inhibitor	9 (5.8)	6 (3.8)	10 (4.3)
Others	54 (34.6)	32 (20.5)	52 (22.2)
Cyclosporin	27 (17.3)	17 (10.9)	27 (11.5)
Methotrexate	27 (17.3)	16 (10.3)	31 (13.2)
Other systemic (non-immunosuppressant), n (%)	15 (9.6)	16 (10.3)	18 (7.7)
Phototherapy, n (%)	28 (17.9)	24 (15.4)	37 (15.8)
Procedures, n (%)[d]	35 (22.4)	31 (19.9)	47 (20.1)

Footnotes: [a] Patients may have had >1 prior therapy; therefore, the n in each group may not add up to the total N for the column. [b] Topical therapies excluding IMT. [c] All immunosuppressant and immunomodulator agents including CS, JAK inhibitors, and others. [d] Procedures include cryotherapy, micro-needling, and platelet-rich plasma injections.
Abbreviations: IMT: immunomodulatory therapy; JAK: janus kinase; PBO: placebo.
Source: BRAVE-AA2 Clinical Study Report; EPAR.[66, 68]

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

B.2.4.1 Analysis sets

The analysis sets used in the analysis of the BRAVE-AA1 and -AA2 are presented in Table 13.

Table 13. Trial populations used for the analysis of outcomes in BRAVE-AA1 and BRAVE-AA2

Analysis set		BRAVE-AA1 (phase III portion)	BRAVE-AA2
Full Analysis Set (FAS)	Description	• Comprises all patients randomised at baseline • The efficacy analysis of the primary and key secondary endpoints was conducted using the FAS population	
	N	N=654	N=546
Modified Full Analysis Set (mFAS) Population	Description	• Comprises all patients randomised who received at least one dose of study intervention, excluding patients with female pattern baldness and male patients with diffuse AGA (grade IV and above) identified at Week 36 • Sensitivity analyses were conducted in the mFAS population to evaluate the impact of protocol deviations and differences among analysis populations	
	N	[redacted]	[redacted]
Per-Protocol Set (PPS)	Description	• Comprises all mFAS patients who were deemed compliant with treatment, who do not have any of the important protocol deviations that exclude patients from the PPS, and whose investigator site does not have significant GCP deviations that require a report to regulatory agencies. Protocol deviations were determined before unblinding prior to the primary outcome database lock • The primary endpoint (supplemental analysis) was analysed in the PPS population	
	N	[redacted]	[redacted]
Safety Population	Description	• Comprises all patients randomised who receive at least one dose of study intervention and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first post-baseline visit • All safety analyses were performed on the safety population	
	N	[redacted]	[redacted]
Pooled Week-36 efficacy population	Description	• Comprises all patients enrolled and randomised in the phase III portion of BRAVE-AA1 and BRAVE-AA2 • Primary and key secondary endpoints were analysed in this population	
	[redacted]		

Abbreviations: AGA: androgenetic alopecia; FAS: full analysis set; GCP: good clinical practice; mFAS: modified full analysis set; PPS: per-protocol set.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65-67]

B.2.4.2 Statistical methods

Treatment comparisons of discrete efficacy variables were made using a logistic regression analysis. In the case when logistic regression model did not produce statistical results due to sparse data, Fisher exact test were used. Patients were generally considered non-responders for the non-responder imputation- (NRI-) based analysis if they did not meet clinical response criteria or if they discontinued study or study treatment at any time prior to the time point of interest for any reason.

Treatment comparisons of continuous efficacy and health outcome variables were be made using analysis of covariance (ANCOVA). Type III tests for LSMs were used for statistical comparison between treatment groups

Fisher exact test were used for all adverse events, discontinuation, and other categorical safety data. Continuous vital signs and laboratory values were analysed by an analysis of covariance (ANCOVA) with treatment and baseline values.

B.2.4.3 Censoring

In the phase III BRAVE-AA1 and in BRAVE-AA2 trials, efficacy data were analysed according to a prespecified censoring rule, where data collected after permanent study drug discontinuation or data collected during remote visits due to the COVID-19 pandemic were censored. This applied to all efficacy analysis populations up to Week 52. Missing data were imputed using non-responder imputation (NRI) for categorical endpoints and modified last observation carried forward (mLOCF) for continuous endpoints.

For endpoints tested outside of the graphical scheme, the statistical significance on the categorical variables concluded throughout are all based on the p-value from the logistic regression model (p-value*b) with a p-value*b ≤0.05 deemed significant.

B.2.4.4 Multiplicity

In the phase III portion of BRAVE-AA1 and in BRAVE-AA2, multiplicity-controlled analyses were performed on the primary and key secondary endpoints using a graphical multiple-testing procedure to control the overall family-wise type I error at a two-sided 5% significance level. Adjustments for multiplicity were not performed on other endpoints. The results of the graphical multiple-testing procedures in BRAVE-AA1 and BRAVE-AA2 are shown in Figure 5 and Figure 6 below.

Figure 5. Results of graphical multiple-testing procedure in BRAVE-AA1

Abbreviations: ClinRO, clinician-reported outcome; EB, eyebrow; EL, eyelash; PCFB, percent change from Baseline; PRO, patient-reported outcome; SALT, Severity of Alopecia Tool; $SALT_{50/90}$, at least 50/90% improvement from baseline in SALT score; SH, scalp hair; Wk, week.

✓Achieved statistical significance after adjustment for multiplicity

⦸ Did not achieve statistical significance after adjustment for multiplicity and stopped graphical testing scheme

✗p≤0.05, without adjustment for multiplicity

✗Not statistically significant, p>0.05

Figure 6. Results of graphical multiple-testing procedure in BRAVE-AA2

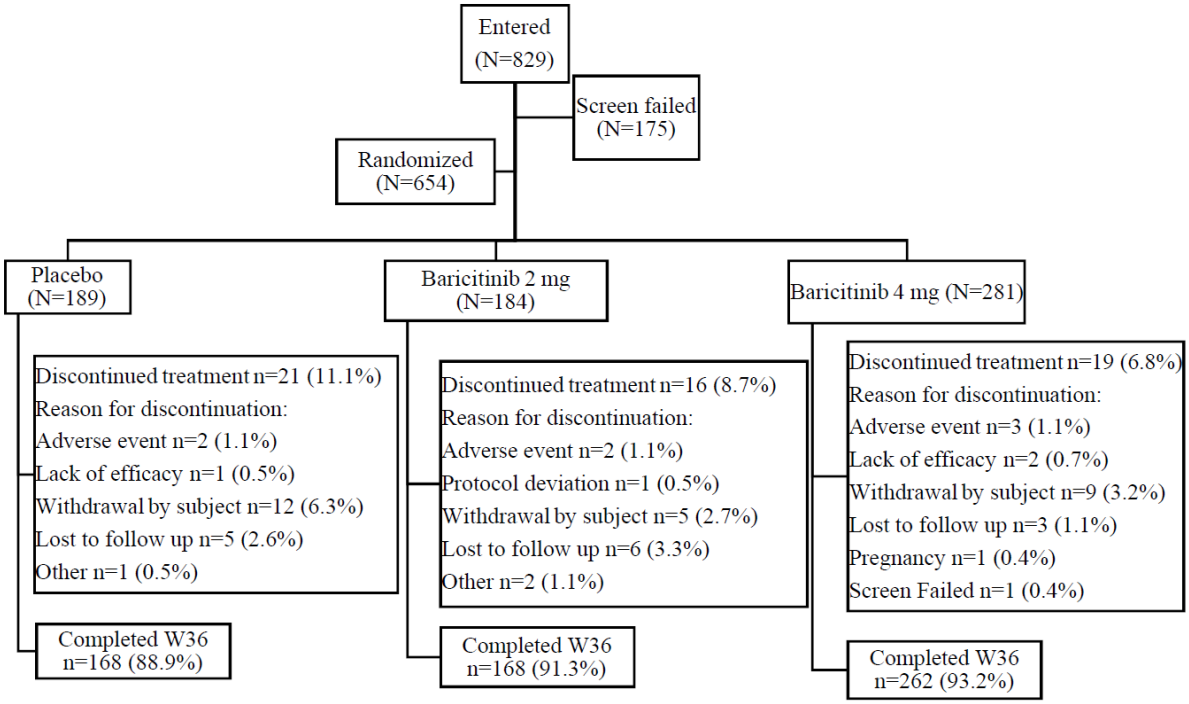
Abbreviations: ClinRO, clinician-reported outcome; EB, eyebrow; EL, eyelash; PCFB, percent change from Baseline; PRO, patient-reported outcome; SALT, Severity of Alopecia Tool; $SALT_{50/90}$, at least 50/90% improvement from baseline in SALT score; SH, scalp hair; Wk, week.
✓Achieved statistical significance after adjustment for multiplicity
⦸ Did not achieve statistical significance after adjustment for multiplicity and stopped graphical testing scheme
✗p≤0.05, without adjustment for multiplicity
✗Not statistically significant, p>0.05

B.2.4.5 Participant disposition

BRAVE-AA1

A total of 829 patients were screened in BRAVE-AA1, of whom 654 were randomised in a 2:2:3 ratio: 189 received placebo, 184 received baricitinib 2 mg, and 281 received baricitinib 4 mg.[67] Overall, study drug discontinuation rates ranged from 6.8% in the baricitinib 4 mg group to 11.1% in the placebo group. The most common reasons for discontinuation were withdrawal in placebo and baricitinib 4 mg groups (6.3% and 3.2%, respectively), and loss to follow-up in the baricitinib 2 mg group (3.3%). Patient disposition for the phase III portion of BRAVE-AA1 through Week 36 is shown in Figure 7.

Figure 7. Patient disposition to Week 36 in BRAVE-AA1

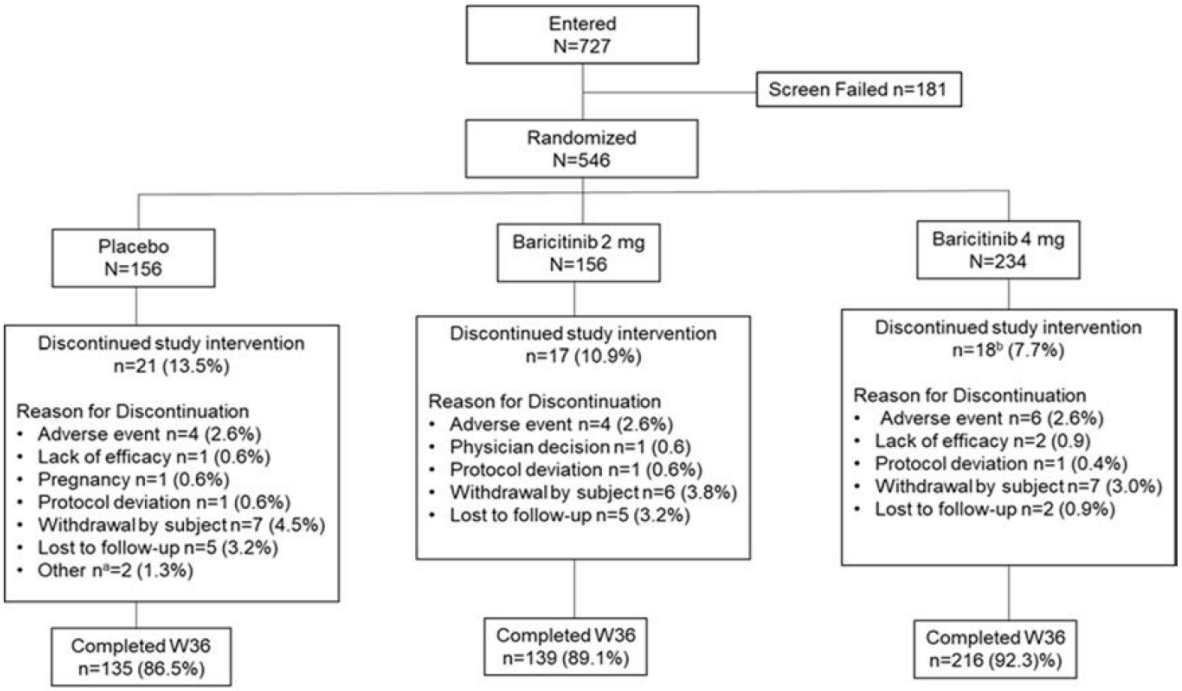


Source: BRAVE-AA1 Clinical Study Report.[65]

BRAVE-AA2

A total of 727 patients were screened in BRAVE-AA2, of whom 546 were randomised in a 2:2:3 ratio: 156 received placebo, 156 received baricitinib 2 mg, and 234 received baricitinib 4 mg.[67] Overall, study drug discontinuation rates ranged from 7.7% in the baricitinib 4 mg group to 13.5% in the placebo group. Across all groups, the most common reason for discontinuation was withdrawal by subject (placebo, 4.5%; baricitinib 2 mg, 3.8%; baricitinib 4 mg, 3.0%). Patient disposition for BRAVE-AA2 through Week 36 is shown in Figure 8.

Figure 8. Patient disposition to Week 36 in BRAVE-AA2



Footnotes: [a] For placebo, the other reasons included lack of adherence to study intervention and screening failure due to criteria 3 but randomised by mistake (this patient never received study intervention).
[b] One patient interrupted study intervention prior to Week 36 and did not resume.
Source: BRAVE-AA2 Clinical Study Report.[66]

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

The trials captured in the clinical SLR were assessed for quality using the Appraisal of RCT checklist by the Cochrane Collaboration,[69] while the quality of all observational studies was assessed using the quality assessment tool developed by the York University CRD.[70] The results of these quality assessments are presented in Appendix D, and a summary of the quality assessment for BRAVE-AA1 and -AA2 is presented in Table 15 below.

Table 14. Critical appraisal of BRAVE-AA1 and BRAVE-AA2

Source of bias	Risk of bias	
	BRAVE-AA1	BRAVE-AA2
Was the method used to generate random allocations adequate?	Low	Low
Was the allocation adequately concealed?	Low	Low
Were the groups similar at the outset of the study in terms of prognostic factors?	Low	Low
Were the care providers, participants, and outcome assessors blind to treatment allocation?	Low	Low
Were there any unexpected imbalances in dropouts between groups?	Low	Low
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Low	Low

Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Low	Low

B.2.6 Clinical effectiveness results of the relevant studies

Summary of clinical effectiveness of baricitinib

- Treatment with baricitinib was associated with a statistically significant hair regrowth on the scalp when compared to placebo. The proportion of patients achieving SALT≤20 at Week 36 in BRAVE-AA1 were 35.2%, 21.7%, and 5.3% for baricitinib 4 mg, 2 mg, or placebo, respectively and 32.5%, 17.3%, and 2.6% in BRAVE-AA2
  - Scalp hair regrowth following treatment with baricitinib 4 mg increased over 36 weeks of treatment, and was associated with a statistically significant higher proportion of patients achieving SALT≤20 at Week 16 and Week 24 as compared with placebo
  - A higher proportion of patients receiving baricitinib 4 mg also achieved SALT≤10 (≥90% scalp hair coverage) compared with placebo at Week 36 in BRAVE-AA1 (26.0% vs 3.7%; p<0.001), and in BRAVE-AA2 (23.5% vs 0.6%; p<0.001), representing almost complete hair regrowth in these patients
- Treatment with baricitinib 4 mg led to a statistically significant eyelash and eyebrow regrowth at Week 36 when compared with placebo treatment, with the proportion of patients achieving an improvement in hair growth increasing over the treatment period
- The SALT≤20 response rate for baricitinib continued to increase from Week 36 through Week 52. Improvements in other secondary endpoints including SALT≤10, $SALT_{50}$ and the ClinRO Measures for EB or EL hair loss also continued to increase from Week 36 through Week 52.
- Among patients who achieved SALT≤20 at Week 52, [redacted] of those who were randomised to remain on baricitinib 4 mg in a down-titration sub-study, and [redacted] of those who were assigned to remain on baricitinib 4 mg in a withdrawal sub-study achieved SALT≤20 at week 76, suggesting that efficacy is maintained over time in most of the patients who have reached SALT≤20. Among the patients who had reached SALT≤20 at week 52, the proportion who had also achieved SALT≤10 increased up to week 76.
- Baricitinib was associated with a notable improvement in HRQoL when compared with placebo. Statistically significant improvements were associated with baricitinib 4 mg treatment after 36 weeks, as measured by the Skindex-16 AA tool across both trials and by HADS-Depression and HADS-Anxiety scores in BRAVE-AA2

The anticipated license dose for baricitinib in AA is 4 mg once daily, although a dose of 2 mg once daily may be appropriate for some patients, or as a down-titration option if patients achieve sustained control of disease with 4 mg once daily. Baricitinib 4 mg will inform the base case. However, for completeness, data for both 4 mg and 2 mg treatment arms will be presented for both trials. Pooled data from BRAVE-AA1 and BRAVE-AA2 will also be presented in Section B.2.8 as these data will be used in the economic model.

B.2.6.1 Primary efficacy endpoint

Treatment with baricitinib is associated with significant improvement in hair regrowth on the scalp, eyebrow and eyelashes at Week 36

Proportion of patients achieving SALT≤20 at Week 36

The SALT score measures the extent of hair loss on the scalp, with higher scores representing a greater extent of hair loss on the scalp. SALT≤20, indicating less than 20% of scalp hair loss (or ≥80% scalp coverage with hair), represents a meaningful outcome for patients with ≥50% scalp hair loss.[71]

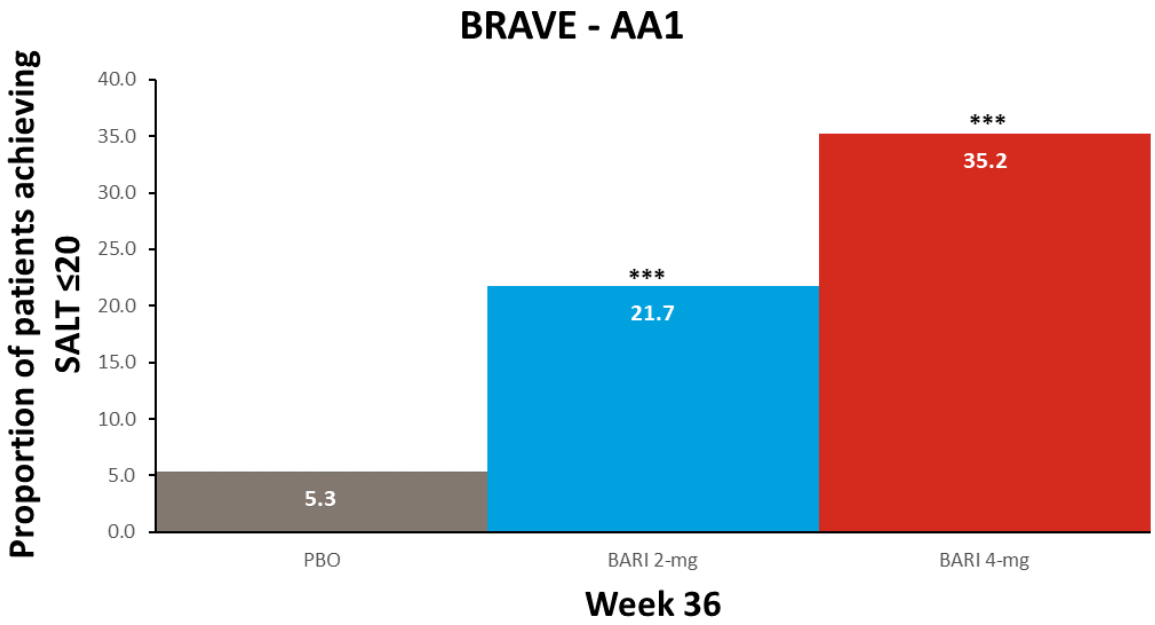
In both BRAVE-AA1 and BRAVE-AA2, treatment with both baricitinib 4 mg and 2 mg resulted in a significantly higher proportion of patients achieving SALT≤20 at Week 36 compared to placebo, after adjusting for multiplicity (Table 15). In BRAVE-AA1, the proportions of patients achieving SALT≤20 at Week 36 were 35.2%, 21.7%, and 5.3% for baricitinib 4 mg, 2 mg, or placebo, respectively (Figure 9). In BRAVE-AA2, 32.5%, 17.3%, and 2.6% of patients who received baricitinib 4 mg, 2 mg, or placebo achieved SALT≤20 at Week 36 (p<0.001 for all comparisons with placebo in both trials) (Figure 10). Across both trials, only 5.3% and 2.6% of patients randomly assigned to placebo, respectively, achieved the primary endpoint of SALT≤20 at 36 weeks. These results confirm published observations that prognosis is particularly poor for patients with chronic extensive hair loss, and that the demonstration of efficacy in BRAVE-AA1 and BRAVE-AA2 was performed in a refractory population of patients. Results for the proportion of patients achieving SALT≤20 at Week 36 according to disease severity are presented in Appendix E.

Table 15. Proportion of patients achieving SALT≤20 at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

Week 36	BRAVE-AA1			BRAVE-AA2		
	PBO	2mg baricitinib	4mg baricitinib	PBO	2mg baricitinib	4mg baricitinib
Response, n (%) (95% CI)[a]	10 (5.3) (2.9, 9.5)	40 (21.7) (16.4, 28.2)	99 (35.2) (29.9, 41.0)	4 (2.6) (1.0, 6.4)	27 (17.3) (12.2, 24.0)	76 (32.5) (26.8, 38.7)
Difference (95% CI) vs PBO[a]	N/A	16.4 (9.7, 23.4)	29.9 (23.2, 36.2)	NA	14.7 (8.3, 21.6)	29.9 (23.1, 36.3)
Odds ratio (95% CI) vs PBO[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	N/A	**<0.001**	**<0.001**	NA	**<0.001**	**<0.001**

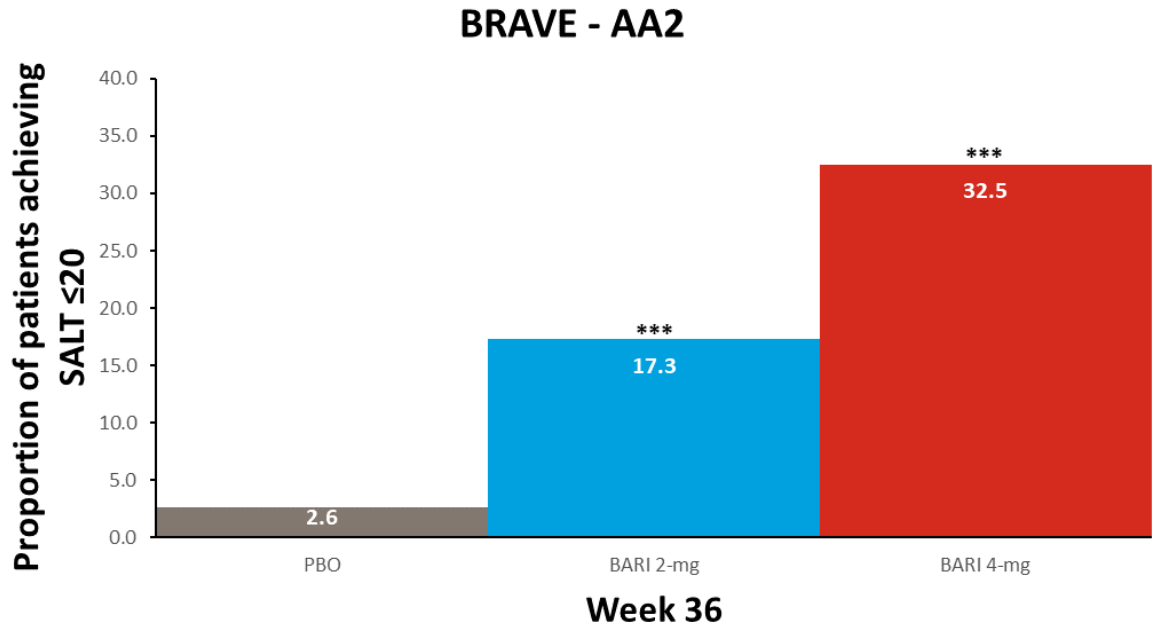
Abbreviations: CI: confidence interval; PBO: placebo; SALT: Severity of Alopecia Areata Tool; NA = not applicable.
Footnotes: Results in bold were statistically significant after adjustment for multiplicity. [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Source: BRAVE-AA1 Clinical Study Report. BRAVE-AA2 Clinical Study Report.[65, 66]

Figure 9. Proportion of patients achieving SALT≤20 at Week 36 in BRAVE-AA1 (FAS population; primary censoring rule [NRI])



Footnotes: ***p<0.001 baricitinib vs placebo.
Abbreviations: AA: alopecia areata; BARI: baricitinib; NRI, non-responder imputation; PBO, placebo; SALT, Severity of Alopecia Tool.
Source: BRAVE-AA1 Clinical Study Report.[65]

Figure 10. Proportion of patients achieving SALT≤20 at Week 36 in BRAVE-AA2 (FAS population; primary censoring rule [NRI])



Footnotes: ***p<0.001 baricitinib vs placebo.
Abbreviations: AA: alopecia areata; BARI: baricitinib; NRI, non-responder imputation; PBO, placebo; SALT, Severity of Alopecia Tool.
Source: BRAVE-AA2 Clinical Study Report.[66]

B.2.6.2 Secondary efficacy endpoints

An overview of results for key secondary endpoints (multiplicity adjusted) from BRAVE-AA1 and BRAVE-AA2 is provided in Table 16 and Table 17, respectively. Data regarding the most relevant endpoints are then described in more detail below.

Table 16. Efficacy results for key secondary endpoints (after adjustment for multiplicity) for BRAVE-AA1 (FAS population)

	PBO (N=189)	2 mg baricitinib (N=184)	p-value vs PBO	4mg baricitinib (N=281)	p-value vs PBO
SALT					
Proportion of patients achieving SALT≤20 at: Week 24, n (% [95% CI]) Week 16, n (%) (95% CI)	 9 (4.8 [2.5, 8.8]) 8 (4.2 [2.2, 8.1])	 **21 (11.4 [7.6, 16.8])** 12 (6.5 [3.8, 11.1])	 **⩽0.05** NS	 **75 (26.7 [21.9, 32.2])** **52 (18.5 [14.4, 23.5])**	 **⩽0.001** **⩽0.001**
Proportion of patients achieving SALT≤10 at: Week 36, n (% [95% CI]) Week 24, n (% [95% CI])	 7 (3.7 [1.8, 7.4]) 5 (2.6 [1.1, 6.0])	 **23 (12.5 [8.5, 18.1])** **14 (7.6 [4.6, 12.4])**	 **⩽0.01** **⩽0.05**	 **73 (26.0 [21.2, 31.4])** **51 (18.1 [14.1, 23.1])**	 **⩽0.001** **⩽0.001**
Mean (SE) change from baseline in SALT score at Week 36	−8.13 (3.10)	**−31.23 (3.16)**	**⩽0.001**	**−45.79 (2.66)**	**⩽0.001**
Proportion of patients achieving $SALT_{90}$ at Week 36, n (% [95% CI])	6 (3.2 [1.5, 6.8])	**21 (11.4 [7.6, 16.8])**	**⩽0.01**	**63 (22.4 [17.9, 27.6])**	**⩽0.001**
Proportion of patients achieving $SALT_{50}$ at Week 12, n (% [95% CI])	9 (4.8 [2.5, 8.8])	18 (9.8 [6.3, 14.9])	⩽0.05	**61 (21.7 [17.3, 26.9])**	**⩽0.001**
PRO Scalp Hair Assessment™					
Proportion of patients with score of 0 or 1 at week 36, n/N[a] (% [95% CI])[b,c]	9/181 (5.0 [2.6, 9.2])	**28/175 (16.0 [11.3, 22.2])**	**⩽0.001**	**91/275 (33.1 [27.8, 38.9])**	**⩽0.001**
ClinRO Measures for eyelash and eyebrow™					
Proportion of patients achieving ClinRO measure for eyebrow hair loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with baseline scores ≥2), n/N (% [95% CI])[a,d]	4/124 (3.2 [1.3, 8.0])	**26/136 (19.1 [13.4, 26.5])**	**⩽0.001**	**59/188 (31.4 [25.2, 38.3])**	**⩽0.001**
Proportion of patients achieving ClinRO measure for eyelash hair loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with baseline scores ≥2), n/N (% [95% CI])[a,d]	3/96 (3.1 [1.1, 8.8])	15/111 (13.5 [8.4, 21.1])	⩽0.05	**56/167 (33.5 [26.8, 41.0])**	**⩽0.001**

Footnotes: Results in **bold** denote statistical significance after adjustment for multiplicity; non-bold results denote significance but not after adjustment for multiplicity.[a] Denominator is number of patients for which data are available and is different from the total patient number in each treatment arm. [b] Patients also achieved ≥2-point improvement from baseline. [c] Only assessed in patients with a score ≥3 at baseline.

Abbreviations: CI: confidence interval; ClinRO: clinician-reported outcome; eyebrow: eyebrow; eyelash: eyelash; NS: not significant; PRO: patient-reported outcome; SALT: Severity of Alopecia Tool; $SALT_{50/90}$: ≥50/90% improvement from baseline in SALT score; SE: standard error.
Source: BRAVE-AA1 Clinical Study Report; EPAR.[65, 67, 68]

Table 17. Efficacy results for key secondary endpoints (after adjustment for multiplicity) for BRAVE-AA2 (FAS population)

	PBO (N=156)	2mg baricitinib (N=156)	p-value vs PBO	4mg baricitinib (N=234)	p-value vs PBO
SALT					
Proportion of patients achieving SALT≤20 at: Week 24, n (% [95% CI]) Week 16, n (%) (95% CI)	 2 (1.3 [0.4, 4.6]) 2 (1.3 [0.4, 4.6])	 17 (10.9 [6.9, 16.8]) 13 (8.3 [4.9, 13.7])	 ⩽0.01 ⩽0.01	 **66 (28.2 [22.8, 34.3])** 41 (17.5 [13.2, 22.9])	 **⩽0.001** ⩽0.001
Proportion of patients achieving SALT≤10 at: Week 36, n (% [95% CI]) Week 24, n (% [95% CI])	 1 (0.6 [0.1, 3.5]) 1 (0.6 [0.1, 3.5])	 17 (10.9 [6.9, 16.8]) 12 (7.7 [4.5, 13.0])	 ⩽0.01 ⩽0.01	 **55 (23.5 [18.5, 29.3])** 44 (18.8 [14.3, 24.3])	 **⩽0.001** ⩽0.001
Mean (SE) change from baseline in SALT score at Week 36	−2.96 (2.72)	−28.21 (2.77)	⩽0.001	**−47.45 (2.23)**	**⩽0.001**
Proportion of patients achieving $SALT_{90}$ at Week 36, n (% [95% CI])	1 (0.6 [0.1, 3.5])	13 (8.3 [4.9, 13.7])	⩽0.01	**50 (21.4 [16.6, 27.1])**	**⩽0.001**
Proportion of patients achieving $SALT_{50}$ at Week 12, n (% [95% CI])	4 (2.6 [1.0, 6.4])	17 (10.9 [6.9, 16.8])	⩽0.01	55 (23.5 [18.5, 29.3])	⩽0.001
PRO Scalp Hair Assessment ™					
Proportion of patients with score of 0 or 1 at Week 36, n/N[a] (% [95% CI])[b,c]	6/151 (4.0 [1.8, 8.4])	**24/149 (16.1 [11.1, 22.8])**	**⩽0.01**	**74/215 (34.4 [28.4, 41.0])**	**⩽0.001**
ClinRO Measures for eyelash and eyebrow ™					
Proportion of patients achieving ClinRO measure for eyebrow hair loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with baseline scores ≥2), n/N (% [95% CI])[a]	5/112 (4.5 [1.9, 10.0])	12/104 (11.5 [6.7, 19.1])	NS	**56/161 (34.8 [27.9, 42.4])**	**⩽0.001**
Proportion of patients achieving ClinRO measure for eyelash hair loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with baseline scores ≥2), n/N (% [95% CI])[a]	5/90 (5.6 [2.4, 12.4])	9/89 (10.1 [5.4, 18.1])	NS	**48/140 (34.3 [26.9, 42.5])**	**⩽0.001**

Footnotes: Results in **bold** denote statistical significance after adjustment for multiplicity; non-bold results denote significance but not after adjustment for multiplicity. [a] Denominator is number of patients for which data are available and is different from the total patient number in each treatment arm. [b] Patients also achieved ≥2-point improvement from baseline. [c] Only assessed in patients with a score ≥3 at baseline.
Abbreviations: CI: confidence interval; ClinRO: clinician-reported outcome; eyebrow: eyebrow; eyelash: eyelash; NS: not significant; PRO: patient-reported outcome; SALT: Severity of Alopecia Tool; $SALT_{50/90}$: ≥50/90% improvement from baseline in SALT score; SE: standard error.
Source: BRAVE-AA2 Clinical Study Report; EPAR.[66-68]

Treatment with baricitinib is associated with significant scalp hair regrowth in responders which increases over 36 weeks of treatment

Proportion of patients achieving SALT≤20 at Weeks 16 and 24

The proportion of patients achieving SALT≤20 at Weeks 16 and 24 across the phase III trials is summarised in Table 18. In BRAVE-AA1, a higher proportion of patients achieving SALT≤20 response was observed in patients treated with baricitinib 4 mg compared to placebo at both Week 16 (18.5% vs 4.2%; p≤0.001) and Week 24 (26.7% vs 4.8%; p≤0.001) after adjustment for multiplicity. A statistically significant increase in the proportion of patients achieving SALT≤20 at Week 24 was also observed in the baricitinib 2 mg group compared to placebo (p≤0.05). In BRAVE-AA2, the proportion of patients receiving baricitinib 4 mg who achieved SALT≤20 was similarly higher compared with placebo at Week 24 (28.2% vs 1.3%; p≤0.001) after adjusting for multiplicity. Statistically significant increases in the proportion of patients achieving SALT≤20 at Week 16 was also observed in both 2 mg (p≤0.01) and 4 mg treatment groups (p≤0.001) (not adjusted for multiplicity).

Overall, the SALT≤20 response rates for baricitinib in each trial continued to increase over the 36 weeks of treatment (Figure 11, Figure 12 and Figure 13). Results for the response rates over time through Week 36 according to disease severity are presented in Appendix E.

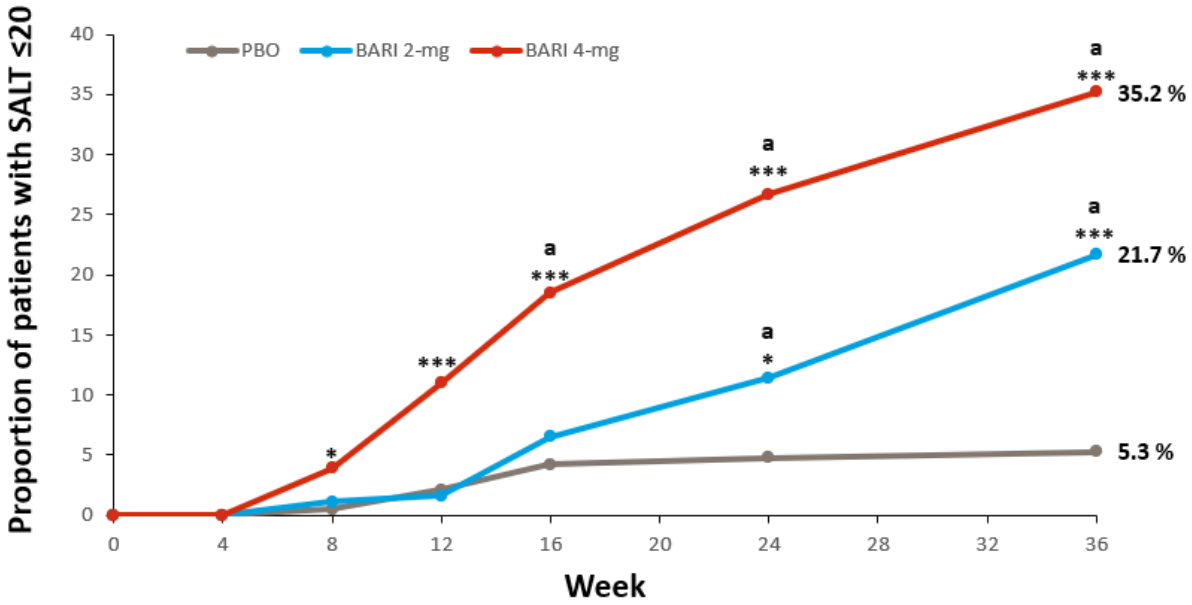
Table 18. Proportion of Patients achieving SALT≤20 at Weeks 16 and 24 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)	PBO (N=189)	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)
Week 16						
Response, n (%) (95% CI)[a]	8 (4.2) (2.2, 8.1)	12 (6.5) (3.8, 11.1)	52 (18.5) (14.4, 23.5)	2 (1.3) (0.4, 4.6)	13 (8.3) (4.9, 13.7)	41 (17.5) (13.2, 22.9)
Difference (95% CI) vs PBO[a]						
Odds ratio (95% CI) vs PBO[b]						
p-Value vs. PBO[b]	NA	0.288	<0.001	NA	0.008	<0.001
Week 24						
Response, n (%) (95% CI)[a]	9 (4.8) (2.5, 8.8)	21 (11.4) (7.6, 16.8)	75 (26.7) (21.9, 32.2)	2 (1.3) (0.4, 4.6)	17 (10.9) (6.9, 16.8)	66 (28.2) (22.8, 34.3)
Difference (95% CI) vs PBO[a]						
Odds ratio (95% CI) vs PBO[b]						

p-Value vs. PBO[b]	NA	0.013	**<0.001**	NA	0.002	**<0.001**

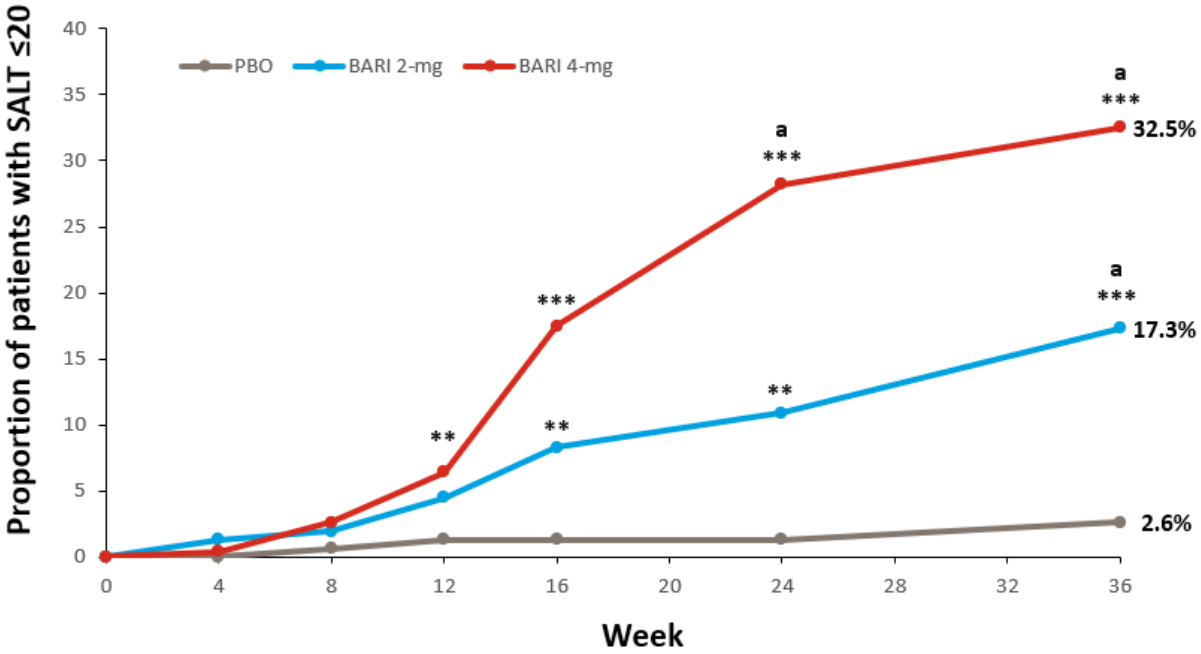
Footnotes: Results in bold were statistically significant after adjustment for multiplicity. [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: SALT: Severity of Alopecia Areata Tool; NA: not applicable.
Source: BRAVE-AA1 Clinical Study Report. BRAVE-AA2 Clinical Study Report.[65-67]

Figure 11. Proportion of patients achieving SALT≤20 over time through Week 36 in BRAVE-AA1 (FAS population; primary censoring rule [NRI])



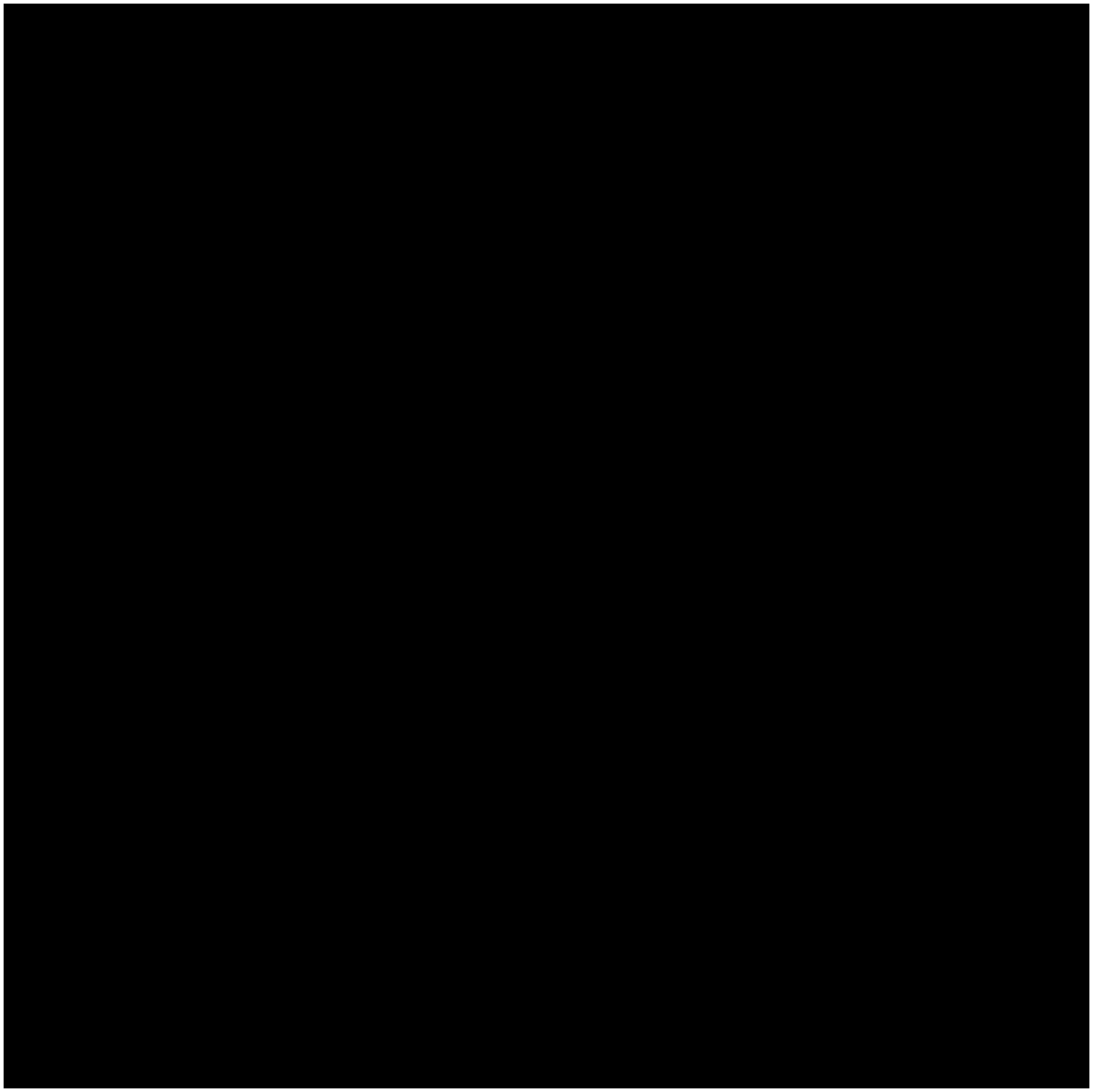
Footnotes: Statistically significant after adjustment for multiplicity (only Weeks 16, 24, and 36 were included in the graphical testing procedure).*p<0.05 vs placebo; **p<0.01 vs placebo; ***p<0.001 vs placebo.
Abbreviations: AA: alopecia areata; BARI: baricitinib; FAS: full analysis set; NRI: non-responder imputation; PBO: placebo; SALT: Severity of Alopecia Tool.
Source: BRAVE-AA1 Clinical Study Report; EPAR.[65, 68]

Figure 12. Proportion of patients achieving SALT≤20 over time through Week 36 in BRAVE-AA2 (FAS population; primary censoring rule [NRI])



Footnotes: Statistically significant after adjustment for multiplicity (only Weeks 16, 24, and 36 were included in the graphical testing procedure).*p<0.05 vs placebo; **p<0.01 vs placebo; ***p<0.001 vs placebo.
Abbreviations: AA: alopecia areata; BARI: baricitinib; FAS: full analysis set; NRI: non-responder imputation; PBO: placebo; SALT: Severity of Alopecia Tool.
Source: BRAVE-AA2 Clinical Study Report; EPAR.[66, 68]

Figure 13. Photographs of changes in absolute SALT scores over time through Week 36 in patients treated with baricitinib in BRAVE-AA1



Abbreviations: BARI: baricitinib; SALT: Severity of Alopecia Tool.

Proportion of patients achieving $SALT_{50}$ at Week 12

Across both trials, both 2 mg and 4 mg baricitinib demonstrated meaningful improvements in scalp hair regrowth before Week 36, with a statically significant increase in the proportion of patients achieving $SALT_{50}$ at Week 12 compared with placebo. In BRAVE-AA1, 21.7% of patients in the baricitinib 4 mg treatment group achieved $SALT_{50}$ at Week 12 compared with only 4.8% of patients in the placebo group ($p<0.001$). Significantly more patients in the 2 mg group also achieved this endpoint compared with placebo at Week 12 ($p<0.05$). Similar results were observed in BRAVE-AA2 with significantly more patients treated with baricitinib 4 mg or 2 mg having $SALT_{50}$ at Week 12 versus placebo (23.5% and 10.9% vs 2.6%; $p<0.001$ [4 mg] and $p<0.01$ [2 mg]).

Table 19. Proportion of Patients Achieving a $SALT_{50}$ at Week 12 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

Week 12	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)	PBO (N=189)	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)
Response, n (%) (95% CI)[a]	9 (4.8) (2.5, 8.8)	18 (9.8) (6.3, 14.9)	61 (21.7) (17.3, 26.9)	4 (2.6) (1.0, 6.4)	17 (10.9) (6.9, 16.8)	55 (23.5) (18.5, 29.3)
Difference (95% CI) vs PBO[a]	N/A	5.0 (-0.3, 10.6)	16.9 (11.0, 22.6)	NA	8.3 (2.8, 14.4)	20.9 (14.7, 27.0)
Odds ratio (95% CI) vs PBO[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	N/A	<0.001	<0.001	NA	0.005	<0.001

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction.
[b] Logistic regression analysis with treatment group, geographic region, duration of current episode at baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: $SALT_{50}$: at least 50% improvement from Baseline in Severity of Alopecia Areata Tool score; NA: not applicable.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Proportion of patients achieving $SALT_{50}$ at Week 36

Across both trials, both 2 mg and 4 mg baricitinib demonstrated meaningful improvements in scalp hair regrowth from baseline at Week 36, with a statically significant increase in the proportion of patients achieving $SALT_{50}$ (at least 50% improvement from baseline) at Week 36 compared with placebo. The proportion of patients achieving $SALT_{50}$ at Week 36 across the phase III trials is summarised in Table 20. In BRAVE-AA1, [redacted] and [redacted] of patients in the baricitinib 4 mg and 2 mg groups, respectively, achieved $SALT_{50}$ at Week 36, compared with [redacted] of placebo-treated patients ([redacted]; not adjusted for multiplicity). In BRAVE-AA2, [redacted] and [redacted] of patients treated with baricitinib 4 mg or 2 mg, respectively, achieved $SALT_{50}$ at Week 36 versus [redacted] of those treated with placebo ([redacted] not adjusted for multiplicity).

Table 20. Proportion of patients achieving an absolute $SALT_{50}$ at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N=281)	PBO (N=156)	2 mg baricitinib (N= 156)	4 mg baricitinib (N=234)
Response, n (%) (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Logistic regression analysis with treatment group, geographic region, duration of current episode at Baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: $SALT_{50}$: at least 50% improvement from Baseline in Severity of Alopecia Areata Tool score; PBO: placebo; NA: not applicable.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Proportion of patients achieving $SALT_{75}$ at Week 36

Baricitinib 2 mg and 4 mg also demonstrated a statically significant increase in the proportion of patients achieving $SALT_{75}$ (at least 75% improvement from baseline) at Week 36 compared with placebo. The proportion of patients achieving $SALT_{75}$ at Week 36 across the phase III trials is summarised in Table 21. In BRAVE-AA1, [redacted] and [redacted] of patients in the baricitinib 4 mg and 2 mg groups, respectively, achieved $SALT_{75}$ at Week 36, compared with [redacted] of placebo-treated patients ([redacted]; not adjusted for multiplicity). In BRAVE-AA2, [redacted] and [redacted] of patients treated with baricitinib 4 mg or 2 mg, respectively, achieved $SALT_{75}$ at week 36 versus [redacted] of those treated with placebo ([redacted] not adjusted for multiplicity).

Table 21. Proportion of patients achieving an absolute $SALT_{75}$ at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N=281)	PBO (N=156)	2 mg baricitinib (N= 156)	4 mg baricitinib (N=234)
Response, n (%) (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO		[redacted]	[redacted]	N/A	[redacted]	[redacted]
Odds ratio (95% CI) vs PBO		[redacted]	[redacted]	N/A	[redacted]	[redacted]
p-Value vs. PBO[a]		[redacted]	[redacted]	N/A	[redacted]	[redacted]

Footnotes: [a] Logistic regression analysis with treatment group, geographic region, duration of current episode at Baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: $SALT_{50}$: at least 50% improvement from Baseline in Severity of Alopecia Areata Tool score; NA: not applicable.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Treatment with baricitinib is associated with almost complete hair regrowth on the scalp by Week 36 in a subset of patients

Proportion of patients achieving an absolute SALT≤10 at Weeks 24 and 36

The proportion of patients achieving SALT≤10 (≥90% scalp coverage with hair) at Weeks 24 and 36 across the phase III trials is summarised in Table 22. Compared with placebo, a higher proportion of patients randomised to receive baricitinib 4 mg achieved SALT≤10 at Weeks 24 ([redacted]) and 36 ([redacted]) in BRAVE-AA1, and at Week 36 in BRAVE-AA2 ([redacted]) after adjustment for multiplicity. In BRAVE-AA1, there was also a higher proportion of patients in the baricitinib 2 mg group that achieved SALT≤10 at Weeks 24 ([redacted]) and 36 ([redacted]) versus placebo after adjustment for multiplicity (Figure 14).

Consistent with SALT ≤20, the SALT≤10 response rates for both doses of baricitinib continued to increase over the 36 weeks of treatment (Figure 14 and Figure 15). In addition, most patients treated with baricitinib 4 mg who had reached SALT≤20 by Week 36 also achieved SALT≤10.

Table 22. Proportion of Patients Achieving an Absolute SALT≤10 at Weeks 24 and 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N=281)	PBO (N=189)	2 mg baricitinib (N= 84)	4 mg baricitinib (N=281)
Week 24						
Response, n (%) (95% CI)[a]						
Difference (95% CI) vs PBO[a]						
Odds ratio (95% CI) vs PBO[b]						
p-Value vs. PBO[b]						
Week 36						
Response, n (%) (95% CI)[a]						
Difference (95% CI) vs PBO[a]						
Odds ratio (95% CI) vs PBO[b]						
p-Value vs. PBO[b]						

Abbreviations: SALT≤10: less than 10% of hair loss observed using the Severity of Alopecia Areata Tool score; NA: not applicable.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Figure 14. Proportion of patients achieving SALT≤10 through Week 36 in BRAVE-AA1 and BRAVE-AA1 (FAS population; primary censoring rule [NRI])

Footnotes: [a] Statistically significant after adjustment for multiplicity (only Weeks 24 and 36 were included in the graphical testing procedure). *p<0.05 vs placebo; **p<0.01 vs placebo; ***p<0.001 vs placebo.
Abbreviations: AA, alopecia areata; BARI, baricitinib; FAS, full analysis set; PBO: placebo; SALT: Severity of Alopecia Tool.
Source: BRAVE-AA1 Clinical Study Report.[65]

Figure 15. Proportion of patients achieving SALT≤10 through Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

Footnotes: [a] Statistically significant after adjustment for multiplicity (only Weeks 24 and 36 were included in the graphical testing procedure). *p<0.05 vs placebo; **p<0.01 vs placebo; ***p<0.001 vs placebo.
Abbreviations: AA, alopecia areata; BARI, baricitinib; FAS, full analysis set; PBO: placebo; SALT: Severity of Alopecia Tool.
Source: BRAVE-AA2 Clinical Study Report.[66]

Treatment with baricitinib is associated with significant eyebrow and eyelash regrowth after 36 weeks

Proportion of patients achieving ClinRO measure for eyebrow hair loss 0 or 1 at Week 36

The ClinRO measure for eyebrow hair loss assesses the extent of eyebrow loss, with higher scores representing greater hair loss. The proportion of patients with a score ≥2 at baseline achieving ClinRO measure for eyebrow hair loss 0 or 1 at Week 36 is summarised in Figure 16 and Figure 17. A higher proportion of patients treated with baricitinib 4 mg achieved a score of 0 or 1 with a ≥2-point improvement from baseline in the ClinRO Measure for eyebrow Hair Loss at Week 36 compared with placebo after adjustment for multiplicity in both BRAVE-AA1(31.4% vs 3.2%; p<0.001) and BRAVE-AA2 (34.8% vs. 4.5%; p<0.001). Baricitinib 2 mg also demonstrated a statistically significant improvement compared with placebo in the proportion of patients achieving ClinRO Measure for eyebrow Hair Loss score of 0 or 1 at Week 36 in BRAVE-AA1 (3.2% vs 19.1%; p<0.001). The proportion of patients treated with baricitinib who achieved a score of 0 or 1, with a ≥2-point improvement from baseline in the ClinRO Measure for eyebrow Hair Loss continued to increase through Week 36.

Figure 16. Proportion of patients achieving a score of 0 or 1 with ≥2-point improvement from baseline on the ClinRO Measure for eyebrow Hair Loss™ through Week 36 in BRAVE-AA1 (FAS population; primary censoring rule [NRI])

Footnotes: Patients with score of ≥2 at baseline. [a] Statistically significant after adjustment for multiplicity (only Week 36 was included in the graphical testing procedure).*p<0.05 vs placebo; **p<0.01 vs placebo; ***p≤0.001 vs placebo.
Abbreviations: BARI: baricitinib; ClinRO: clinician-reported outcome; eyebrow: eyebrow; FAS: full analysis set; PBO: placebo. **Source:** BRAVE-AA1 Clinical Study Report.[65]

Figure 17. Proportion of patients achieving a score of 0 or 1 with ≥2-point improvement from baseline on the ClinRO Measure for eyebrow Hair Loss™ through Week 36 in BRAVE-AA2 (FAS population; primary censoring rule [NRI])

Footnotes: Patients with score of ≥2 at baseline. [a] Statistically significant after adjustment for multiplicity (only week 36 was included in the graphical testing procedure).*p<0.05 vs placebo; **p<0.01 vs placebo; ***p≤0.001 vs placebo.
Abbreviations: BARI: baricitinib; ClinRO: clinician-reported outcome; eyebrow: eyebrow; FAS: full analysis set; PBO: placebo.
Source: BRAVE-AA2 Clinical Study Report.[66]

Proportion of patients achieving ClinRO measure for eyelash hair loss 0 or 1 at Week 36

The ClinRO measure for eyelash hair loss assesses the extent of eyelash hair loss, with higher scores representing greater eyelash hair loss The proportion of patients with score of ≥2 at baseline achieving ClinRO measure for eyelash hair loss 0 or 1 at Week 36 is summarised in Figure 18 and Figure 19. A higher proportion of patients treated with baricitinib 4 mg achieved a score of 0 or 1 with a ≥2-point improvement from baseline in the ClinRO Measure for eyelash Hair Loss at week 36 compared with placebo after adjustment for multiplicity in in BRAVE-AA1 (33.5% vs 3.1%; p<0.001) and BRAVE-AA2 (34.3% vs 5.6%; p<0.001). Baricitinib 2 mg did not achieve a statistically significant improvement compared to placebo in either trial. Nevertheless, similar findings were observed with respect to improvements over time, as the proportion of patients treated with baricitinib who achieved a score of 0 or 1 for eyelash Hair Loss continued to increase through to Week 36.

Figure 18. Proportion of patients achieving a score of 0 or 1 with ≥2-point improvement from baseline on the ClinRO Measure for eyelash Hair Loss™ through Week 36 in BRAVE-AA1 (FAS population; primary censoring rule [NRI])

Footnotes: Patients with a score of ≥2 at baseline [a] Statistically significant after adjustment for multiplicity (only Week 36 was included in the graphical testing procedure).*p<0.05 vs placebo; **p<0.01 vs placebo; ***p≤0.001 vs placebo.
Abbreviations: BARI: baricitinib; ClinRO: clinician-reported outcome; eyelash: eyelash; FAS: full analysis set; PBO: placebo.
Source: BRAVE-AA1 Clinical Study Report. [65]

Figure 19. Proportion of patients achieving a score of 0 or 1 with ≥2-point improvement from baseline on the ClinRO Measure for eyelash Hair Loss™ through Week 36 in BRAVE-AA2 (FAS population; primary censoring rule [NRI])

Footnotes: Patients with a score of ≥2 at baseline.[a] Statistically significant after adjustment for multiplicity (only Week 36 was included in the graphical testing procedure).*p<0.05 vs placebo; **p<0.01 vs placebo; ***p≤0.001 vs placebo.
Abbreviations: BARI: baricitinib; ClinRO: clinician-reported outcome; eyelash: eyelash; FAS: full analysis set; PBO: placebo.
Source: BRAVE-AA2 Clinical Study Report.[66]

B.2.6.3 Health-related quality of life (HRQoL) outcomes

Treatment with baricitinib demonstrates a notable improvement in HRQoL when compared with placebo

Mean change from baseline in Skindex-16 AA domain scores at Week 36

The Skindex-16 adapted for AA score assesses the impact of AA on quality of life by measuring the degree to which the subjects are bothered by AA and associated symptoms. The Skindex-16 scores in patients enrolled in BRAVE-AA1 and BRAVE-AA2 across the 3 domains (emotions, functioning and symptoms) are shown in Table 23. Overall, baricitinib 4 mg and 2 mg demonstrated a statistically significant improvement in mean change from baseline to Week 36 in the Skindex-16 AA Emotions domain compared with placebo in both the BRAVE-AA1 and BRAVE-AA2 trials ([redacted]). In addition, baricitinib 4 mg was associated with a statistically significantly greater improvement in the Functioning domain compared with placebo at Week 36 in both phase III trials (BRAVE-AA1: [redacted]; BRAVE-AA2: [redacted]). Finally, statistically significant differences in the Symptoms domain versus placebo were observed in patients treated with baricitinib 4 mg in BRAVE-AA2 ([redacted]) and 2 mg in BRAVE-AA1 ([redacted]).

Table 23. Mean change from baseline in Skindex-16 AA domain scores (Emotions, Functioning and Symptoms) at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [mLOCF])

Week 36	BRAVE-AA1			BRAVE-AA2		
	PBO (N=119)	2 mg baricitinib (N=108)	4 mg baricitinib (N=171)	PBO (N=156)	2 mg baricitinib (N=156)	4 mg baricitinib (N=234)
Emotions						
Mean (SD) baseline score	[redacted]					
LSM (SE) change from baseline[a]	[redacted]					
p-value vs PBO	[redacted]					
Functioning						
Mean (SD) baseline score	[redacted]					
LSM (SE) change from baseline[a]	[redacted]					
p-value vs PBO	[redacted]					
Symptoms						
Mean (SD) baseline score	[redacted]					
LSM (SE) change from baseline[a]	[redacted]					
p-value vs PBO	[redacted]					

Footnotes: Baseline is defined as the last non-missing assessment recorded on or prior to the date of first study drug administration. If a patient is randomised but does not receive study drug, then the date of randomization is used instead of the first dose date. [a] number of BRAVE-AA1 patients: n=110; baricitinib 2 mg, n=102; baricitinib 4 mg, n=165; number of BRAVE-AA2 patients: placebo, n=146; baricitinib 2 mg, n=147; baricitinib 4 mg, n=227.
Abbreviations: AA: alopecia areata; NA: not applicable; LSM, least squares mean; Skindex-16 AA: Skindex-16 Adapted for Alopecia areata; SD: standard deviation; SE: standard error.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Skindex-16 AA domain scores at Week 36: Responder analyses

A separate analysis of the Skindex-16 AA scores by $SALT_{50}$ responder status at Week 36 (responders versus non-responders) was conducted. The Skindex-16 scores in patients enrolled in the BRAVE-AA studies (pooled Week 36 efficacy population) across the 3 domains (emotions, functioning and symptoms) for responders and non-responders are shown in Table 24.

Overall, baricitinib 4 mg and 2 mg demonstrated a statistically significant improvement in mean change from baseline for responders versus non-responders in the Emotions domain at Week 36 in the pooled efficacy analysis ([redacted]). Statistically significant improvements in mean change for responders versus non-responders for baricitinib 4 mg and 2 mg were also observed for the Functioning domain ([redacted] and [redacted], respectively). For the Symptoms domain, a numerical improvement in mean change for responders versus non-responders was observed, however this difference was not statistically significant. These data demonstrate that achieving a $SALT_{50}$ response at Week 36 is associated with notable improvement in Skindex-AA scores, particularly in the Emotions and Functions domains that are likely most affected by AA.

Table 24: Mean change in Skindex-16 AA domain scores (Emotions, Functioning and Symptoms) at Week 36 in the BRAVE-AA studies for $SALT_{50}$ Responders versus Non-responders (pooled Week 36 efficacy population; primary censoring rule [mLOCF])

Week 36	Responders				Non-responders			
	PBO (N=32)	2 mg baricitinib (N=100)	4 mg baricitinib (N=240)	Total (N=372)	PBO (N=313)	2 mg baricitinib (N=240)	4 mg baricitinib (N=275)	Total (N=828)
Emotions								
Mean (SD) baseline score	[redacted]				[redacted]			
LSM (SE) difference versus non-responders	[redacted]							
p-value vs non-responders								
Functioning								
Mean (SD) baseline score	[redacted]			[redacted]	[redacted]			[redacted]
LSM (SE) difference versus non-responders	[redacted]							
p-value vs non-responders								
Symptoms								
Mean (SD) baseline score	[redacted]				[redacted]			
LSM (SE) difference versus non-responders	[redacted]							
p-value vs non-responders								

Footnotes: Baseline is defined as the last non-missing assessment recorded on or prior to the date of first study drug administration. If a patient is randomised but does not receive study drug, then the date of randomization is used instead of the first dose date.
Abbreviations: AA: alopecia areata; NA: not applicable; LSM, least squares mean; Skindex-16 AA: Skindex-16 Adapted for Alopecia areata; SD: standard deviation; SE: standard error.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Mean change from baseline in HADS-Anxiety and HADS-Depression scores at Week 36

The HADS is a 14 item self-assessment scale that determines the levels of anxiety and depression that a patient is experiencing using a 4-point Likert scale. Mean change in HADS-Anxiety and HADS-Depression scores from baseline in participants from BRAVE-AA1 and BRAVE-AA2 are shown in Table 25. In BRAVE-AA1, both baricitinib 2 mg and 4 mg were associated with statistically significant improvements in HADS-A scores compared with placebo at Week 36 ([redacted] respectively). In BRAVE-AA2, patients treated with 4 mg baricitinib experienced a significant improvement in HADS-Anxiety scores at Week 36 ([redacted]). These patients also had statistically significant improvements in HADS-Depression at Week 36 compared to placebo ([redacted]). Of note, patients with significant uncontrolled neuropsychiatric disorder or who were clinically judged by the investigator to be at risk of suicide were excluded from the study.

Table 25. Mean change from Baseline in HADS-A and HADS-D Total Scores at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [mLOCF])

Week 36	BRAVE-AA1			BRAVE-AA2		
	PBO (N = 189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N = 281)	PBO (N = 156)	2 mg baricitinib (N = 156)	4 mg baricitinib (N = 234)
HADS Anxiety						
Mean (SD) baseline score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LSM (SE)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS Depression						
Mean (SD) baseline score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LSM (SE)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a] BRAVE-AA1 patient, n: placebo, n=177; baricitinib 2 mg, n=172; baricitinib 4 mg, n=272. BRAVE-AA2 patient, n: placebo, n=146; baricitinib 2 mg, n=147; baricitinib 4 mg, n=227.
Abbreviations: CSR: clinical study report; FAS: full analysis set; HADS: Hospital Anxiety Depression Scale-Anxiety; LSM: least squares mean; mLOCF: modified last observation carried forward; SD: standard deviation; SE: standard error.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Mean change from baseline in EQ-5D-5L: Health State Index at Week 36

The EQ-5D-5L Health State Index Score measures self-rated patient health status, with lower scores indicating worse disease state. EQ-5D-5L scores for BRAVE-AA1 and BRAVE-AA2 at Week 36 are summarised in Table 26. In BRAVE-AA1, neither baricitinib groups demonstrated a statistically significant improvement or reductions in EQ-5D-5L scores compared to placebo at Week 36. In contrast, statistically significant improvements in EQ-5D-5L scores were observed in BRAVE-AA2 among patients receiving 4 mg baricitinib at Week 36 ([redacted]).

Table 26. Mean change from baseline in EQ-5D-5L health state index (UK algorithm) scores at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [mLOCF])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N = 189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N = 281)	PBO (N = 156)	2 mg baricitinib (N = 156)	4 mg baricitinib (N = 234)
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 36						
LSM (SE)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Baseline is defined as the last non-missing assessment recorded on or prior to the date of first study drug administration. If a patient is randomised but does not receive study drug, then the date of randomization is used instead of the first dose date.
Abbreviations: AA: alopecia areata; NA: not applicable; LSM, least squares mean; PBO: placebo; SE: standard error; EQ-5D-5L: European Quality of Life - 5 Dimensions-5 Levels.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Mean change from baseline in EQ-5D-5L VAS score at Week 36

The EQ-5D-5L VAS score uses a visual analogue scale (VAS) to measure self-rated patient health status, with lower scores indicating worse disease state. EQ-5D-5L VAS scores for participants in BRAVE-AA1 and BRAVE-AA2 at Week 36 are summarised in Table 27. Consistent with the EQ-5D-5L Health State Index results, there were no significant improvements or reductions in EQ-5D-5L VAS scores in BRAVE-AA1 among either group receiving baricitinib at Week 36. In BRAVE-AA2, no statistically significant improvements or reductions in EQ-5D-5L VAS scores were observed across either baricitinib group compared to placebo at Week 36.

Table 27. Mean change from baseline in EQ-5D-5L VAS scores at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [mLOCF])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)	PBO (N=156)	2 mg baricitinib (N=156)	4 mg baricitinib (N=234)
Baseline mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 36						
LSM (SE)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Baseline is defined as the last non-missing assessment recorded on or prior to the date of first study drug administration. If a patient is randomised but does not receive study drug, then the date of randomization is used instead of the first dose date.
Abbreviations: AA: alopecia areata; CI: confidence interval. NA: not applicable; PBO: placebo; LSM, least squares mean; SE: standard error; EQ-5D-5L= European Quality of Life - 5 Dimensions-5 Levels; VAS: visual analogue scale.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

While EQ-5D-5L data were collected as part of BRAVE-AA1 and BRAVE-AA2, HRQoL outcomes in this indication are not expected to be adequately captured by the EQ-5D-5L instrument, with implications for the derivation of utility values for use in the economic analysis.

This arises because AA is characterised by non-scarring hair loss that, unlike some other dermatological conditions, does not usually cause physical symptoms (beyond hair loss) or disability.[1, 72] The impact of AA on HRQoL is instead attributed to the significant psychological distress caused by hair loss.[2, 11, 13, 47] Owing to this mono-symptomatic aspect of AA, the five dimensions of health covered by the generic EQ-5D instrument, comprised of mobility, self-care, usual activities, pain/discomfort and anxiety/depression domains, do not adequately capture the dimensions of HRQoL that are affected by AA (in this case the psychological aspects), demonstrating a lack of content validity for the EQ-5D instrument in AA.[73, 74] Thus, even a post hoc responder analysis of EQ-5D may not differ significantly between responders and non-responders, despite the obvious health benefits that are gained due to hair regrowth in those that respond to baricitinib treatment. Furthermore, this lack of content validity for EQ-5D simultaneously contributes towards a significant ceiling effect in the trial EQ-5D data, whereby many patients at baseline report almost “perfect health” on the EQ-5D instrument and therefore are unable to report an improvement from treatment in a responder analysis, despite entering the trial with severe AA (>50% scalp hair loss). Similar limitations have been reported from a recent trial funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme to treat vitiligo.[75] The observed ceiling effect may also be partly due to the exclusion of patients with significant uncontrolled neuropsychiatric disorders from the BRAVE-AA studies, for whom the HRQoL impact of AA may have been most pronounced.

Mean change from baseline in SF-36 physical component score (PCS) at Week 36

The SF-36 PCS is a patient-completed measure designed to assess the impact of a disease on various physical aspects of health, with higher scores indicating better levels of function and/or better health. PCS at Week 36 for participants in BRAVE-AA1 and BRAVE-AA2 are summarised in Table 28. Neither baricitinib 2 mg nor 4 mg demonstrated a statistically significant improvement compared to placebo for the mean change in PCS from baseline at Week 36 in either phase III trial, though there were also no reductions in PCS due to baricitinib treatment in either trial.

Table 28. Mean change from baseline in SF-36 PCS at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [mLOCF])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N = 189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N = 281)	PBO (N = 156)	2 mg baricitinib (N = 156)	4 mg baricitinib (N = 234)
Baseline Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 36						
LSM (SE)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Baseline is defined as the last non-missing assessment recorded on or prior to the date of first study drug administration. If a patient is randomised but does not receive study drug, then the date of randomization is used instead of the first dose date.
Abbreviations: AA: alopecia areata; CI: confidence interval; NA: not applicable; LSM, least squares mean; SE: standard error; PCS: physical component score.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Mean change from baseline in SF-36 mental component score (MCS) at Week 36

The SF-36 MCS is a patient-completed measure designed to assess the impact of a disease on mental health, with higher scores indicating better mental health. Mental Component Scores at Week 36 for participants in BRAVE-AA1 and BRAVE-AA2 are summarised in Table 29. In BRAVE-AA1, neither baricitinib 2 mg nor 4 mg demonstrated a statistically significant improvement compared to placebo in MCS from Baseline at Week 36. However, patients randomised to baricitinib 4 mg in BRAVE-AA2 achieved significant improvements compared with placebo in MCS scores at Week 36 ([redacted]).

Table 29. Mean change from baseline in SF-36 MCS at Week36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [mLOCF])

	BRAVE-AA1			BRAVE-AA2		
	PBO (N = 189)	2 mg baricitinib (N = 184)	4 mg baricitinib (N = 281)	PBO (N = 156)	2 mg baricitinib (N = 156)	4 mg baricitinib (N = 234)
Baseline Mean	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 36						
LSM (SE)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Baseline is defined as the last non-missing assessment recorded on or prior to the date of first study drug administration. If a patient is randomised but does not receive study drug, then the date of randomization is used instead of the first dose date.
Abbreviations: AA: alopecia areata; CI: confidence interval; NA: not applicable; LSM, least squares mean; SE: standard error; MCS: mental component score.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

B.2.7 Subgroup analyses

Prespecified subgroup analyses were conducted on the BRAVE-AA1 and BRAVE-AA2 study populations for the primary endpoint (SALT ≤20). The proportion of patients achieving SALT≤20 at Week 36 for subgroups in BRAVE-AA1 and BRAVE-AA2 are presented in Table 30 and Table 31, respectively. All statistical tests of treatment-by-subgroup interaction were tested at the 0.1 significance level. If the p-value of subgroup-by-treatment interaction term was <0.1, further diagnostics were performed to describe the nature of the interaction. If any group within a subgroup comprised <10% of the overall sample size, only descriptive summary statistics were provided for treatment groups, and no treatment group comparisons were performed within these subgroup levels. No significant differences in response were observed in BRAVE-AA1 or BRAVE-AA2.

Table 30. Proportion of patients in BRAVE-AA1 achieving SALT≤20 at Week 36 (FAS population; primary censoring rule [NRI])

Subgroup	Category	Response at Week 36, n (%)			p-value
		PBO (N=189)	2 mg baricitinib (N=184)	4 mg baricitinib (N=281)	
BRAVE-AA1					
Gender	Male (N=271)	[redacted]	[redacted]	[redacted]	[redacted]
	Female (N=383)	[redacted]	[redacted]	[redacted]	
Age	<40 (N=388)	[redacted]	[redacted]	[redacted]	[redacted]
	≥40 (N=265)	[redacted]	[redacted]	[redacted]	
	<65 (N=638)	[redacted]	[redacted]	[redacted]	
	≥65 (N=15)	[redacted]	[redacted]	[redacted]	
Baseline weight	<60 (N=144)	[redacted]	[redacted]	[redacted]	[redacted]
	≥60 to <100 (N=449)	[redacted]	[redacted]	[redacted]	
	≥100 (N=61)	[redacted]	[redacted]	[redacted]	
Baseline BMI	<25 (N=326)	[redacted]	[redacted]	[redacted]	[redacted]
	≥25 to <30 (N=190)	[redacted]	[redacted]	[redacted]	
	≥35 (N=137)	[redacted]	[redacted]	[redacted]	
Renal function	Impaired (N=4)	[redacted]	[redacted]	[redacted]	[redacted]
	Not impaired (N=650)	[redacted]	[redacted]	[redacted]	
Race	American Indian or Alaska Native (N=21)	[redacted]	[redacted]	[redacted]	[redacted]
	Asian (N=268)	[redacted]	[redacted]	[redacted]	
	Black or African American (N=52)	[redacted]	[redacted]	[redacted]	
	Native Hawaiian or Other Pacific Islander (N=3)	[redacted]	[redacted]	[redacted]	
	White (N=299)	[redacted]	[redacted]	[redacted]	

	Multiple (N=8)	[redacted]	[redacted]	[redacted]	
Geographic region	North America (N=358)	[redacted]	[redacted]	[redacted]	[redacted]
	Asia (N=247)	[redacted]	[redacted]	[redacted]	
	Rest of world (N=49)	[redacted]	[redacted]	[redacted]	
Duration of current episode of AA category (years)	<4 (N=450)	[redacted]	[redacted]	[redacted]	[redacted]
	≥4 (N=204)	[redacted]	[redacted]	[redacted]	
Baseline SALT category	Severe (N=302)	[redacted]	[redacted]	[redacted]	[redacted]
	Very severe (N=352)	[redacted]	[redacted]	[redacted]	

Abbreviations: AA, alopecia areata; BMI, body mass index; PBO, placebo; SALT, Severity of Alopecia Tool.
Source: BRAVE-AA1 Clinical Study Report.[65]

Table 31. Proportion of patients in BRAVE-AA2 achieving SALT≤20 at Week 36 (FAS population; primary censoring rule [NRI])

Subgroup	Category	Response at Week 36, n (%)			p-value
		PBO (N=156)	2 mg baricitinib (N=156)	4 mg baricitinib (N=234)	
BRAVE-AA2					
Gender	Male (N=201)	[redacted]	[redacted]	[redacted]	[redacted]
	Female (N=345)	[redacted]	[redacted]	[redacted]	
Age	<40 (N=304)	[redacted]	[redacted]	[redacted]	[redacted]
	≥40 (N=241)	[redacted]	[redacted]	[redacted]	
	<65 (N=531)	[redacted]	[redacted]	[redacted]	[redacted]
	≥65 (N=14)	[redacted]	[redacted]	[redacted]	
Baseline weight	<60 (N=117)	[redacted]	[redacted]	[redacted]	[redacted]
	≥60 to <100 (N=390)	[redacted]	[redacted]	[redacted]	
	≥100 (N=39)	[redacted]	[redacted]	[redacted]	
Baseline BMI	<25 (N=263)	[redacted]	[redacted]	[redacted]	[redacted]
	≥25 to <30 (N=169)	[redacted]	[redacted]	[redacted]	

	≥35 (N=114)	[redacted]	[redacted]	[redacted]	
Renal function	Impaired (N=2)	[redacted]	[redacted]	[redacted]	[redacted]
	Not impaired (N=543)	[redacted]	[redacted]	[redacted]	
Race	Asian (N=167)	[redacted]	[redacted]	[redacted]	[redacted]
	Black or African American (N=46)	[redacted]	[redacted]	[redacted]	
	Native Hawaiian or Other Pacific Islander (N=1)	[redacted]	[redacted]	[redacted]	
	White (N=321)	[redacted]	[redacted]	[redacted]	
	Multiple (N=11)	[redacted]	[redacted]	[redacted]	
Geographic region	North America (N=190)	[redacted]	[redacted]	[redacted]	[redacted]
	Asia (N=147)	[redacted]	[redacted]	[redacted]	
	Rest of world (N=209)	[redacted]	[redacted]	[redacted]	
Duration of current episode of AA category (years)	<4 (N=337)	[redacted]	[redacted]	[redacted]	[redacted]
	≥4 (N=209)	[redacted]	[redacted]	[redacted]	
Baseline SALT category	Severe (N=259)	[redacted]	[redacted]	[redacted]	[redacted]
	Very severe (N=286)	[redacted]	[redacted]	[redacted]	

Abbreviations: AA, alopecia areata; BMI, body mass index; PBO, placebo; SALT, Severity of Alopecia Tool.
Source: BRAVE-AA1 Clinical Study Report.[65]

B.2.8 Meta-analysis

BRAVE-AA1 and BRAVE-AA2 are the only two trials identified evaluating baricitinib in this setting and due to the low number of studies, no formal meta-analysis has been conducted. As the trial designs of the phase III portion of BRAVE-AA1 and of BRAVE-AA2 were identical, efficacy data for both trials at Week 36 have been pooled for use in the economic model (see Section B.3.3.2).

B.2.8.1 Proportion of patients achieving $SALT_{30}$, $SALT_{50}$, and $SALT_{75}$ at Week 36

Data from patients in BRAVE-AA1 and BRAVE-AA2 with a baseline SALT score of SALT50–100 were pooled as described above and stratified into SALT response categories ($SALT_{30}$, $SALT_{50}$, and $SALT_{75}$) at Week 36 (Table 32). The $SALT_{50}$ and $SALT_{75}$ data are used in the economic model base case.

Table 32. Proportion of patients responding to treatment at Week 36 in the BRAVE-AA studies (pooled Week 36 efficacy population)

Intervention	$SALT_{30}$		$SALT_{50}$		$SALT_{75}$	
	Efficacy	SE	Efficacy	SE	Efficacy	SE
Baricitinib 2mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinib 4mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SALT: Severity of Alopecia Tool; SE: standard error.

B.2.8.2 Proportion of patients achieving SALT≤20 at Week 52

The response rate to baricitinib treatment continues to increase from Week 36 through Week 52

The proportion of patients with SALT≤20 continued to increase in both the baricitinib 4 mg and 2 mg treatment groups from week 36 through Week 52 in the pooled population (Figure 20). At Week 52, [redacted] and [redacted] of patients receiving baricitinib 4 mg or 2 mg, respectively, achieved this endpoint (improved from [redacted] and [redacted] at Week 36, respectively).

Figure 20. Proportion of patients achieving SALT≤20 through Week 52 in the BRAVE-AA studies (pooled Week 52 efficacy population; primary censoring [NRI])

Abbreviations: BARI: baricitinib; NRI: non-responder imputation; SALT: Severity of Alopecia Tool.

Proportion of patients achieving SALT≤10 at Week 52

The proportion of patients with SALT≤10 continued to increase in both the baricitinib 4 mg and 2 mg treatment groups from week 36 through Week 52 in the pooled population Figure 21). At Week 52, [redacted] and [redacted] of patients receiving baricitinib 4 mg or 2 mg, respectively, achieved SALT≤10 (improved from [redacted] and [redacted] at Week 36, respectively).

Figure 21. Proportion of patients achieving SALT≤10 through Week 52 in the BRAVE-AA studies (pooled Week 52 efficacy population; primary censoring [NRI])

Abbreviations: BARI: baricitinib; NRI: non-responder imputation; SALT: Severity of Alopecia Tool.

Proportion of patients achieving $SALT_{50}$ at Week 52

The proportion of patients achieving $SALT_{50}$ continued to increase in both the baricitinib 4 mg and 2 mg treatment groups from Week 36 through Week 52 in the pooled population (Figure 22). The proportion of patients achieving $SALT_{50}$ at Week 52 was [redacted] and [redacted] in the baricitinib 4 mg and 2 mg groups, respectively.

Figure 22. Proportion of patients achieving $SALT_{50}$ through Week 52 in the BRAVE-AA studies (pooled Week 52 efficacy population; primary censoring [NRI])

Abbreviations: BARI: baricitinib; NRI: non-responder imputation; SALT: Severity of Alopecia Tool.

The efficacy of baricitinib is maintained over time from Week 52 to Week 76 among patients achieving SALT≤20

Proportion of patients achieving SALT≤20 at Week 76: withdrawal sub-study (BRAVE-AA1)

In the withdrawal sub-study of BRAVE-AA1, which comprised a small number of patients who achieved adequate clinical response on baricitinib 4 mg at week 52 (SALT≤20), response was reduced by week 76 in those who were transitioned to placebo but was retained in those who remained on baricitinib 4 mg (Figure 23), though it should be noted that these data do not exactly align with the expected stopping rule in clinical practice. At week 76, [redacted] of patients randomised to baricitinib 4 mg achieved SALT≤20 compared with [redacted] of individuals randomised to placebo. Similar results were observed among patients who achieved adequate clinical response on baricitinib 2 mg at week 52 (SALT≤20) (Figure 23): [redacted] of patients randomised to baricitinib 2 mg achieved SALT≤20 at week 76 compared with [redacted] of those randomised to placebo.

Among the [redacted] patients who were initially treated with baricitinib 4 mg and then re-randomised to placebo, [redacted] had reached SALT≤20 before/at week 36; [redacted] of these individuals (n=[redacted]) maintained SALT≤20 at week 76. [redacted] had reached SALT≤20 between weeks 40 and 52; only [redacted] of these individuals (n=[redacted]) maintained SALT≤20 at week 76. Among the [redacted] patients who were initially treated with baricitinib 2 mg and then re-randomised to placebo, [redacted] had reached SALT≤20 before/at week 36, with [redacted] of these individuals (n=[redacted]) maintaining SALT≤20 at week 76. [redacted] of the patients who were initially treated with baricitinib 2 mg and then re-randomised to placebo reached SALT≤20 between weeks 40 and 52 and [redacted] of these individuals (n=[redacted]) maintained SALT≤20 at week 76.

Frequency of relapses were comparable between patients who had reached SALT≤20 before or at week 36 compared with those who had reached response after week 36 (up to week 52).

Figure 23. Proportion of patients with SALT≤20 response from week 52 through week 76 (randomised withdrawal population; tertiary censoring rule [NRI])

Footnotes: Data censored after permanent study drug discontinuation, retreatment with the original dose of baricitinib or data collected during remote visits due to the COVID-19 pandemic.
Abbreviations: BARI, baricitinib; NRI, non-responder imputation; SALT, Severity of Alopecia Tool.

Proportion of patients achieving SALT≤20 at Week 76: down-titration sub-study (BRAVE-AA2)

In the down-titration sub-study of BRAVE-AA2, [redacted] of the patients who achieved SALT≤20 at Week 52 and who were down-titrated to baricitinib 2 mg achieved SALT≤20 at Week 76. Of the patients who remained on baricitinib 4 mg, [redacted] achieved SALT≤20 at Week 76. Among the [redacted] patients who were down-titrated to baricitinib 2 mg, [redacted] patients reached SALT≤20 before/at Week 36; of these, [redacted] (n=[redacted]) had a SALT≤20 at Week 76. [redacted] patients reached SALT≤20 for the first time between Weeks 40 and 52; of these, [redacted] (n=[redacted]) achieved SALT≤20 at Week 76 (Figure 24).

Figure 24. Maintenance of efficacy through week 76 following down-titration at week 52 (baricitinib 4 mg week 52 re-randomised responder population [BRAVE-AA2]; tertiary censoring rule [NRI])

Footnotes: Data censored after permanent study drug discontinuation, retreatment with the original dose of baricitinib or data collected during remote visits due to the COVID-19 pandemic.
Abbreviations: BARI: baricitinib; NRI: non-responder imputation; SALT: Severity of Alopecia Tool.

Proportion of patients achieving SALT≤10 at Week 76: BRAVE-AA1 and BRAVE-AA2

At week 52, [redacted] of the [redacted] patients who participated in the randomised down-titration sub-study (BRAVE-AA2) or the withdrawal sub-study (BRAVE-AA1) and remained on baricitinib 4 mg had achieved a SALT≤10 (Table 33). In this subgroup of patients, SALT≤10 response rates continued to increase over time reaching [redacted] at Week 76.

Table 33. Response rates at week 52 and week 76 among patients who remained on baricitinib 4 mg in BRAVE-AA1 and BRAVE-AA2 (re-randomised responder population; tertiary censoring rule [NRI])

	Baricitinib 4 mg (BRAVE-AA1/BRAVE-AA2) (N=129)	
	Week 52	Week 76
Proportion of patients achieving SALT ≤20, n (%) [95% CI]	[redacted]	[redacted]
Proportion of patients achieving SALT≤10, n (%) [95% CI]	[redacted]	[redacted]

Footnotes: Tertiary censoring rule: data after permanent study drug discontinuation, after treatment, or collected at remove visits were excluded.
Abbreviations: CI: confidence interval; NRI: non-responder imputation; SALT: Severity of Alopecia Tool.

B.2.9 Indirect and mixed treatment comparisons

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 safety and efficacy of baricitinib and potential comparators for the treatment of adults with severe AA, that identified a total of 47 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 range of comparators than are relevant in UK clinical practice. Of the 47 studies included and extracted in the SLR, 22 were not included in the feasibility assessed as the treatments studied were not considered part of established clinical management, or due to reported endpoints not being comparable. The 25 remaining studies were included in the feasibility assessment, and comprised 4 were placebo-controlled RCTs, 8 non placebo-controlled RCTs and 13 observational studies (Table 34).

Table 34. Studies included in the feasibility assessment (n=25)

Study	Treatment	Dose	Treatment route and class	Total N	Study design
Placebo-controlled RCTs (n=4)					
BRAVE-AA1 (NCT03570749)[67]	Placebo	-	-	189	Double-blind RCT (phase II/III)
	Baricitinib	2mg/day	Oral JAK inhibitor	184	
	Baricitinib	4mg/day	Oral JAK inhibitor	281	
BRAVE-AA2 (NCT03899259)[67]	Placebo	-	-	156	Double-blind RCT (phase III)
	Baricitinib	2mg/day	Oral JAK inhibitor	156	
	Baricitinib	4mg/day	Oral JAK inhibitor	234	
Lai 201916 (ACTRN12618001084279)[76]	Placebo	4mg/kg/day	-	18	Double-blind RCT
	Cyclosporine		Oral IS	18	
Kar 2005[60]	Placebo	200mg/week	-	20	Double-blind RCT
	Prednisolone		Oral Steroid	23	
Non-placebo-controlled RCTs (n=8)					
Asilian 2021 (IRCT20181226042136N1)[77]	Methotrexate	15mg/week	Oral IS	12	Double-blind RCT
	Betamethasone	3mg/week	Oral Steroid	12	
	Both	As above	Combination	12	
Ghandi 2021[78]	DPCP alone	DPCP: 1x week (in 0.01-1% increases until prespecified tolerance)	Topical IT	25	Open-label RCT
	DPCP + anthralin	DPCP: 1x week (in 0.01-1% increases		25	

		until prespecified tolerance) Anthralin: 0.5% concentration 5*week (days 3-7)			
Kurosawa 2006[79]	Dexamethasone	0.5mg/day	Oral Steroids	19	Open-label RCT
	Triamcinolone acetonide	40mg/month		43	
	Prednisolone	80mg for 3 days every 3 months		29	
Thuangtong 2017[80]	DPCP standard	0.0001%/week	Topical IT	20*	Open-label RCT
	DPCP new	Up to 0.5%/week			
Shapiro 1993[81]	DPCP + vehicle	DPCP 1xweek + vehicle 2xday	Topical IT	15*	Double-blind RCT
	DPCP + Minoxidil	DPCP 1xweek + 5% minoxidil 2xday			
Tiwary 2016[82]	DPCP	1x week (in 0.01-1% increases until prespecified tolerance)	Topical IT	12	Single-blind RCT
	SADBE	1x week (in 0.01-1% increases until prespecified tolerance)		12	
Al Bazzal 2021[83]	DPCP novel regime	Initial dose chosen based on a prespecified reaction to 6 different concentrations of DPCP	Topical IT	15	Open-label RCT
	DPCP standard regime	1x week (in 0.01-1% increases until prespecified tolerance)		15	
Rocha 2021[84]	DPCP	Gradual dose increase until mild eczema	Topical IT	13	Single-blinded RCT
	Anthralin	2% anthralin in petroleum jelly for 30 min	Topical anti-psoriatic	11	
Observational studies (n=13)					
Alsufyani 2017[85]	Methotrexate	Starting dose: 10 to 25 mg (±15 mg) with cumulative dose to onset of response: 30 to 630 mg (±180 mg)	Oral IS	28	Retrospective
English 2015[86]	Methotrexate	15-20mg/week	Oral IS	31	Retrospective
Ferrando 1999[87]	Cyclosporine	Starting dose 5 mg/kg/day adjusted to between 100 and 350 ng/ml; average 150 mg twice a day	Oral IS	15	Retrospective
Firooz 2013[88]	Methotrexate	5-10mg/week	Oral IS	10	Prospective case-series

Maryam 2009[89]	DPCP	1x week (in 0.01-1% increases until prespecified tolerance)	Topical IT	54	Retrospective
Joly 2006[90]	Methotrexate	15-25mg/week	Oral IS (n=15) Subcutaneous IS injection (n=7)	22	Retrospective
Shapiro 1997[91]	Cyclosporine	5mg/kg/day	Oral IS	8	Prospective case-series
Gupta 1990[92]	Cyclosporine	6mg/kg/day	Oral IS	6	Prospective case-series
Jang 2016[93]	Cyclosporine	50-400mg/day	Oral IS	51	Retrospective
	Betamethasone	2-6mg/2xweek	Oral Steroid	37	
Shin[a] 2018[94]	Tofacitinib[a]		Oral JAK inhibitor		Retrospective
	Steroid + cyclosporine	24 mg/day in week 1 then 4mg/day + 150-300 mg/day	Combination	26	
	DPCP	0.0001 to 1%	Topical IT	30	
Sriphojanart 2017[95]	DPCP standard	0.0001%/week	Topical IT	23	Retrospective
	DPCP new regime	Up to 0.5%/week		16	
Vano-Galvan 2016[96]	Pulse corticosteroid with dexamethasone	0.1/mg/kg/day twice weekly	Oral Steroid	31	Prospective
Vano-Galvan 2016[97]	Azathioprine	2.5 mg/kg/day	Oral	14	Prospective

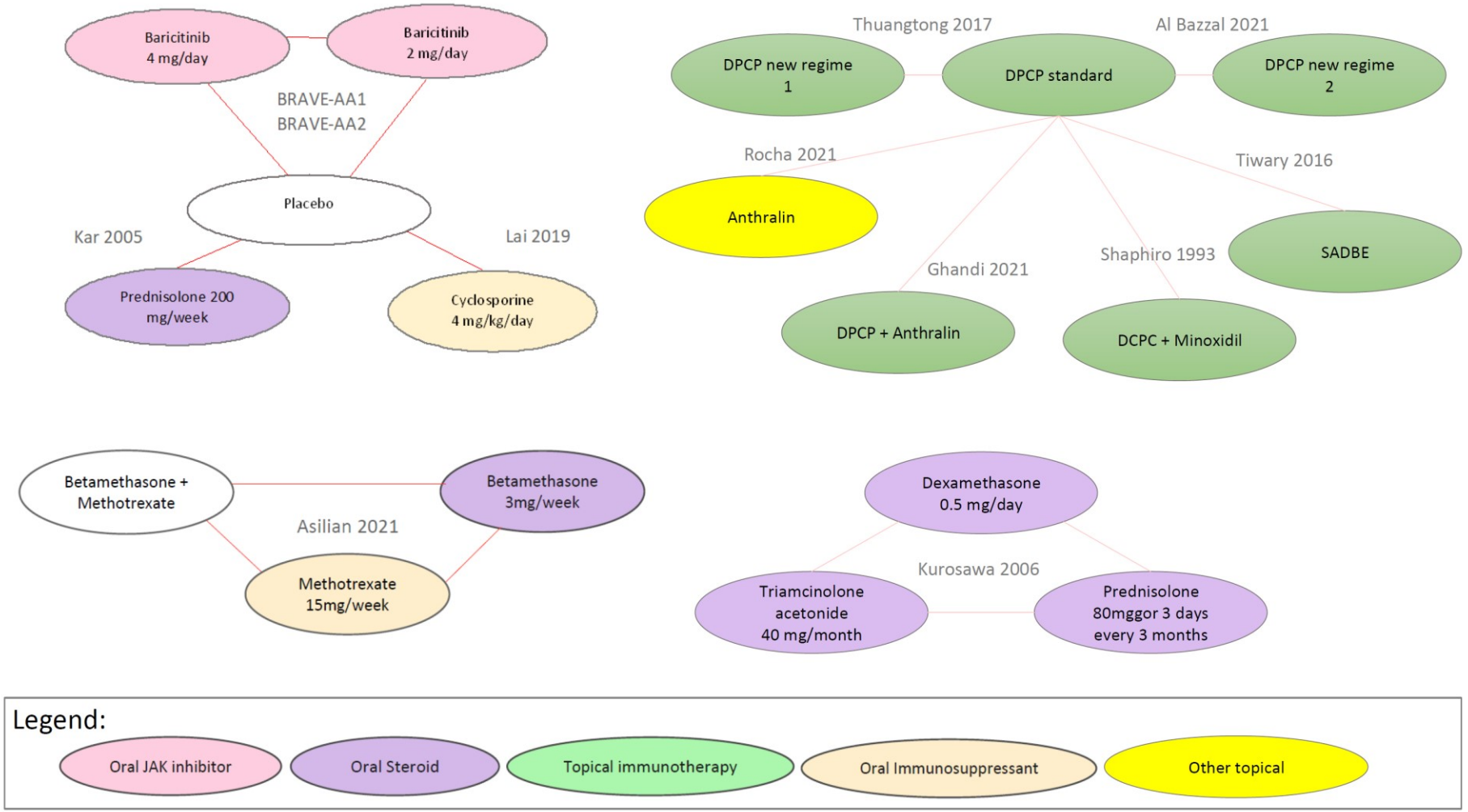
Footnotes: [a] The tofacitinib arm was excluded from further analysis due to lack of regulatory approval and established clinical management
Abbreviations: DPCP 2,3-diphenylcyclopropenone; IFN interferon; IS Immunosuppressant; IT Immunotherapy; JAK Janus kinase; RCT randomised controlled trial; SADBE squaric acid dibutyl ester. * Within-patient control.

B.2.9.2 Feasibility assessment

An evidence network for the 12 RCTs included in the feasibility assessment is shown in Figure 25. While the four placebo-controlled trials form a connected evidence network, the remaining studies in the feasibility assessment comprise a disconnected network of uncontrolled studies.

A summary of the baseline characteristics and outcomes of all included studies in the feasibility assessment is presented in Table 35 and Table 36.

Figure 25. Evidence networks based on the 4 placebo-controlled and 7 non placebo-controlled RCTs



Abbreviations: DPCP: diphenylcyclopropenone; IFN: interferon; JAK: Janus kinase; SADBE: squaric acid dibutyl ester.

Table 35. Baseline characteristics of the 25 studies included in the feasibility assessment

Study	Total N	Treatment	Current age Mean (years)	Age at onset Mean (years)	Disease duration Mean (years)	Current episode Mean (years)	Mean SALT score/% hair loss	Ophiasis subtype (%)	Family history (%)
Placebo-controlled RCTs (n=4)									
BRAVE-AA2	189	Placebo	37.4	≥18y: 59.3%	12.6	3.5	84.7	6.9	-
	184:281	Baricitinib (2:4mg)	38.0, 36.3	≥18y: 67.9%, 61.6%	12.1, 11.8	3.9, 3.5	86.8, 85.3	9.2, 9.6	-
BRAVE-AA1	156	Placebo	37.1	25.1	11.8	4.7	85.0	7.7	-
	156:234	Baricitinib (2:4mg)	39.0, 38.0	25.9, 26.0	13.1, 11.9	4.4, 3.9	85.6, 84.8	10.3, 10.3	-
Lai 2019	18	Placebo	45.7	29.3	-	5.7	81.1	-	5.6
	18	Cyclosporine	36.4	19.7	-	7.4	77.8	-	16.7
Kar 2005	20	Placebo	30.2	27.5	2.8	-	≥40	-	18.8
	23	Prednisolone	26.3	23.0	3.1	-	≥40	-	10.0
Non-placebo-controlled RCTs (n=8)									
Asilian 2021	12	Methotrexate	31.3	-	6	-	100	-	-
	12	Betamethasone	27.5	-	4.5	-	100	-	-
	12	Both	25.8	-	6.4	-	100	-	-
Al Bazzal 2021	15	DPCP novel regime	24.1	16.6	6.8	-	69.5	-	13.3
	15	DPCP standard regime	30.8	28.0	4.8	-	85.2	-	13.3
Ghandi 2021	21	DPCP alone	>30y: 47.6%	<10y:14.3%; 11-30y: 57.2%; >31y: 28.6%	-	-	By quartiles: 9.5%, 33.3%, 23.8%, 33.33%	23.8	-
	22	DPCP + anthralin	>30y: 36.4%	<10y:13.6%; 11-30y: 54.5%; >31y: 31.8%	-	-	By quartiles: 22.7%, 13.6%, 18.2%, 45.5%	18.2	-
Kurosawa 2006	19	Dexamethasone	-	-	-	-	-	-	-
	43	Triamcinolone	-	-	-	-	-	-	-

	29	Prednisolone	-	-	-	-	-	-	-
Rocha 2021	13	DPCP	36.8	23.7	-	3**	98**	-	-
	11	Anthralin	34.1	17.8	-	2**	10**	-	-
Thuangtong 2017	20*	DPCP standard	35	-	1	>1	-	-	-
		DPCP new							
Shapiro 1993	7*	DPCP	32.7	-	11.1	-	≥75 (42.9%)	-	-
	6*	DPCP + Minoxidil	44.8	-	20.8	-	≥75 (33.3%)	-	-
Tiwary 2016	12	DPCP	-	-	-	-	34.45	16.7	-
	12	SADBE	-	-	-	-	52.25	16.7	-
Observational studies (n=13)									
Alsufyani 2017	28	Methotrexate	39.0**	-	5.0**	-	-	-	-
English 2015	31	Methotrexate	40.0	-	-	-	≥50 (100%)	16.0	-
Ferrando 1999	15	Cyclosporine	28.9	-	9.9	-	≥70 (46.7%)		
Firooz 2013	10	Methotrexate	29.6	-	8.1	-	-	10.0	-
Gupta 1990	6	Cyclosporine	26	-	8.0	-	5.6***	-	-
Jang 2016	51	Cyclosporine	36.3	32.4	19.6% >5 years	-	≥50 (78.4%)	-	7.8
	37	Betamethasone	38.7	30.4	18.9 % >5 years	-	≥50 (73.0%)	-	10.8
Joly 2006	22	Methotrexate	37.6	-	11.1	-	-	-	-
Maryam 2009	54	DPCP	-	-	7.8	-	>75 (68.5%) 26-75 (29.6%) <25 (1.9%)	-	-
Shapiro 1997	8	Cyclosporine	-	-	7.5	-	≥95	-	-
Shin[a] 2018	26	Steroid + cyclosporine	30.5**	23.5**	5.0**	30.8% >5y	98.5**	-	-
	30	DPCP	33**	25.0**	5.0**	50.0% >5y	98.1**	-	-

Sriphojanart 2017	23	DPCP standard	35.5	26.9	1.1	-	-	-	8.6
	16	DPCP new	32.9	23.5	1.2	-	-	-	18.8
Vano-Galvan 2016	31	Pulse corticosteroid with dexamethasone	35.2	-	-	-	-	-	-
Vano-Galvan 2016	14	Azathioprine	37.7	-	2	-	-	-	-

Footnotes: [a] The tofacitinib arm excluded from further analysis due to lack of regulatory approval and established clinical management. * Within-patient control. ** Median. *** Severity score of hair loss on the scalp. a The tofacitinib arm excluded from analysis due to lack of regulatory approval and established clinical management
Abbreviations: DPCP 2,3-diphenylcyclopropenone; RCT randomised controlled trial; SADBE squaric acid dibutyl ester; SALT Severity of Alopecia Tool.

Table 36. Availability of outcome data for the studies included in the feasibility assessment

Study	Treatment	Total N	Timepoints (weeks)	Continuous SALT score	Dichotomous SALT score	Categorical response (% hair regrowth)	Eyelash/ brow score
Placebo-controlled RCTs (n=4)							
BRAVE-AA1	Baricitinib	654	12, 24, 36	Mean	≤10, ≤20, $SALT_{50, 75, 90, 100}$	-	ClinRO/PRO
BRAVE-AA2	Baricitinib	546	12, 24, 36	Mean	≤10, ≤20, $SALT_{50, 75, 90, 100}$	-	ClinRO/PRO
Lai 2019	Cyclosporine	36	12	Mean	$SALT_{30, 50, 75, 100}$	-	ClinRO
Kar 2005	Prednisolone	43	12	-	-	≥31%, >60%	-
Non-placebo-controlled RCTs (n=8)							
Asilian 2021	Methotrexate	36	12, 24, 36	-	-	Mean %	-
	Betamethasone						
Al Bazzal 2021	DPCP	30	24	Mean		>75%	
Ghandi 2021	DPCP	50	12, 24	Mean		Mean %; 0%, quartiles, 100%	
Kurosawa 2006	Dexamethasone	91	-	-	-	>40%	-
	Triamcinolone						
	Prednisolone						

Thuangtong 2017	DPCP	20	34	-	-	>75%	-
Rocha 2021	DPCP	24	12, 24	Median	-	>75%	
	Anthralin						
Shapiro 1993	DPCP	15	24	-	-	>75%	-
Tiwary 2016	DPCP	48	24	Median	-	0%, quartiles 100%	-
	SADBE						
Observational studies (n=13)							
Alsufyani 2017	Methotrexate	28	-	-	-	0%, quartiles 100%	-
English 2015	Methotrexate	31	-	-	-	100%	-
Ferrando 1999	Cyclosporine	15	-		-	IPD published	-
Firooz 2013	Methotrexate	10	-	-	-	100%	-
Joly 2006	Methotrexate	22	-	-	-	100%	-
Maryam 2009	DPCP	54	-	-	-	quartiles	-
Shapiro 1997	Cyclosporine	8	24	-	-	75-100%	-
Gupta 1990	Cyclosporine	6	-	-	-	>60	-
Jang 2016	Cyclosporine	51	36-42	-	-	-	-
	Betamethasone						
Shin[a] 2018	Steroid+cyclosporin	56	12, 24	Median	$SALT_{50}$	>5%, >50%, >90%	-
	DPCP						
Sriphojanart 2017	DPCP	39	52	-	-	0%, quartiles 100%	-
Vano-Galvan 2016	Pulse corticosteroid with dexamethasone	31	-	-	-	0%, <75%, ≥75%	-
Vano-Galvan 2016	Azathioprine	14	-	-	-	0%, <75%, ≥75%	-

Footnotes: [a] The tofacitinib arm excluded from further analysis due to lack of regulatory approval and established clinical management
Abbreviations: ClinRO clinician reported outcome; DPCP 2,3-diphenylcyclopropenone; IFN interferon; PRO patient reported outcome RCT randomised controlled trial; SADBE squaric acid dibutyl ester; SALT Severity of Alopecia Tool.

Heterogeneity of trial populations

The patients recruited to the BRAVE-AA studies were notably older (mean age 38 years), with longer disease duration (mean 12 years) and greater disease severity (>85% hair loss), compared to the 21 studies outside of the connected network (where data were reported). The average age in these studies mainly ranged between 25–35 years and disease duration was generally around 4–8 years. Only four other studies (beyond the BRAVE-AA and Lai 2019 studies) reporting hair regrowth outcomes focussed on a patient population with severe or very severe AA. Although current age and disease duration were generally widely reported by the 18 studies outside of the connected network (76%), other patient characteristics were not, so it was difficult to evaluate how comparable study populations were. Disease severity at baseline was only reported by 9/21 (29%), and age of onset and duration of current episode by 3/21 (14%) (Table 35).

Endpoint assessment

Four other studies investigated cyclosporine and one other study investigated prednisolone, but these either did not have comparable outcomes to the three placebo-controlled trials, or did not report data at the same time points (Table 36).

Aside from cyclosporine, other oral immunosuppressants investigated included methotrexate in six studies, five of which reported 100% regrowth as an outcome and one reported mean percentage regrowth. One study reported data at 24 weeks and four studies did not specify follow-up time. One study investigated azathioprine in 14 patients and reported response as regrowth of >75% of the scalp.

Aside from the prednisolone, other oral steroids investigated included betamethasone in two studies, which reported mean percentage of regrowth, but at different timepoints. Two studies investigated dexamethasone and one triamcinolone acetonide and reported percentage regrowth but did not specify follow-up time.

Six RCTs and three observational studies examined the topical immunotherapies DPCP, and SADBE. The British Association of Dermatologists (BAD) guidelines recommends topical immunotherapy for extensive hair loss and AT/AU, so despite being short course treatments, it is possible that a comparison with baricitinib would be of interest. However, such a comparison would be very limited, given that these were all small, heterogeneous studies reporting different thresholds of hair regrowth as outcomes. Furthermore, since topical immunotherapy is a short-course treatment, patients may relapse on treatment cessation and baricitinib may have benefits beyond this time frame.

Feasibility of conducting indirect or mixed treatment comparisons

For the connected evidence-network formed by the four placebo-controlled trials, it would be possible to conduct separate pairwise anchored indirect comparisons to compare the JAK inhibitor baricitinib with the oral immunosuppressant cyclosporine, and with the oral steroid prednisolone due to the linking of a common placebo in the trials. However, although it is statistically possible to perform such network comparisons, it would not be recommended due to some major limitations. Firstly, the sample sizes in the cyclosporine and prednisolone trials are very small compared to the available evidence for baricitinib from the BRAVE-AA studies (32 and 36 patients with outcome data in Lai 2019 and Kar 2005, respectively versus ~1200 patients combined in the BRAVE-AA studies), hence results would be highly uncertain with very limited

interpretation. In addition, complete data for potential prognostic factors and effect modifiers is not available for either of these comparisons. While it may be possible to apply a population-matching method, given that IPD are available for the two BRAVE-AA studies, it is still likely to yield biased comparisons due to missing prognostic factor data, and the effective sample sizes would be reduced even further. Furthermore, these are short-term treatments used across the spectrum of AA patients, whereas baricitinib is a longer-term therapy targeted at severe AA patients, so the clinical utility of this is unlikely to be warranted.

The remaining 18 studies included in the feasibility assessment comprise a disconnected network of uncontrolled studies with very small sample sizes and heterogeneity between patient populations. Baseline characteristics and timing of follow-up were not consistently reported, and outcome measures varied across studies. Although unanchored indirect comparisons are statistically possible, this would not be recommended unless there was a strong clinical imperative to do so, due to the strong assumptions underlying any analysis and uncertainty due to limited data.

Due to these reasons, no indirect or mixed treatment comparisons were conducted.

B.2.10 Adverse reactions

The safety of baricitinib in the treatment of severe AA is consistent with the known safety profile of baricitinib in other indications

- The BRAVE-AA trials found baricitinib to have a tolerable safety profile, with nasopharyngitis and headaches representing the most common AEs in BRAVE-AA1 and BRAVE-AA2, respectively
- A numerically higher proportion of baricitinib-treated patients reported TEAEs and SAEs as compared with placebo
- No deaths occurred in the placebo or baricitinib treatment groups across both trials

B.2.10.1 Summary of adverse events

The safety of baricitinib versus placebo was evaluated in the BRAVE-AA1 and BRAVE-AA2 trials and is presented separately for each trial.

Across the BRAVE-AA1 and -AA2 trials, a numerically higher proportion of baricitinib-treated patients reported TEAEs and AEs compared with placebo. In BRAVE-AA1, there was also a higher proportion of AEs and discontinuations from the study drug due to AEs, with the greatest frequency of TEAEs being reported in the 4 mg baricitinib group. However, none of these differences across either trial appeared clinically meaningful, and most events were assessed as being mild to moderate in severity. No patients died during either of the BRAVE-AA studies (Table 37).

Table 37. Overview of adverse events in the BRAVE-AA studies

	BRAVE-AA1			BRAVE-AA2		
	PBO (N=189)	2 mg baricitinib (N=183)	4 mg baricitinib (N=280)	PBO (N=154)	2 mg baricitinib (N=155)	4 mg baricitinib (N=233)
Patients with ≥1 TEAE, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Deaths	0	0	0	0	0	0
SAEs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs leading to permanent discontinuation from study intervention, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs leading to discontinuation from study, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; PBO: placebo; SAE: serious adverse event; TEAE: treatment emergent adverse event.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

B.2.10.2 Treatment emergent adverse events

Common TEAEs across both BRAVE-AA trials, defined as TEAEs that occurred in ≥2% of patients in any treatment group including placebo, are summarised in Table 73. Across both trials, the most common TEAEs were:

- Upper respiratory tract infection

- Headache
- Nasopharyngitis
- Acne
- Urinary tract infection.

These events have been recognised as adverse drug reactions (ADRs) in the established safety profile of baricitinib. Incidence rates of TEAEs did not increase with longer exposure.

Table 38. Summary of TEAEs affecting ≥2% of patients in any treatment groups in BRAVE-AA1 and BRAVE-AA2

TEAEs affecting ≥2% of patients, n (%)	PBO, n (%)	2 mg baricitinib, n (%)	4 mg baricitinib, n (%)
BRAVE-AA1	**(N=189)**	**(N=183)**	**(N=280)**
Nasopharyngitis	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]
Influenza	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]
Viral upper respiratory tract infection	[redacted]	[redacted]	[redacted]
Folliculitis	[redacted]	[redacted]	[redacted]
BRAVE-AA2	**(N = 154)**	**(N = 155)**	**(N = 233)**
Headache	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]
Acne	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]	[redacted]
Blood creatine phosphokinase increased	[redacted]	[redacted]	[redacted]
Dyslipidaemia	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]
Dyspepsia	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]
Folliculitis	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]
Oropharyngeal pain	[redacted]	[redacted]	[redacted]
Menstruation irregular	[redacted]	[redacted]	[redacted]

Abbreviations: TEAE: treatment emergent adverse event.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

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 both BRAVE-AA RCTs are presented in Table 39. A numerically higher proportion of patients in the baricitinib treatment groups reported SAEs compared to the placebo group across both trials. However, most SAE preferred terms (PTs) were single cases, spread across the treatment groups that had no cluster of events to indicate a potential safety issue. The most common SAEs in the baricitinib-treated participants were reported in the injury, poisoning, and procedural complications system organ class (SOC), with traumatic fractures accounting for most events; and in the infections and infestations SOC, with COVID-19-related events being most frequently reported.

Table 39.Serious adverse events across all treatment groups in the BRAVE-AA trials

SAEs by SOC and PT	PBO n (%)	2 mg baricitinib n (%)	4 mg baricitinib n (%)
BRAVE-AA1	**(N = 189)**	**(N = 183)**	**(N = 280)**
Cardiac disorders	[redacted]	[redacted]	[redacted]
Ventricular tachycardia	[redacted]	[redacted]	[redacted]
Acute myocardial infarction	[redacted]	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]	[redacted]
Food poisoning	[redacted]	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]
Chest pain	[redacted]	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]	[redacted]
Facial bones fracture	[redacted]	[redacted]	[redacted]
Ankle fracture	[redacted]	[redacted]	[redacted]
Foot fracture	[redacted]	[redacted]	[redacted]
Hand fracture	[redacted]	[redacted]	[redacted]
Humerus fracture	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]
Guillain-Barre syndrome	[redacted]	[redacted]	[redacted]
Pregnancy, puerperium, and perinatal conditions	[redacted]	[redacted]	[redacted]
Abortion missed	[redacted]	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]
Rhabdomyolysis	[redacted]	[redacted]	[redacted]
Renal and urinary disorders	[redacted]	[redacted]	[redacted]
Nephrolithiasis	[redacted]	[redacted]	[redacted]
BRAVE-AA2	**(N = 154)**	**(N = 155)**	**(N = 233)**
Patients with ≥1 SAE	[redacted]	[redacted]	[redacted]
Gastrointestinal Disorders	[redacted]	[redacted]	[redacted]
Inguinal hernia	[redacted]	[redacted]	[redacted]

Strangulated umbilical hernia	[redacted]	[redacted]	[redacted]
Hepatobiliary disorders	[redacted]	[redacted]	[redacted]
Cholecystitis acute	[redacted]	[redacted]	[redacted]
Infections and Infestations	[redacted]	[redacted]	[redacted]
Pyelonephritis	[redacted]	[redacted]	[redacted]
COVID-19 pneumonia	[redacted]	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]	[redacted]
Lumbar vertebral fractures	[redacted]	[redacted]	[redacted]
Ankle Fracture	[redacted]	[redacted]	[redacted]
Investigations	[redacted]	[redacted]	[redacted]
SARS-Cov-2 test positive	[redacted]	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (including cysts and polyps)	[redacted]	[redacted]	[redacted]
B-cell lymphoma	[redacted]	[redacted]	[redacted]
Prostate cancer	[redacted]	[redacted]	[redacted]
Product issues	[redacted]	[redacted]	[redacted]
Device dislocation	[redacted]	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]	[redacted]
Cardiac failure congestive	[redacted]	[redacted]	[redacted]

Abbreviations: BARI: baricitinib; PBO: placebo; PT: preferred term; SOC: system organ class
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

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 the study protocol. All AEs which resulted in permanent discontinuation from study treatment in either treatment arm across all the BRAVE-AA studies are presented in Table 40. In BRAVE-AA1, a numerically higher proportion of patients in the baricitinib groups reported AEs leading to permanent discontinuation from the study intervention compared to placebo, though no event PTs were reported more than once in any treatment group. In BRAVE-AA2, the same proportion of patients across all treatment groups reported AEs leading to permanent discontinuation from the study intervention. Consistent with BRAVE-AA1, no event preferred terms (PTs) were reported more than once in any treatment group. Overall, across both trials, no clinically meaningful differences were noted between treatment groups. Furthermore, incidence of AEs leading to permanent discontinuation of study drug did not increase with longer exposure.

Table 40. Adverse events leading to permanent discontinuation from study treatment across all treatment groups in the BRAVE-AA trials

	PBO, n (%)	2 mg baricitinib, n (%)	4 mg baricitinib, n (%)
BRAVE-AA1	**(N = 189)**	**(N = 183)**	**(N = 280)**
Subjects with ≥1 AE	[redacted]	[redacted]	[redacted]
Blood and lymphatic system disorders	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]	[redacted]
Ventricular tachycardia	[redacted]	[redacted]	[redacted]
Investigations	[redacted]	[redacted]	[redacted]
Weight increased	[redacted]	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]
Hepatic enzyme increased	[redacted]	[redacted]	[redacted]
Transaminases increased	[redacted]	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]
Bursitis	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]
Guillain-Barre syndrome	[redacted]	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]	[redacted]
Anxiety	[redacted]	[redacted]	[redacted]

BRAVE-AA2	(N = 154)	(N = 155)	(N = 233)
Patients with ≥1 AE	[redacted]	[redacted]	[redacted]
Blood and lymphatic system disorders	[redacted]	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]	[redacted]
General disorders and administration	[redacted]	[redacted]	[redacted]
Non-cardiac chest pain	[redacted]	[redacted]	[redacted]
Hepatobiliary disorders	[redacted]	[redacted]	[redacted]
Cholecystitis acute	[redacted]	[redacted]	[redacted]
Investigations	[redacted]	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]
Weight increased	[redacted]	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (includes cysts and polyps)	[redacted]	[redacted]	[redacted]
B-cell lymphoma	[redacted]	[redacted]	[redacted]
Prostate cancer[a]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Denominator adjusted because gender-specific event for males: N = 57 (PBO), N = 53 (BARI 2 mg), N = 89.
Abbreviations: AE: adverse event; PBO: placebo; PT: preferred term; SOC: system organ class.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

B.2.10.5 Adverse events of special interest (AESI)

AESIs defined in the study protocol include infections, malignancies, hepatic events, gastrointestinal perforations, major adverse cardiovascular events (MACE) and thromboembolic events, such as venous thromboembolism (VTE) and arterial thromboembolism (ATE). All AESIs recorded in either treatment arm across all BRAVE-AA studies are presented in Table 41.

Table 41. Adverse events of special interest across all treatment groups in the BRAVE-AA trials

	PBO n (%)	2 mg baricitinib n (%)	4 mg baricitinib n (%)
BRAVE-AA1	**(N = 189)**	**(N = 183)**	**(N = 280)**
Patients with ≥1 TE infection	[redacted]	[redacted]	[redacted]
TE herpes zoster	[redacted]	[redacted]	[redacted]
TE herpes simplex	[redacted]	[redacted]	[redacted]
Positively adjudicated MACE	[redacted]	[redacted]	[redacted]
BRAVE-AA2	**(N = 154)**	**(N = 155)**	**(N = 233)**
Patients with ≥1 TE infection	[redacted]	[redacted]	[redacted]
Serious infections	[redacted]	[redacted]	[redacted]
TE herpes zoster	[redacted]	[redacted]	[redacted]
TE herpes simplex	[redacted]	[redacted]	[redacted]
Malignancies other than NMSC	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; NMSC: nonmelanoma skin cancer; MACE: major adverse cardiovascular event; PBO: placebo; TE: treatment emergent.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[65, 66]

Infections

Across both BRAVE-AA trials, a similar proportion of patients reported treatment-emergent infections in the baricitinib and placebo groups. No dose-relationship was observed for infections. Most infections were mild or moderate in severity. The most frequently reported infections across all treatment groups were upper respiratory tract infection and nasopharyngitis, and frequencies were similar for both baricitinib groups compared with placebo. Baricitinib treatment did not increase the risk for serious infections or temporary or permanent discontinuation of study drug due to infections compared with placebo. All patients recovered, and no patient discontinued study drug due to serious infection.

Herpes zoster, recognised as an ADR for baricitinib, was numerically more frequent with baricitinib compared with placebo. No disseminated infections were observed during the placebo-controlled period of either trial. No dose response was observed. Herpes simplex, recognised as an ADR for baricitinib, occurred at a similar frequency in placebo and baricitinib 4 mg groups in BRAVE-AA2. In BRAVE-AA2, a lower proportion of patients in the baricitinib groups reported treatment emergent (TE) herpes simplex compared to the placebo group. All herpes simplex infections across both trials were mild or moderate in severity, with no serious infection. The most frequently reported event term was oral herpes. There were no reports of tuberculosis and no reports of viral hepatitis in the BRAVE-AA clinical trials.

Other AESI

Across both trials, there were no reports of positively adjudicated VTE, or ATE and no reports of nonmelanoma skin cancer (NMSC) or confirmed gastrointestinal perforation in any treatment group. In BRAVE-AA1, a myocardial infarction occurred in a patient who received 2 mg baricitinib and had cardiovascular risk factors; this event was not considered to be related to the study drug by the investigator. In BRAVE-AA2, there were two instances of malignancies reports, a prostate cancer in the placebo group, and a B-cell lymphoma in the 4 mg baricitinib group.

B.2.10.6 Deaths

There were no deaths in either BRAVE-AA clinical trials.

B.2.10.7 Treatment-emergent laboratory changes

Haematologic changes

Changes in neutrophils, lymphocytes, haemoglobin, and platelet counts for patients in the BRAVE-AA trials receiving baricitinib were consistent with the established safety profile of baricitinib.

Abnormal hepatic tests

Alanine aminotransferase and aspartate aminotransferase increases of ≥3× the upper limit of normal (ULN) are recognised as ADRs for baricitinib. Increases of transaminases to ≥3× ULN were seen less frequently with baricitinib treatment than with placebo across both trials. Increases to ≥5× or ≥10× ULN were uncommon.

Creatine phosphokinase increases

Creatine phosphokinase (CPK) increases of ≥5× ULN are recognised as ADRs for baricitinib. Across both trials, dose related CPK increases were observed in the blinded period of the BRAVE-AA trials. In addition, a higher proportion of baricitinib-treated patients had CPK elevations compared with placebo. Most CPK increases on baricitinib treatment resulted in increases to Common Terminology Criteria for Adverse Events (CTCAE) Grades 1 and 2. The frequency of CTCAE Grade 3 and 4 increases was similar for baricitinib 4 mg and placebo with no such increases within the baricitinib 2 mg group.

Lipid changes

The proportion of patients who had a categorical increase (according to NCEP criteria) for total cholesterol, low-density lipoprotein-cholesterol (LDL-C), and high-density lipoprotein-cholesterol (HDL-C) was higher in the baricitinib 2 mg and 4 mg groups, compared with the placebo group. No clinically meaningful differences were observed for triglycerides in the baricitinib groups compared with the placebo group. Lipid changes were not associated with MACE.

B.2.10.8 Extended safety data

Safety data are also available from the long-term extension period of the BRAVE-AA trials, covering the period from randomization to the 23rd of August 2021 and the 20th of August 2021 for BRAVE-AA1 and BRAVE-AA2, respectively. No new safety findings were observed during this period, and the incidence of SAEs, TEAEs and AEs leading to permanent discontinuation did not increase with longer exposure (Table 42).

Table 42. Summary of AEs from the extended period in the pooled BRAVE-AA1 and BRAVE-AA2 analysis set

	2 mg baricitinib (N=365; PYE=371.5) n (%) [IR]	4 mg baricitinib (N=540; PYE=624.3) n (%) [IR]
Deaths	[redacted]	[redacted]
SAE	[redacted]	[redacted]
TEAE[a]	[redacted]	[redacted]
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]
Severe	[redacted]	[redacted]
AEs leading to permanent discontinuation from study intervention, n (%)	[redacted]	[redacted]
AEs leading to discontinuation from the study, n (%)	[redacted]	[redacted]

Footnotes: [a] Patients with multiple occurrences of the same event are counted under the highest severity.
Abbreviations: AA: alopecia areata; AE: adverse event; BARI: baricitinib; DC: discontinuation; IR: incidence rate; PBO: placebo; PYE: patient-year exposure; PYR: patient-years at risk; SAE: serious adverse event; TEAE: treatment-emergent adverse event.

B.2.11 Ongoing studies

No ongoing studies of baricitinib which have not been discussed in this submission are expected to be published in the next 12 months.

B.2.12 Interpretation of clinical effectiveness and safety evidence

Principal findings from the clinical evidence base

Baricitinib is a clinically effective treatment for patients with severe AA. This is a disease with a significant burden on patients' health-related quality of life, with limited treatment options. As the first evidence-based treatment, baricitinib is an innovative therapy that represents an important milestone in the treatment of severe AA.

The efficacy and safety of baricitinib has been demonstrated in two randomised, placebo-controlled, phase III studies (BRAVE-AA1 and BRAVE-AA2). The demonstration of efficacy as performed in a refractory population of patients with extensive hair loss (median SALT score of 96) and chronic disease (mean current episode duration of 3.9 years). In addition, approximately 90% of patients in the BRAVE-AA studies reported prior AA therapy, and over 50% had used systemic immunosuppressant or immunomodulator therapy.[67] Despite the reported poor prognoses of this patient population, results from both trials consistently demonstrate that baricitinib is an effective treatment option for severe AA, associated with robust improvement of symptoms, including significant scalp hair regrowth (SALT≤20) by Week 36, with outcomes far exceeding those for placebo treatment.[1] In addition, Week 52 data indicate that hair regrowth continues to increase among responders beyond 36 weeks, demonstrating that the plateau of response is not reached for all patients at Week 36 and that responding patients may experience further clinical benefit from baricitinib beyond this timepoint. Furthermore, data from Week 76 suggest that efficacy is maintained over time in most of the patients who have reached SALT≤20 at Week 52.

Beside the significant scalp hair regrowth observed in responders across both trials after 36 weeks, treatment with baricitinib 4 mg also led to a statistically significant eyelash and eyebrow regrowth at Week 36 when compared with placebo treatment, with the proportion of patients achieving an improvement in hair growth increasing over the treatment period. In addition, baricitinib was associated with a notable improvement in HRQoL and symptoms of anxiety and depression when compared with placebo. The BRAVE-AA trials also found baricitinib to have a tolerable safety profile, with nasopharyngitis and headaches representing the most common AEs in BRAVE-AA1 and BRAVE-AA2, respectively. No new safety findings were identified in the BRAVE-AA trials compared with the known safety profile of baricitinib established in other indications. All results were consistent across the two trial populations of BRAVE-AA1 and BRAVE-AA2, trials, highlighting that the results were robust and likely transferable to the general population with severe AA.

Baricitinib thus allows patients with severe AA who respond to treatment to benefit from significant hair regrowth as compared to established clinical management. In the context of current clinical practice within the NHS in England, this submission positions baricitinib as a treatment for adults with severe AA. This is in line with the anticipated marketing authorisation for baricitinib and the population specified in the NICE scope.

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 AA. The BRAVE-AA trials represent the primary sources of evidence for baricitinib as a treatment for adult patients with severe AA. The BRAVE-AA trials are large, placebo-controlled RCTs of good quality and thus provide robust evidence for the safety and efficacy of baricitinib for the treatment of adult patients with severe AA.

The variety of endpoints considered in both BRAVE-AA trials measuring hair regrowth and HRQoL are clinically relevant and important to both patients and clinicians. The benefits demonstrated in these trials will therefore translate to meaningful improvements for patients in clinical practice.

As discussed in Section B.2.5, the BRAVE-AA1 and BRAVE-AA2 trials were methodologically robust and well reported. The results were considered at a low risk of bias in all categories considered.

As a feasibility assessment deemed that the evidence base uncovered by the SLR was not suitable to conduct an indirect or mixed treatment comparison in this setting, there was no direct comparative efficacy evidence comparing baricitinib with current alternative treatment options. As treatment options for severe AA are limited and frequently includes no further pharmacological treatment, the clinical effectiveness evidence in this submission focusses on the placebo comparator used within the BRAVE-AA trials. This comparator reflects UK clinical practice, given that many patients with severe AA experience treatment failure with current off-label treatments, and therefore have no option left but to leave their AA untreated.

Overall conclusions

Severe AA leads to a significant burden on patients' health-related quality of life, with limited treatment options available. Baricitinib is a clinically effective treatment for patients with severe AA and addresses the substantial unmet need for an evidence-based, effective and well-tolerated medication in this indication, by providing patients who respond to treatment the opportunity to experience significant hair regrowth and improved health-related quality of life. As the first evidence-based treatment specific to AA, baricitinib therefore is an innovative therapy that represents an important step-change in the treatment of severe AA.

B.3 Cost-effectiveness

Cost-effectiveness model methodology

- An SLR of economic evaluations did not identify any prior cost-effectiveness analyses for any treatment in AA, including baricitinib. Accordingly, a de novo Markov model was developed that closely follows the model structures of other dermatological disorders such as psoriasis and atopic dermatitis
- The model includes four health states: “Induction”, “Maintenance”, “BSC” and “Death”, which were deemed representative of the treatment and disease progression in AA
- The pooled FAS populations from BRAVE-AA1 and BRAVE-AA2, comprised of patients with SALT scores of 50–100, were included in the base case
- The main comparator included in the model is a ‘watch and wait’ period, followed by patients transitioning to BSC after non-response, reflecting established clinical management in the UK
- The base case analysis categorised patients based on achieving at least a $SALT_{50}$ response status (i.e. an improvement of >50% from baseline in SALT score), while also accounting separately for those with a response of $SALT_{75}$ at Week 36, as $SALT_{50}$ was deemed an effective method for capturing sufficient clinical benefit to justify continuing treatment after the trial Induction period. Response rates were informed by the BRAVE-AA trials
- EQ-5D data were derived from a real-world evidence study as the EQ-5D data collected during the BRAVE-AA trials exhibited clear ceiling effects and implausibly high baseline values, which result in underestimations of the HRQoL gain associated with treatment response. Trial-based utilities were explored in scenario analyses
- Health state unit cost and resource use were sourced from NHS reference costs, NICE guidance for depression in adults and the PSSRU, with input from clinical experts as necessary
- In line with the NICE reference case, the analysis was conducted from the perspective of the UK NHS and Personal Social Services (PSS) over a lifetime horizon

Cost-effectiveness model results

- At the confidential PAS price, the ICER for baricitinib versus ‘watch and wait’ fell within the range considered to be cost-effective. At £29,111/QALY gained, it is below the NICE willingness-to-pay (WTP) threshold of £30,000
- These results demonstrate baricitinib to be a cost-effective option for the treatment of patients with severe AA when compared to ‘watch and wait’

Cost-effectiveness model conclusions

- Overall, the introduction of baricitinib for the treatment of severe AA into UK clinical practice is anticipated to bring substantial benefits to patients with severe AA for whom current treatment options are unable to fulfil a substantial unmet need for an effective, well-tolerated treatment that is able to restore hair growth and improve health-related quality of life
- This analysis demonstrates that baricitinib is a cost-effective treatment option that would offer value for money for the NHS. If recommended, baricitinib would represent the first treatment available that is specifically for the treatment of AA in the UK

B.3.1 Published cost-effectiveness studies

A *de novo* economic SLR was conducted on 19th August 2021 and updated on 11th January 2022 to identify cost-effectiveness, health-state utility values (HSUVs) and cost and healthcare resource use data for the treatment of adult patients with severe AA. The SLR identified a total of 35 relevant studies, including 4 studies reporting on cost and resource use for AA patients, as well as 30 studies reporting on HRQoL in AA patients. Full details of the SLR search strategy, study selection process and results are reported in Appendix G.

As the SLR did not identify any evaluations investigating the cost-effectiveness of any treatment in AA, including baricitinib, a *de novo* cost-effectiveness analysis of baricitinib versus the comparator relevant to the decision problem for this submission was performed.

B.3.2 Economic analysis

The objective of this economic analysis was to assess the cost-effectiveness of once-daily 4 mg baricitinib compared with established clinical management for the treatment of severe AA patients. Established clinical management was defined as a 'watch and wait' approach, informed by the placebo arm outcomes in the BRAVE-AA trials, followed by best supportive care. 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. A Markov structure was deemed appropriate to adequately represent the treatment and disease progression of AA patients. 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.[98] 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 4 mg baricitinib once-daily in adult patients with severe AA, in line with the indication of relevance for this submission.

The patient population included in the economic evaluation is in line with the eligibility criteria for the BRAVE-AA1 and BRAVE-AA2 trials: patients with a SALT score higher or equal to 50 at baseline. As such, the pooled population of patients from BRAVE-AA1 and BRAVE-AA2 at Week 36 (FAS population, N=[redacted]) is used in the base case, in line with the population specified in the decision problem in Section B.1 (Table 1) and the anticipated marketing authorisation for baricitinib in AA.

Consistent with the final NICE scope, the economic evaluation also considers two severity subgroups which are explored in subgroup analyses:

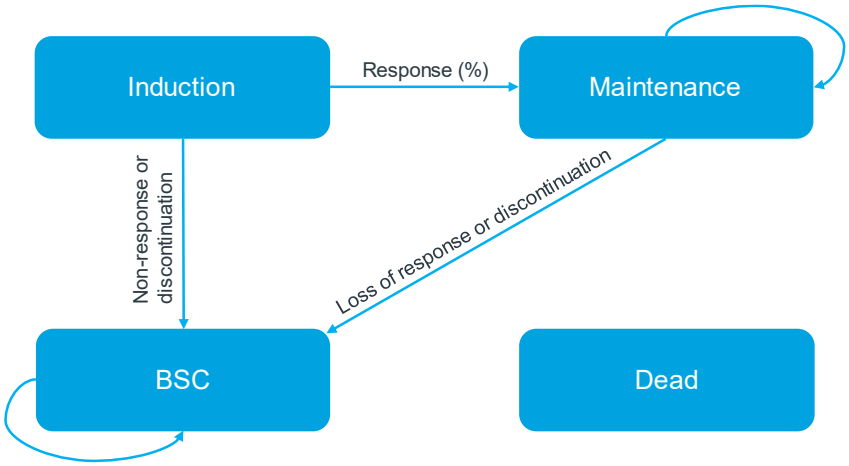
- SALT 50–94, severe population (N=[redacted])
- SALT 95–100, very severe population (N=[redacted])

B.3.2.2 Model structure

A cohort Markov state transition model was chosen to evaluate the cost-effectiveness of 4 mg baricitinib versus established clinical management in the target population and was constructed in Microsoft Excel. The model structure aimed to adequately capture the key features of AA and to be reflective of clinical practice in the UK. Given the lack of AA models identified in the literature, it was considered appropriate that the model structure reflected previous economic analyses in other dermatological disorders, such as psoriasis and atopic dermatitis. This model structure was discussed with a health economics expert and was deemed representative of the treatment and disease progression of AA patients.

The model structure is presented in Figure 26. The model includes four health states: "Induction", which is represented by a set of tunnel states, followed by "Maintenance", "BSC" and "Death".

Figure 26. 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 'watch and wait' and enter the Induction period. The length of the Induction period is 36 weeks for baricitinib and 'watch and wait', aligning with the double-blinded treatment period and primary response timepoint in the BRAVE-AA trials for baricitinib and placebo. This Induction period is achieved through using 9 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. Patients may remain in the induction period for 36 weeks or they may discontinue, in which case they move to the BSC state.

At the end of the Induction period, patient response to treatment is assessed. In the base case, response is defined as $SALT_{50}$, defined as the proportion of patients achieving at least a 50% improvement from baseline SALT score, as this was deemed an effective method for capturing sufficient clinical benefit to justify continuing treatment after the trial induction period. Using a higher cut-off (such as $SALT_{75}$) was deemed to be a less appropriate indicator of the potential to respond at this stage, due to the gradual nature of hair regrowth and the possibility that in some cases pattern baldness may lead to a ceiling on the maximum possible response. In addition, selection of the $SALT_{50}$ response level in the model captures the differential HRQoL gain associated with the achievement of $SALT_{50}$ versus $SALT_{75}$. Therefore, the $SALT_{50}$ endpoint is clinically meaningful, and also allows accurate determination of the HRQoL gain associated with varying levels of response to baricitinib treatment.

Patients who respond to treatment transition to the Maintenance health state, where they remain until loss of response or treatment discontinuation for other reasons. Non-responders are those who do not meet the definition of response at the end of the trial period. These patients transition to the BSC state, given there are no further lines of treatment available for AA. Therefore, upon entering BSC, patients remain there until the end of the simulation or death.

Entering the Death state is possible from each health state and a rate is applied each cycle. 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 UK general 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 43. Health state utility values are derived by cross-walking EQ-5D-5L scores collected in from a real-world evidence study, the Adelphi Disease Specific Programme (DSP) in AA, to EQ-5D-3L scores to be valued using the 3L value sets, in line with the NICE reference case.[99] Costs considered within the model include treatment acquisition costs and monitoring and disease management costs. Effectiveness measures include 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 rate 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 AA. Maximal lifetime for patients is set to 100 years, reflecting that the Office for National Statistics (ONS) life tables for mortality end at 100.[100]

The cycle length employed in the Markov model is four weeks; a half-cycle correction was not included in the model due to the short cycle length. The cycle length reflects the time points of baricitinib response assessment in the BRAVE-AA clinical trials and is also aligned to cycle lengths used in other dermatological conditions, such as AD. Given the different time reference of model inputs, including annually or per 36-weeks, calculations are performed in the model to rescale all variables to four-week duration. Two methods for rescaling were 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, 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 43. Features of the economic analysis

Factor	Current appraisal	
	Chosen values	Justification
Model structure	Markov state transition model with 4-week cycles	A Markov state transition model was chosen as this model structure reflects that of other dermatological disorders, such as psoriasis and atopic dermatitis, and was deemed representative of the treatment and disease progression of AA patient.
Time horizon	Lifetime	In line with the NICE reference case, and considered to reflect the fact that AA 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	Health state utilities are derived from the Adelphi DSP, which is a real-world evidence study conducted in Germany, Spain, Italy, France, and the UK	The Adelphi DSP will be used in the base case as the EQ-5D data captured directly from the baricitinib trials were insensitive to changes in the severity of AA. This is due to the lack of content validity of the EQ-5D instrument in this indication, leading to implausible high baseline utility

		values and subsequently, a ceiling effect. In addition, patients with severe psychological disorders were excluded from the trial, meaning that patients suffering the most severe psychiatric impact of AA would have been excluded from the utility estimates. The HRQoL improvement due to each response level is therefore more accurately captured using the Adelphi DSP data compared to the HRQoL data from the BRAVE-AA trials, and the Adelphi DSP better reflects the population expected to be treated in clinical practice.
Source of costs and resource use	National schedule of NHS costs,[101] NHS Drug Tariff,[102] NHS wigs and fabric supports costs[103] and Personal Social Services Research Unit (PSSRU)[104] and NICE guideline CG90.[105] Healthcare resource utilisation was obtained from the Adelphi RWE study and UK clinical expert opinion.	Established sources of costs within the NHS, and in line with the NICE reference case
Health effects measure	QALYs	In line with the NICE reference case

Abbreviations: AA: alopecia areata; EQ-5D: EuroQol-5 Dimension; DSP: Disease Specific Programme; HRQoL: health related quality of life; NICE: National institute for health and care excellence.

B.3.2.3 Intervention technology and comparators

The intervention of interest is 4 mg baricitinib administered orally once a day, as listed in Table 44. This is in line with the regimen used in the phase III BRAVE-AA trials supporting the submission, and the SmPC for baricitinib.[8] Patients who do not respond to baricitinib transition to BSC. As BSC remains poorly defined in UK clinical practice, a basket of treatments was used in the model. The composition of the BSC basket of treatments was sourced from a real-world evidence (RWE) study, specifically the Adelphi DSP in AA, starting in October 2021. Dermatologists who were actively treating patients with severe/ very severe AA in Germany, Spain, Italy, France, and the UK were enrolled in an online survey. They were asked to complete patient record forms for at least 7 consecutive adult patients with mild (n=1), moderate (n=3) and severe (n=3) AA; current disease severity was rated based on their clinical judgement. Physicians completed questionnaires regarding patient demographics, clinical status, and current treatments. In addition, each patient was invited to complete a self-completion form that included an EQ-5D-5L questionnaire. Current treatments for severe/very severe patients in the Adelphi DSP sample only for the UK (n=117) have been used in the base case. In a second option, explored as a scenario analysis, the estimates were based on separate discussions held with 3 UK-based key opinion leaders (KOLs) with current experience of treating patients with severe AA. The composition of BSC in the model is detailed in Table 45.

Table 44. Treatments included in the model

Treatment	Induction period (weeks)	Reference
Baricitinib 4mg	36	BRAVE-AA1/-AA2 clinical data
'Watch and wait'	36	BRAVE-AA1/-AA2 clinical data

The main comparator included in the model is established clinical management, defined as a 'watch and wait' period with patients transitioning to BSC following non-response. The 'watch and wait' period is informed by the efficacy data of the placebo comparator used in the BRAVE-AA trials. This approach reflects established clinical management in the UK, whereby patients presenting with severe AA may undergo a period of 'watch and wait', in which they receive no treatment, before trialling other interventions included in the BSC basket if no hair regrowth is observed. Other comparators of baricitinib used in clinical practice are included in a "basket" of BSC treatments, which are not associated with efficacy estimates in the model. This is a reasonable simplification of the model, due to a lack of robust clinical data for these comparators, meaning that comparison is not advisable based on the results of the feasibility analysis, as well as reflecting the fact that there is no established management pathway in severe AA.

Table 45. Best supportive care treatments estimated to be included in the BSC state

Treatment	Percentage of use	
	Adelphi DSP (base case)	KOL input (scenario analysis)
Ciclosporin	12.39%	12.50%
Methotrexate	14.29%	7.50%
Azathioprine	2.86%	8.67%
Intralesional steroids (triamcinolone acetonide)	9.53%	30.83%
DPCP (contact immunotherapy) treatment	20.96%	27.50%
Prednisolone	17.15%	25.00%
TCS: Mometasone ointment	24.77%	63.33%
Minoxidil 5% foam (topical)	5.72%	37.50%
Minoxidil tablets	0.00%	7.50%
Mycophenolate Mofetil	2.86%	0.00%
Anthralin 0.1% cream	5.72%	0.00%
Patients not currently on treatment *	13.00%	0.00%
Wig use (modacrylic wig)	80.00%+	80.00%

Footnotes: * used as BSC management to avoid double counting; + not sufficient number of responses received about wig use in the Adelphi DSP and the suggestion from the KOLs was used instead.
Abbreviations: DPCP, Diphenylcyclopropenone; DSP: Disease Specific Programme; KOL: key opinion leader; TCS, Topical corticosteroids.

In light of the above, baricitinib is positioned as a first-line therapy to treat patients newly diagnosed with severe AA, as well as in later-line treatment to treat patients who did not respond to other treatment strategies such as 'watch and wait', corticosteroid treatment or contact immunotherapy, as an alternative to established clinical management consisting of a 'watch and wait' strategy followed by BSC. The comparator included in the model therefore reflects the standard of care for patients in this setting in UK clinical practice.

B.3.3 Clinical parameters and variables

As described in Section B.3.2.2, four distinct health states are defined. In the base case, these are defined based on the achievement of $SALT_{50}$, defined as the proportion of patients achieving at least a 50% improvement from baseline SALT score. Patients transition between the Induction and Maintenance (i.e. responder) or BSC health states depending on the achievement of the $SALT_{50}$ endpoint following treatment with the intervention or comparators. Patients in the Maintenance health state may over time transition to the BSC or Death health states. Once patients enter the BSC 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 and established clinical management are derived from the phase III portion of the pivotal BRAVE-AA trials, given that these trials provide head-to-head evidence between baricitinib and placebo, which informs the 'watch and wait' period used as the comparator in the model.

B.3.3.1 Baseline characteristics

The baseline characteristics of the modelled cohort used in the base case and their source are presented in Table 46. Patient baseline characteristics were obtained from the pooled BRAVE-AA studies (Table 46). Baseline characteristics for the SALT 50–94 and SALT 95–100 populations utilised in scenario analyses are presented in Table 47.

Table 46. Baseline patient characteristics for the pooled BRAVE-AA FAS populations utilised in the base case

Characteristic	Mean value	SE	Reference
SALT 50–100 (FAS population)			
Age	[redacted]	[redacted]	Pooled BRAVE-AA1/-AA2 clinical studies
% Male	[redacted]	[redacted]	

Abbreviations: FAS: full analysis set; SALT: severity of alopecia Tool; SE: standard error.

Table 47. Baseline patient characteristics for the SALT 50–94 and SALT 95–100 populations utilised in scenario analyses

Characteristic	Mean value	SE	Reference
SALT 50–94 (severe population)			
Age	[redacted]	[redacted]	Pooled BRAVE-AA1/-AA2 clinical studies
% Male	[redacted]	[redacted]	
SALT 95–100 (very severe population)			
Age	[redacted]	[redacted]	Pooled BRAVE-AA1/-AA2 clinical studies
% Male	[redacted]	[redacted]	

Abbreviations: SALT: severity of alopecia Tool; SE: standard error.

B.3.3.2 Treatment response

The treatment response rates have been obtained from pooled data from the BRAVE-AA clinical trial programme, using the $SALT_{50}$ threshold measured at the end of the Induction period. In the base case, the Induction period is aligned with the primary endpoint of BRAVE-AA1 and BRAVE-

AA2 (36 weeks), though data from Week 52 suggests that some patients continue to respond after this time point (Section B.2.8.2). The base case response rates used in the model are presented in Table 48. The response rates for the SALT 50–94 and SALT 95–100 populations utilised in scenario analyses are presented in Table 50.

Table 48. Proportion of patients in the pooled BRAVE-AA FAS populations responding to treatment after the Induction period (36 Weeks) in the base case

Intervention	$SALT_{50}$		$SALT_{75}$	
	Response rate	SE	Response rate	SE
SALT 50–100 patients (FAS population)				
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]	[redacted]	[redacted]

Footnote: 'Watch and wait' treatment is informed by the placebo treatment groups in the BRAVE-AA trials.
Abbreviations: FAS: full analysis set; SALT, Severity of Alopecia Tool; SE, Standard Error.

Table 49. Proportion of SALT 50–94 and SALT 95–100 populations responding to treatment after the Induction period (36 Weeks)

Intervention	$SALT_{50}$		$SALT_{75}$	
	Response rate	SE	Response rate	SE
SALT 50–94 patients (severe population)				
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]	[redacted]	[redacted]
SALT 95–100 patients (very severe population)				
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]	[redacted]	[redacted]

Footnote: FAS: full analysis set; 'Watch and wait' treatment is informed by the placebo treatment groups in the BRAVE-AA trials.
Abbreviations: SALT, Severity of Alopecia Tool; SE, Standard Error.

B.3.3.3 Sustained response and long-term treatment discontinuation

Discontinuation from treatment is applied in the model in two stages. This can occur during the Induction period from model initiation until Week 36, and from Week 36 onwards during the Maintenance period. In the base case analysis, discontinuation is applied on a cycle basis according to the discontinuation rates presented in Table 50 for the base case analysis and in Table 51 for the SALT50–94 and SALT95–100 populations utilised in scenario analyses. Treatment-specific discontinuation is informed by the pooled data from the baricitinib phase III trials. Given that response is assessed at the end of the Induction phase, all-cause discontinuation during the Induction phase excludes lack of efficacy to avoid double counting this type of discontinuation. Discontinuation during the Maintenance phase is due to all causes, including lack of efficacy, as patients who lose their response are switched to BSC.

Due to lack of data beyond Week 36 for placebo, the Week 0–36 all-cause discontinuation rate is used to estimate the annual discontinuation rate in the Maintenance period. For baricitinib, data were available up to Week 52 and therefore the Week 0–52 all-cause discontinuation rate is used in the Maintenance period.

Table 50. Proportion of patients in the base case discontinuing treatment during the Induction and Maintenance periods

Intervention	Induction Period		Maintenance Period	
	Discontinuation	SE	Discontinuation	SE
SALT 50–100 patients (FAS population)				
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]	[redacted]	[redacted]

Footnote: 'Watch and wait' treatment is informed by the placebo treatment groups in the BRAVE-AA trials.
Abbreviations: FAS: full analysis set; SALT, Severity of Alopecia Tool; SE, Standard Error.

Table 51. Proportion of patients in SALT 50–94 and SALT 95–100 populations discontinuing treatment during the Induction and Maintenance periods utilised in scenario analyses

Intervention	Induction Period		Maintenance Period	
	Discontinuation	SE	Discontinuation	SE
SALT 50–94 patients (severe population)				
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]	[redacted]	[redacted]
SALT 95–100 patients (very severe population)				
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]	[redacted]	[redacted]

Footnote: 'Watch and wait' treatment is informed by the placebo treatment groups in the BRAVE-AA trials.
Abbreviations: SALT, Severity of Alopecia Tool; SE, Standard Error.

B.3.3.4 Adverse events

Costs and disutilities associated with TEAEs are not included in the base case, as observed AEs were mild, meaning that it is not expected that a significant detriment in HRQoL, or a significant increase in cost, would be associated with these events. This approach is in line with previous NICE appraisals of baricitinib and dupilumab in AD (TA681 and TA534).[7, 106]

Given the very low incidence of SAEs observed in the BRAVE-AA clinical trials ([redacted] of patients had ≥1 SAE), the impact of SAEs on the results was considered negligible. As such, the model does not include any specific SAEs.

B.3.3.5 Mortality

All-cause mortality was considered in the cost-effectiveness analysis based on Office for National Statistics lifetables between 2017–2019, to avoid the impact of the COVID-19 pandemic on these data.[100] Age- and gender-specific rates were combined to a blended rate, based on the proportion of men and women in the model, 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 data from clinical trials

As described in Section B.2.3, the BRAVE-AA trials assessed HRQoL using several different instruments up to Week 36, including EQ-5D-5L and HADS. However, in the base case analysis,

health state utility values were derived from the Adelphi DSP, a real-world evidence study in which EQ-5D-5L data were collected from patients with AA in Germany, Spain, Italy, France, and the UK. During the study, which was initiated in October 2021, physician-completed questionnaires were used to rate disease severity, while patients were invited to complete a questionnaire that included EQ-5D-5L. Utilities were generated via transforming the EQ-5D-5L to EQ-5D-3L values using the cross-walk algorithm by Hernandez *et al*., as recommended by NICE.[107] Given that the model captures the differential HRQoL gain associated with the achievement of $SALT_{50}$ and $SALT_{75}$ in the base case analysis (Section B.3.2.2), health state utility values (HSUV) were applied for both response levels. Patients in the Induction and the BSC health states were assigned a baseline utility.

The Adelphi DSP HRQoL data were used in the base case analysis as the EQ-5D data captured directly from the baricitinib trials were insensitive to changes in the severity of AA. This may partly be due to the fact that patients with significant uncontrolled neuropsychiatric disorders were excluded from the BRAVE-AA studies, meaning that patients suffering the most severe psychological impact of AA would have likely been excluded from the trial utility estimates. In addition, due to the lack of content validity of the EQ-5D instrument in this indication, the baseline utility values generated from the BRAVE-AA EQ-5D data are almost the same as the age and sex-adjusted utility value for the general population, resulting in a ceiling effect. For these reasons, the trial-based utilities are explored in scenario analyses, presented in Section B.3.11.3.

The lack of content validity of the EQ-5D instrument in AA is also likely to have affected the HRQoL data generated from the Adelphi DSP. Therefore, while utilities generated from the Adelphi DSP are more appropriate than those based on the trial entry criteria, the utility estimates from these data used in the model should still be considered conservative.

B.3.4.2 Mapping

No mapping was used in the base case analysis. A scenario analysis is presented where the HADS data collected in the baricitinib phase III studies was mapped to EQ-5D-3L based on a mapping algorithm reported in Brazier (2014).[108]

B.3.4.3 Health-related quality of life studies

As described in Section B.3.1, a *de novo* SLR was conducted to identify any cost-effectiveness, HSUVs and cost and healthcare resource use data for adult patients with severe AA. The SLR yielded two results related to utility data associated with baricitinib treatment of adults with AA. One study used a US tariff, and the second included a small number of patients (N=37) and reported no severity information for used in the model. In addition, both studies reported baseline utility values for AA that were higher than the age- and sex-adjusted utility value for the general population. For these reasons, the utilities values reported in these studies were not utilised as inputs for the cost-effectiveness model. Full details on the methods and results of this SLR are presented in Appendix H.

B.3.4.4 Adverse reactions

Disutilities associated with AEs are not included in the model since the AEs observed in the BRAVE-AA1 and BRAVE-AA2 trials were mild. Therefore, it is not expected that a significant detriment in HRQoL would be associated with these events. This approach is in line with the previous NICE appraisals in AD (TA534 and TA681).[7, 106]

B.3.4.5 Health-related quality of life data used in the cost-effectiveness analysis

Utility values were estimated based on the data collected from the baricitinib phase III trials and from the Adelphi DSP study, as described above. The baseline utility values generated from EQ-5D collected in the Adelphi DSP study are lower compared with the trial and the general population, and therefore the HRQoL improvement due to each response level can be more accurately captured. In addition, the Adelphi DSP study was conducted in 5 European countries, including the UK, whereas the baricitinib clinical trials did not include any European patients. The trial-based utilities are explored in scenario analyses. Utility values used in the base case analyses are shown in Table 52. Utility values generated from the EQ-5D and HADS data from the BRAVE-AA trials explored in scenario analyses are presented in Table 53 and Table 54, respectively.

Table 52. Utility values assigned to $SALT_{50}$ and $SALT_{75}$ response based on Adelphi EQ-5D-3L values

Population	Baseline		CFB for $SALT_{50}$		CFB for $SALT_{75}$	
	Value	SE	Value	SE	Value	SE
SALT 50–100	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CFB: change from baseline; SALT: severity of alopecia tool; SE: standard error.

Table 53. Utility values assigned to $SALT_{50}$ and $SALT_{75}$ response based on trial collected EQ-5D data

Population	Baseline		CFB for $SALT_{50}$		CFB for $SALT_{75}$	
	Value	SE	Value	SE	Value	SE
SALT 50–100	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CFB: change from baseline; SALT: severity of alopecia tool; SE: standard error.

Table 54. Utility values assigned to $SALT_{50}$ and $SALT_{75}$ response based on trial collected HADS data mapped to EQ-5D-3L

Population	Baseline		CFB for $SALT_{50}$		CFB for $SALT_{75}$	
	Value	SE	Value	SE	Value	SE
SALT 50–100	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CFB: change from baseline; SALT: severity of alopecia tool; SE: standard error.

Age adjustment

Utility values included in the model were age-adjusted based on the HSE 2014 dataset as recommended in the latest NICE guidance.[109] That is, age-and gender-specific disutilities are estimated for each year from the general population utility values to account for the loss of HRQoL due to aging over time.

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

As described in B.3.1, an SLR was conducted to identify any relevant cost or resource use data for adult patients with severe AA. No studies featuring relevant cost and resource use data

associated with the treatment of adult patients with severe AA in the UK were identified. Full details of the SLR methods and results are reported in Appendix I.

The following cost categories are included in the model:

- Drug acquisition and 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 UK-specific sources such as the National schedule of NHS costs,[101] NHS Drug Tariff,[102] NHS wigs and fabric supports costs[103], and Personal Social Services Research Unit (PSSRU).[104]

B.3.5.1 Intervention and comparators' costs and resource use

B.3.5.1.1 Drug acquisition and administration costs

Costs associated with baricitinib and 'watch and wait'

In the model, patients incur treatment and monitoring costs which differ depending on treatment and health state. The resource use related to the administration and monitoring of treatment was aligned with local clinical guidelines and routine clinical practice.

Drug acquisition costs, dose and frequency were based on approved doses obtained from Summary of Product Characteristics (SmPC). Table 55 presents the treatment acquisition costs for included treatments in the model. A patient access scheme (PAS) discount has been applied to the cost of baricitinib 4mg. Based on the estimates reported in Table 55, a drug acquisition cost for baricitinib 4mg of £[redacted] in the Induction state and of £[redacted] in the Maintenance state per year is estimated.

Table 55. Drug acquisition costs for baricitinib 4 mg and 'watch and wait' (PAS price)

Treatment	Number of doses in induction (36 weeks)	Number of doses in maintenance (annual)	Cost per pack (with PAS)	Units in pack	Frequency of administration	Reference
Baricitinib 4 mg	252	365	£[redacted]	28	QD	Eli Lilly
'Watch and wait'	N/A	N/A	£0.00	N/A	N/A	Assumption

Abbreviations: N/A: not applicable; PAS: patient access scheme; QD: once daily.

Costs associated with BSC treatments

When patients transition to the BSC health state, they are assumed to receive additional therapy which is comprised of various BSC treatments. The composition of the BSC treatments and the proportion of patients that are assumed to receive such treatments are based on the Adelphi DSP study in the base case and are presented in Table 56. Input for the BSC composition is also available from the clinical experts and is explored in a scenario analysis.

Unit costs were taken from the NHS Drug Tariff [102], whereas the average daily dosage was obtained from the summary of product characteristics (SmPC). The total costs for each pharmacological intervention consist of drug acquisition and monitoring costs (Table 56 and Table 57, respectively). The annual drug acquisition cost for each intervention is calculated by multiplying the pack cost, the proportion of patients (Adelphi DSP in the base case), the number of doses per year and dividing by the number of units per pack.

Table 56. Drug acquisition costs for BSC

Item	Dose and frequency **	Pack size	Units per pack	Pack cost [+]	Number of doses per year	Proportion of patients in BSC		Annual cost
						Adelphi DSP (base case)	KOL input (scenario analysis)	
Ciclosporin	4 mg/kg QD	50mg, 30 capsules	1,500	£35.97	108,114 (4mg * 74kg * 365.25 days)	12.39%	12.50%	£321.12
Methotrexate	20 mg per week	2.5mg, 28 tablets	70	£1.70	1,040 (20mg * 52 weeks)	14.29%	7.50%	£3.61
Azathioprine	2 mg/kg body weight QD, for 1 year	25mg, 28 tablets	700	£1.56	54,057 (2mg * 74kg * 365.25 days)	2.86%	8.67%	£3.44
Intralesional steroids (triamcinolone acetonide)	5 mg repeated every other week	40mg/1ml, 5 vials	200	£7.45	130 (5mg * 26 weeks)	9.53%	30.83%	£0.46
DPCP (contact immunotherapy) treatment	Weekly treatment for 9 months	NA	1	£114.40	36 (4 times per month for 9 months)	20.96%	27.50%	£863.27
Prednisolone	0.4 mg/kg QD	2.5mg, 28 tablets	70	£3.91	10,811 (0.4mg * 74kg * 365.25 days)	17.15%	25.00%	£103.57
TCS: Mometasone ointment	32 g QD for six months	0.1%, 100g	100	£7.33	5,844 (32g * 183 days)	24.77%	63.33%	£106.12
Minoxidil 5% foam (topical)	1 g BID (women) or 1 g QD (men) - discontinue if no improvement after 16 weeks (women) or 24 weeks (men)	5%, 180g	180	£50.01	202 [60.7%[§] *(2g * 7 days * 16 weeks) + 39.3%[§] *(1g * 7 days * 24 weeks)]	5.72%	37.50%	£3.21
Minoxidil tablets	20 mg QD	10mg, 60 tablets	600	£30.68	7,305 (20mg * 365.25 days)	0.00%	7.50%	£0.00
Mycophenolate Mofetil	1 g BID, for 1 year	50 mg, 500 tablets	25000	£5.49	730,500 (2,000mg * 365.25 days)	2.86%	0.00%	£4.59
Anthralin 0.1% cream	1.5 g QD	0.1%, 50g	50	£3.77	242 (1.5g * 7 days * 23 weeks)	5.72%	0.00%	£1.04

Patients not currently on treatment	NA	NA	NA	£0.00	NA	13.00%	0.00%	£0.00
Total cost	**£1,410.42**							

Footnotes: **based on SmPC. * Source: NHS Drug Tariff[102], apart from DPCP treatment.[110] § The percentages of females (60.7%) and males (39.3%) reported above are specific to the SALT 50-100 population (base case), as reported in Table 46.
Abbreviations: BID: twice per day; DPCP: diphenylcyclopropenone; DSP: disease specific programme; KOL: key opinion leader; QD: once per day; SmPC: summary of product characteristics; TCS: topical corticosteroids.

B.3.5.1.2 Monitoring costs and resource use

Costs associated with baricitinib and 'watch and wait'

Included monitoring and related resource use items are shown in Table 57, based on the feedback received by clinical experts. Frequency was stratified by Induction (reflecting the induction period resource use), and Maintenance (annual frequency for responders).

Table 57. Annual frequency of physician visits and monitoring tests per treatment

Item	Baricitinib 4 mg		'Watch and wait'		Reference
	Induction (36 weeks)	Maintenance (annual)	Induction (36 weeks)	Maintenance (annual)	
Consultation visits					
Dermatologist outpatient consultation	4.0	2.0	4.0	2.0	Clinician expert opinion
Dermatologist nurse visit	1.0	0.5	1.0	0.5	Clinician expert opinion
Tests and investigations					
Full blood count	0.0	4.0	0.0	0.0	NICE (2020)[7]
Other					
Wig use (modacrylic wig)	2.0	0.0	2.0	0.0	Clinician expert opinion
Orthotics	1.0	0.0	1.0	0.0	Clinician expert opinion

Disease management unit costs for included treatments are presented in Table 58. Costs were sourced from UK-specific available sources.

Table 58. Disease management unit costs

Item	Unit cost	Description unit cost	Proportion of patients *	Reference (cost)
Consultation visits				
Dermatologist outpatient consultation	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	100%	NHS 2019/20 National Cost Collection data (2020)[101]
Dermatologist nurse visit	£28.25	PSSRU, 15 minutes of hospital nurse Band 6 patient related time	100%	NHS 2019/20 National Cost Collection data (2020)[101]
Tests and investigations				
Full blood count	£2.55	DAPS05 - Haematology	100%	NHS 2019/20 National Cost

				Collection data (2020)[101]
Other				
Wig use (modacrylic wig)	£75.70	Wigs and fabric supports on the NHS	80%	NHS Wigs and fabric supports on the NHS (2020)[103]
Orthotics for wig fitting	£132.00	Service Code 658 - Total Outpatient Attendances	80%	NHS 2019/20 National Cost Collection data (2020)[101]

Notes: * based on clinician expert opinion.
Abbreviations: NHS, National Health Service; PSSRU, personal social services research unit.

Based on the resource use estimates and the associated unit costs reported in Table 57 and Table 58, the total disease management costs for baricitinib 4mg and 'watch and wait' are displayed in Table 59.

Table 59. Total disease management costs for induction and maintenance

Item	Cost for baricitinib 4 mg		Cost for 'watch and wait'	
	Induction (36 weeks)	Maintenance (Annual)	Induction (36 weeks)	Maintenance (Annual)
Dermatologist outpatient consultation	£499.16	£249.58	£499.16	£249.58
Dermatologist nurse visit	£28.25	£14.13	£28.25	£14.13
Full blood count	£0.00	£10.20	£0.00	£0.00
Wig use	£121.12	£0.00	£121.12	£0.00
Orthotics	£105.60	£0.00	£105.60	£0.00
Total cost	£754.13	£273.91	£754.13	£263.71

Costs associated with BSC treatments

Monitoring costs and health-care resource use for the BSC health state are presented in Table 60. Resource use has been obtained from clinical experts during the model development phase, whereas the proportion of patients on disease management costs is taken from either the Adelphi DSP study (dermatologist visits) or from clinical experts (tests, wig use and orthotics). Costs were sourced from publicly available sources. Based on the resource use estimates and the associated unit costs reported in Table 60, the total annual disease management costs associated with BSC are £354.20.

Table 60. Annual frequency and disease management costs associated with BSC

Resource	Unit cost	Description unit cost	Proportion of patients	Annual frequency*	Annual cost	Reference (cost)
Consultation visits						
Dermatologist outpatient consultation	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	13%**	2.00	£32.45	NHS (2020)[101]
Dermatologist nurse visit	£28.25	PSSRU, 15 minutes of hospital nurse Band 6	13%**	0.50	£1.84	NHS (2020)[101]

		patient related time				
Tests and investigations						
Thyroid function	£2.55	DAPS05 - Haematology	100%*	4.00	£10.20	NHS (2020)[101]
Vitamin D	£2.55	DAPS05 - Haematology	100%*	4.00	£10.20	NHS (2020)[101]
Ferritin	£2.55	DAPS05 - Haematology	100%*	4.00	£10.20	NHS (2020)[101]
Full blood count	£2.55	DAPS05 - Haematology	100%*	4.00	£10.20	NHS (2020)[101]
Liver function	£1.20	DAPS04 - Clinical biochemistry	100%*	4.00	£4.80	NHS (2020)[101]
Renal function	£1.20	DAPS04 - Clinical biochemistry	100%*	4.00	£4.80	NHS (2020)[101]
Tuberculosis	£8.15	DAPS07 - Microbiology	100%*	4.00	£32.60	NHS (2020)[101]
Lipids	£2.55	DAPS05 - Haematology	100%*	4.00	£10.20	NHS (2020)[101]
Other						
Wig use (modacrylic wig)	£75.70	Wigs and fabric supports on the NHS	80%*	2.00	£121.12	NHS (2020)[103]
Orthotics	£132.00	Service Code 658 - Total Outpatient Attendances	80%*	1.00	£105.60	NHS (2020)[101]
Total costs	**£354.20**					

Notes: *based on clinical expert opinion; ** .
Abbreviations: GBP, great British pound; NA, not applicable; NHS, National Health Service; PSSRU, personal social services research unit.

The annual monitoring cost is calculated by multiplying the visit unit cost, the frequency of the visits, and the proportion of patients (Adelphi DSP is used as the base case). For DPCP treatment and intralesional steroids monitoring, the cost of a minor skin procedure was used from the NHS reference costs,[101] whereas for the remaining BSC treatments the cost of a dermatologist outpatient consultation was used. The frequency of visits is based on KOL input and is presented in Table 61 below.

The annual cost for each BSC treatment is calculated by summing the annual drug acquisition costs and the annual monitoring costs and then multiplying by the percentage of patients that utilise each treatment. Based on the resource use estimates and the associated unit costs reported below, the total annual drug acquisition costs associated with BSC are £1,410.42 (Table 56), whereas the total annual drug monitoring costs associated with BSC are £2,272.68 (Table 26). The total annual drug acquisition and monitoring costs associated with BSC are £3,683.10.

Table 61. Drug monitoring costs for BSC

Item	Unit cost [+]	Description unit cost	Proportion of patients in BSC		Frequency of visits per year *	Annual cost
			Adelphi DSP (base case)	KOL input (scenario analysis)		
Ciclosporin	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	12.39%	12.50%	9 (weekly visits for the first 4 weeks, monthly visit for the next 2 months, then 1 every 3 months)	£139.11
Methotrexate	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	14.29%	7.50%	9 (weekly visits for the first 4 weeks, monthly visit for the next 2 months, then 1 every 3 months)	£160.51
Azathioprine	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	2.86%	8.67%	9 (assumed same as cyclosporin and methotrexate)	£32.10
Intralesional steroids (triamcinolone acetonide)	£156.56	JC43C – OPROC – Minor Skin Procedures, 19 years and over – Dermatology	9.53%	30.83%	18 (1 session every 6 weeks for 6 sessions, repeated 3 times)	£268.50
DPCP (contact immunotherapy) treatment	£156.56	JC43C – OPROC – Minor Skin Procedures, 19 years and over – Dermatology	20.96%	27.50%	36 (weekly visits for up to 9 months depending on response)	£1,181.41
Prednisolone	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	17.15%	25.00%	13 (every 4-6 weeks)	£278.22
TCS: Mometasone ointment	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	24.77%	63.33%	4 (every 3 months)	£123.65
Minoxidil 5% foam (topical)	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	5.72%	37.50%	4 (every 3 months)	£28.54
Minoxidil tablets	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	0.00%	7.50%	4 (every 3 months)	£0.00

Mycophenolate Mofetil	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	2.86%	0.00%	9 (assumed same as cyclosporin and methotrexate)	£32.10
Anthralin 0.1% cream	£124.79	Weighted average of WF01A-D and WF02A-D - Dermatology	5.72%	0.00%	4 (assumed same as TCS)	£28.54
Patients not currently on treatment	£0.00	NA	13.00%	0.00%	0	£0.00
Total cost	**£2,272.68**					

Notes: [+] National Cost Collection for the NHS [111] . * Based on clinician expert opinion.
Abbreviations: DPCP, diphenylcyclopropenone; DSP, disease specific programme; NA, not applicable; TCS, topical corticosteroids.

B.3.5.2 Health-state unit costs and resource use

Costs and resource use associated with different health states are described in section B.3.5.1.

B.3.5.3 Adverse reaction unit costs and resource use

Costs and resource use associated with adverse reactions were not included in the base case, as described above.

B.3.5.4 Miscellaneous unit costs and resource use

Costs associated with baricitinib and 'watch and wait'

Since AA is associated with a significant psychological and emotional distress, the psychological burden of the disease was included in the model. Costs of managing the psychological burden of AA were based on NICE guidance for depression in adults.[105] Overall, these costs can be divided into non-pharmacological and pharmacological treatments costs. Following clinical expert opinion, the resource use in the Maintenance period is estimated to be zero due to the patient improvement in emotional symptoms following hair regrowth, thus the costs for psychological burden only occur in during Induction and BSC.

Table 62 presents the costs for non-pharmacological management of the psychological burden of AA included in the model for baricitinib 4 mg and 'watch and wait'. Following the approach used in the NICE guidance for depression, in addition to therapists' time, the intervention costs of all psychological therapies included an initial general practitioner (GP) visit for referral to psychological services.[105] It is acknowledged that this assumption (100% GP referral to psychological services) is a conservative estimate, as a proportion of people with a new episode of depression may self-refer to psychological services. On the other hand, it is possible that some of the people self-referring may have consulted their GP prior to self-referral. The resource use and proportion of patients with AA that would utilise each type of psychological support service was based on clinical expert opinion.

Table 62. Costs of non-pharmacological management of the psychological burden of alopecia areata.

Item	Unit cost	Description	Proportion of patients*	Resource use in induction*	Cost in induction
Psychiatrist visit	£112.00	NICE GID Table 86 - band 7 HI therapist (with MBCT qualification).	5.00%	3.00	£16.80
Psychologist visit	£50.00	NICE GID Table 85 - One-hour direct contact (band 5 PWP).	10.00%	3.00	£15.00
Self-help with support	£39.23	1 GP session.	12.38%	0.75	£3.64
Group exercise+	£186 + £39.23	30 sessions x 1 hour each; 1 therapist (band 5 PWP) and 8 participants per group = 30 therapist.	0.75%	0.75	£1.27
Interpersonal psychotherapy+	£873 + £39.23	8 sessions x 1 hour each = 8 therapist hours per service user (band 7 HI therapist).	0.75%	0.75	£5.13
Counselling+	£873 + £39.23	12 sessions x 1 hour each = 12 therapist hours per service user (band 7 HI therapist).	1.13%	0.75	£7.70
Total costs	**£49.54**				

Footnote: * based on clinician expert opinion. [+]Cost of psychological intervention plus 1 GP referral visit, at a GP unit cost £39.23 per patient contact lasting 9.22 minutes;[104] cost of psychological intervention based on resource use combined with unit cost of the appropriate level of therapist, estimated as described in Table 85, Table 86 and Table 87 in NICE guidance CG90 for depression in adults.[105]
Abbreviations: HI: high intensity; PWP: psychological well-being practitioner.

Table 63 presents the costs of pharmacological interventions for the management of the psychological burden of AA for baricitinib 4mg and 'watch and wait' based on the NICE guideline CG90 for the management of depression.[105] The total costs for each pharmacological intervention consist of drug acquisition and GP visit costs, comprising of 4 GP visits based on the committee's expert advice in the NICE guideline for the management of depression.[105] The unit cost of GP care was sourced at £39.23 per patient contact lasting 9.22 minutes.[104]

The drug acquisition costs were taken from NHS Drug Tariff,[102] whereas the average daily dosage was taken from the NICE guideline for the management of depression.[105] Pharmacological treatment was administered over 12 weeks as it was assumed that at the end of this period adults with less severe depression would have achieved remission. This is a conservative assumption as it does not account for adults with severe depression that would continue maintenance pharmacological treatment. Gradual discontinuation (tapering) beyond 12 weeks was not allowed in the model. However, it was assumed that patients would be re-treated at the start of each year in the maintenance phase of the economic model. The proportion of patients with AA that would utilise each pharmacological intervention was based on clinical expert opinion.

Based on the resource use estimates and the associated unit costs reported in Table 62 and Table 63, the total annual costs of psychological burden for baricitinib 4 mg and 'watch and wait' are £95.52. As with the approach for the non-pharmacological management of the psychological burden of AA and following clinical expert advice, these costs only occur during Induction and BSC and the resource use in maintenance period is zero.

Table 63. Costs of pharmacological treatment management of the psychological burden of alopecia areata

Item	Proportion of patients*	Resource use in induction*	Mean daily dosage	Drug acquisition cost	12-week drug cost	Total cost **	Cost in induction	Reference (cost)
Sertraline	16.50%	0.75	50% 50mg; 25% 100mg; 15% 150mg; 10% 200mg	50mg, 28 tab, £1.04 100mg, 28 tab, £1.16	£4.50	£161.42	£19.98	Table 83 in NICE CG90[105]
Escitalopram	16.50%	0.75	80% 10mg; 20% 20mg	10mg, 28 tab, £1.07 20mg, 28 tab, £1.29	£3.34	£160.26	£19.83	
Duloxetine	5.00%	0.75	80% 60mg; 20% 120mg	60mg, 28 caps, £2.13	£7.67	£164.59	£6.16	
GP care (4 visits)	N/A	N/A	N/A	N/A	N/A	£156.92	N/A	
Total costs	**£45.98**							

Footnotes: * based on clinician expert opinion. ** The total cost for each pharmacological intervention is calculated by summing the 12-week drug cost and GP care.
Abbreviations: caps: capsules; GP: general practitioner; N/A: not applicable; NICE: National Institute for Health and Care Excellence; tab: tablet.

Costs associated with BSC treatments

The costs of managing the psychological burden of AA were included in the model for patients receiving BSC. The same types of pharmacological and non-pharmacological costs with those for patients in induction were considered for patients that receive BSC. The resource use and the proportion of patients that would receive non-pharmacological and pharmacological treatment options for psychological burden were based on clinical expert opinion. The same proportion of patients with induction was assumed to utilise pharmacological and non-pharmacological costs in the BSC state, whereas the resource frequency was slightly higher in the BSC state to reflect the longer time that patients spend in BSC compared to induction (i.e. 52 weeks vs 36 weeks).

Table 64 presents the costs for non-pharmacological management of psychological burden included in the model for BSC and Table 65 presents the costs of pharmacological interventions for the management of psychological burden for BSC. The total annual costs of managing the psychological burden of AA for BSC are £127.35.

Table 64. Costs of non-pharmacological management of the psychological burden of alopecia areata in BSC

Items	Unit cost	Description	Proportion of patients*	Resource use*	Annual cost
Psychiatrist visit	£112.00	NICE GID Table 86 - band 7 HI therapist (with MBCT qualification).	5.00%	4.00	£22.40
Psychologist visit	£50.00	NICE GID Table 85 - One-hour direct contact (band 5 PWP).	10.00%	4.00	£20.00
Self-help with support	£39.23	1 GP session.	12.38%	1.00	£4.85
Group exercise+	£186 + £39.23	30 sessions x 1 hour each; 1 therapist (band 5 PWP) and 8 participants per group = 30 therapist.	0.75%	1.00	£1.69
Interpersonal psychotherapy+	£873 + £39.23	8 sessions x 1 hour each = 8 therapist hours per service user (band 7 HI therapist).	0.75%	1.00	£6.84
Counselling+	£873 + £39.23	12 sessions x 1 hour each = 12 therapist hours per service user (band 7 HI therapist).	1.13%	1.00	£10.26
Total costs	**£66.05**				

Footnotes: * based on clinician expert opinion. +Cost of psychological intervention plus 1 GP referral visit, at a GP unit cost £39.23 per patient contact lasting 9.22 minutes;[104] cost of psychological intervention based on resource use combined with unit cost of the appropriate level of therapist, estimated as described in Table 85- 87 in NICE guidance CG90 for depression in adults.[105]
Abbreviations: CBT: cognitive behavioural therapy; HI: high intensity; IPT: interpersonal psychotherapy; LS: less severe; MBCT: Mindfulness-based cognitive therapy; MS: more severe; NA: not applicable; PWP: psychological well-being practitioner.

Table 65. Costs of pharmacological treatment management of the psychological burden of alopecia areata in BSC

Items	Proportion of patients *	Annual frequency *	Mean daily dosage	Drug acquisition cost	12-week drug cost	Total cost**	Annual cost	Reference (cost)
Sertraline	16.50%	1.00	50% 50mg; 25% 100mg; 15% 150mg; 10% 200mg	50mg, 28 tab, £1.04 100mg, 28 tab, £1.16	£4.50	£161.42	£26.63	Table 83 of document "Evidence review B"[105]
Escitalopram	16.50%	1.00	80% 10mg; 20% 20mg	10mg, 28 tab, £1.07 20mg, 28 tab, £1.29	£3.34	£160.26	£26.44	
Duloxetine	5.00%	1.00	80% 60mg; 20% 120mg	60mg, 28 caps, £2.13	£7.67	£164.59	£8.23	
GP care (4 visits)	N/A	N/A	N/A	N/A	N/A	£156.92	N/A	
Total costs							**£61.31**	

Footnotes: *based on clinician expert opinion. ** The total cost for each pharmacological intervention is calculated by summing the 12-week drug cost and GP care.
Abbreviations: caps: capsules; GP: general practitioner; N/A: not applicable; tab: tablets.

B.3.6 Severity

This technology does not meet the criteria for a severity weight.

B.3.7 Uncertainty

Neither the condition, nor the technology impact the ability to generate high-quality clinical evidence.

As discussed in Section B.2.6.3, AA is characterised by non-scarring hair loss that, unlike some other dermatological conditions, does not usually cause physical symptoms (beyond hair loss) or disability.[1, 72] The impact of AA on HRQoL is instead attributed to the significant psychological distress caused by hair loss.[2, 11, 13, 47] Owing to this mono-symptomatic aspect of AA, the five dimensions of health covered by the generic EQ-5D instrument, comprised of mobility, self-care, usual activities, pain/discomfort and anxiety/depression domains, do not adequately capture the dimensions of HRQoL that are affected by AA (in this case the psychological aspects), demonstrating a lack of content validity for the EQ-5D instrument in AA.[73, 74] Similar limitations have been reported from a recent trial funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme to treat vitiligo.[75]

B.3.8 Managed access proposal

It is anticipated that the appraisal will result in routine commissioning and therefore no managed access proposal is required.

B.3.9 Summary of base-case analysis inputs and assumptions

B.3.9.1 Summary of base-case analysis inputs

A summary of the base case model inputs and settings are presented in Table 66.

Table 66: Summary of variables applied in the economic model base case

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty (distribution)	Reference to section in submission
Model characteristics			
Time horizon	Lifetime	NA	B.3.2
Cycle length	4 weeks	NA	
Discount rate effects	3.5%	NA	
Discount rate costs	3.5%	NA	
Patient characteristics			
Mean starting age, years	37.52	SE 0.372	B.3.2.1
Proportion male, %	39.33	SE 1.410	
Efficacy data			
Response	Baricitinib trial data	SE	B.3.3.2

Drug costs			
Baricitinib 4 mg (Induction)	£[redacted]	Assumed to be 10% of the mean (gamma)	B.3.5.1
Baricitinib 4 mg (Maintenance)	£[redacted]		
'Watch and wait'	£0.00		
BSC treatment (drug costs)	£1,410.42		
BSC treatment (drug monitoring)	£2,272.68		
Disease management costs			
Baricitinib 4 mg (Induction)	£754.13	Assumed to be 10% of the mean (gamma)	B.3.5.1
Baricitinib 4 mg (Maintenance)	£273.91		
'Watch and wait' (Induction)	£754.13		
'Watch and wait' (Maintenance)	£263.71		
BSC treatment	£354.20		
Costs associated with managing psychological burden of AA			
Non-pharmacological management (Induction)	£49.54	Assumed to be 10% of the mean (gamma)	B.3.5.1
Pharmacological management (Induction)	£45.98		
Non-pharmacological management (BSC)	£66.05		
Pharmacological management (BSC)	£61.31		

Abbreviations: AA: alopecia areata; BSC: best supportive care.

B.3.9.2 Assumptions

A list of the assumptions made in the base case analysis and their justification is provided in Table 67 where appropriate, the exploration of the potential impact of these assumptions in a scenario analysis is noted.

Table 67: Key modelling assumptions

Assumption	Description of assumption for the base case	Justification	Addressed in scenario analysis
Response definition	Based on $SALT_{50}$	This response definition is anticipated to be clinically relevant in NHS clinical practice, as it will capture sufficient clinical benefit to justify continuing treatment beyond the trial period. In	The use of $SALT_{75}$ is explored in scenario analyses

		addition, selection of the $SALT_{50}$ response in the model captures the differential HRQoL gain associated with the achievement of $SALT_{50}$ versus $SALT_{75}$.	
Response maintenance	Patients maintain their end of Induction period response until they discontinue treatment.	Loss of response is assumed to result in treatment discontinuation and is reflected in the all-cause discontinuation rate. This is a conservative assumption given that no utility gain will be captured for patients moving from $SALT_{50}$ to $SALT_{75}$ during maintenance.	No
Discontinuation rate during induction	The discontinuation rate applied during the Induction phase encompasses all causes except for loss of efficacy.	To avoid double counting efficacy, as efficacy is captured by response assessment measured at the end of the Induction period.	No
Discontinuation rate during maintenance	Patients discontinue treatment in the maintenance phase utilising the all-cause discontinuation rates from the baricitinib phase III studies, with data obtained from Week 0 to 36 for the placebo arm and Week 0 to 52 for the baricitinib arm	Lack of discontinuation rates for the placebo arm in the maintenance phase.	No
BSC	Assumed that patients in the BSC state remain in that state until the end of model simulation or death	Reflects patients who have failed treatment with baricitinib or failed to spontaneously recover during a 'watch and wait' approach and may go on to receive one or more of a basket of BSC therapies	No
Adverse events risk	Assumed to remain constant over the treatment duration	Assumption appears reasonable given the lack of long-term data from the baricitinib phase III trial programmes.	No
Health state utility over time	Assumed to decline with age, with the model applying an age adjustment factor based on the Health Survey for England (HSE) 2014 dataset[109]	Assumption is recommended by the latest NICE guidance and based on a well-established UK literature source preferred by NICE in many previous appraisals.	No
Treatment adherence	Adherence to treatment was not modelled separately	Reflects that compliance rates are high in the	No

		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.	
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 baricitinib incurs all costs of treatment, but not the benefits of treatment during the trial period.	No

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; SALT: severity of alopecia tool.

B.3.10 Base-case results

B.3.10.1 Base-case incremental cost-effectiveness analysis results

A summary of results in the base case analysis are presented in Table 68 (ICER at PAS price) and Table 69 (net health benefit).

At PAS price, baricitinib 4 mg and 'watch and wait' accumulated costs of £[redacted] and £[redacted], and total QALYs of [redacted] and [redacted], respectively. The with-PAS ICER was within the range considered cost-effective; at £29,111/QALY, it falls below the NICE WTP threshold of £30,000. These results demonstrate baricitinib to be a cost-effective option for the treatment of patients with severe AA versus a 'watch and wait' treatment strategy, the comparator relevant to UK.

Disaggregated deterministic results of the base case incremental cost-effectiveness analysis are presented in Appendix J.

Table 68. Base case cost-effectiveness results at baricitinib PAS price (probabilistic)

	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
'Watch and wait'	£[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4mg	£[redacted]	22.60	[redacted]	£[redacted]	0.00	[redacted]	£29,111

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life-years gained; PAS: patient access scheme; QALY: quality-adjusted life year.

Table 69. Net health benefit

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	NHB at £20,000	NHB at £30,000
'Watch and wait'	£[redacted]	[redacted]	-	-	-	-
Baricitinib 4mg	£[redacted]	[redacted]	£[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; NHB: net health benefit

B.3.11 Exploring uncertainty

B.3.11.1 Probabilistic sensitivity analysis (PSA)

A probabilistic sensitivity analysis (PSA) was run with 1,000 Monte Carlo simulations in order to assess the uncertainty associated with model input parameters. Use of 1,000 iterations was deemed appropriate based on the results of an ICER convergence test, shown in Figure 27.

Figure 27. Convergence plot for NMB

Abbreviations: CI: confidence interval; NMB, net monetary benefit; PSA: probabilistic sensitivity analysis; QALY: quality-adjusted life year.

A visual representation of the PSA results comparing baricitinib 4mg and placebo is provided in the cost-effectiveness plane (see Figure 28 below). Each dot resembles one Monte Carlo simulation where the input parameters are sampled from the distributions in a total of 1,000 loops. The results of the cost-effectiveness plane show moderate uncertainty with regards to the extent of the additional costs for baricitinib 4 mg compared with 'watch and wait', but little uncertainly with regards to the existence of these additional costs, as most dots fall on the North quadrants of the plane. On the effectiveness side, PSA results show moderate uncertainty with regards to the extent of additional benefits.

Figure 28. Cost-effectiveness plane for baricitinib 4 mg compared with 'watch and wait'

Abbreviations: QALY, quality-adjusted life year.

The cost-effectiveness acceptability curve shows a [redacted]% probability of baricitinib 4 mg being cost-effective compared with 'watch and wait' at a cost-effectiveness threshold of £30,000/QALY (Figure 29), with the probability increasing with higher thresholds.

Figure 29. Cost-effectiveness acceptability curve for baricitinib 4mg compared with placebo

Footnotes: "placebo" in the figure represents "watch and wait"

B.3.11.2 Deterministic sensitivity analysis (DSA)

A deterministic one-way sensitivity analysis (OWSA) has been performed and the ten most important drivers of the model have been plotted in a tornado diagram (Figure 30). The three most influential parameters in the model were the HSUV for $SALT_{75}$, the proportion of patients on DPCP treatment in the BSC state and the discontinuation rate for baricitinib 4 mg in the Maintenance phase. Another driver of the model was the response rate at 36 weeks using $SALT_{50}$ for baricitinib 4 mg and the resource frequency for the monitoring cost of DPCP treatment in the BSC state.

Figure 30. Tornado diagram for baricitinib 4 mg compared with 'watch and wait' at PAS price

Footnotes: "placebo" in the figure represents "watch and wait"
Abbreviations: BSC: best supportive care; DPCP: diphenylcyclopropenone; HSUV: health state utility value; SALT: severity of alopecia tool.

B.3.11.3 Scenario analysis

A number of scenario analysis were explored in which model assumptions or parameters were altered. The rationale and results of the scenario analyses carried out are presented below.

B.3.11.3.1 Scenario 1: Starting population with SALT 50–94

A summary of the baseline demographic characteristics, treatment response and discontinuation inputs for the SALT 50–94 population are provided in Table 47, Table 49 and Table 51, respectively. The cost-effectiveness outcomes of the scenario using a population with baseline SALT 50–94 are presented in Table 70 below. On average, a patient on baricitinib 4 mg accumulated [redacted] QALYs (discounted) compared to [redacted] QALYs on 'watch and wait', which amounts to an incremental difference of [redacted] QALYs. Total treatment costs were [redacted] for baricitinib 4 mg and [redacted] in the 'watch and wait' arm of the model; an incremental difference of [redacted]. The ICER increased from the base case analysis from £29,111/QALY to £35,533/QALY, indicating that baricitinib is less cost-effective in the subgroup of severe patients compared to the ITT population.

Table 70. Cost-effectiveness results Scenario 1: Starting population with SALT 50–94

	Total			Incremental			ICER versus baseline (£/QALY)
Technologies	**Costs (£)**	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
'Watch and wait'	[redacted]	22.45	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.45	[redacted]	[redacted]	0.00	[redacted]	£35,533

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year; WTP, willingness to pay.

B.3.11.3.2 Scenario 2: Staring population with SALT 95–100

A summary of the baseline demographic characteristics, treatment response and discontinuation inputs for the SALT 95–100 population are provided in Table 47, Table 49 and Table 51, respectively. The cost-effectiveness outcomes of the scenario using a population with baseline SALT 95–100 are presented in Table 71 below. On average, a patient on baricitinib 4 mg accumulated [redacted] QALYs (discounted) compared to [redacted] QALYs on 'watch and wait', which amounts to an incremental difference of [redacted] QALYs. Total treatment costs were [redacted] for baricitinib 4 mg and [redacted] in the 'watch and wait' arm of the model; an incremental difference of [redacted]. The ICER decreased from the base case analysis from £29,111/QALY to £24,268/QALY, indicating that baricitinib is more cost-effective in the subgroup of very severe patients compared to the ITT population.

Table 71. Cost-effectiveness results Scenario 2: Starting population with SALT 95-100

	Total			Incremental			ICER versus baseline (£/QALY)
Technologies	**Costs (£)**	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
'Watch and wait'	[redacted]	22.73	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.73	[redacted]	[redacted]	0.00	[redacted]	£24,268

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year; WTP, willingness to pay.

B.3.11.3.3 Scenario 3: Response based on SALT$_{75}$

A summary of the treatment response inputs for a SALT$_{75}$ response is provided in Table 48. The cost-effectiveness outcomes of the scenario using a response level of SALT$_{75}$ are presented in Table 72 below. On average, a patient on baricitinib 4 mg accumulated [redacted] QALYs (discounted) compared to [redacted] QALYs on 'watch and wait', which amounts to an incremental difference [redacted] QALYs. Total treatment costs were [redacted] for baricitinib 4 mg and [redacted] in the 'watch and wait' arm of the model; an incremental difference of £[redacted]. The ICER decreased from the base case analysis from £29,111/QALY to £24,049/QALY, indicating that baricitinib is more cost-effective when response is defined by SALT$_{75}$ compared to SALT$_{50}$.

Table 72. Cost-effectiveness results Scenario 3: Response based on SALT$_{75}$

	Total			Incremental			ICER versus baseline (£/QALY)
Technologies	**Costs (£)**	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
'Watch and wait'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	£[redacted]	0.00	[redacted]	£24,049

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year; WTP, willingness to pay.

B.3.11.3.4 Scenario 4: Utilities based on EQ-5D data from baricitinib phase III studies

A summary of the utility inputs using the EQ-5D data from the phase II trials is provided in Table 53. The cost-effectiveness outcomes of the scenario using utilities based on EQ-5D data from

the baricitinib phase III studies are presented in Table 73 below. On average, a patient on baricitinib 4 mg accumulated [redacted] QALYs (discounted) compared to [redacted] QALYs on 'watch and wait', which amounts to an incremental difference of [redacted] QALYs. Total treatment costs were [redacted] for baricitinib 4 mg and [redacted] in the 'watch and wait' arm of the model; an incremental difference of [redacted]. The ICER increased from the base case analysis from £29,111/QALY to £217,186/QALY. This indicates that baricitinib is less cost-effective when utilities are estimated from the EQ-5D data collected in the baricitinib phase III studies compared to the EQ-5D data collected in the Adelphi DSP study, though due to the unsuitability of the EQ-5D data generated from the BRAVE-AA trials, this result is not relevant to decision making and should be interpreted with caution.

Table 73. Cost-effectiveness results Scenario 4: Utilities based on EQ-5D data from baricitinib phase III studies

	Total			Incremental			ICER versus baseline (£/QALY)
Technologies	**Costs (£)**	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
'Watch and wait'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	£217,186

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year; WTP, willingness to pay.

B.3.11.3.5 Scenario 5: Utilities based on HADS data mapped to EQ-5D-3L data collected from baricitinib phase III studies

A summary of the utility inputs using the HADS data from the phase III studies mapped into EQ-5D is provided in Table 54. The cost-effectiveness outcomes of the scenario using utilities based on HADS data from the baricitinib phase III studies are presented in Table 74 below. On average, a patient on baricitinib 4 mg accumulated [redacted] QALYs (discounted) compared to [redacted] QALYs on 'watch and wait', which amounts to an incremental difference of [redacted] QALYs. Total treatment costs were [redacted] for baricitinib 4 mg and [redacted] in the 'watch and wait' arm of the model; an incremental difference of [redacted]. The ICER increased from the base case analysis from £29,111/QALY to £90,482/QALY, indicating that baricitinib is less cost-effective when utilities are estimated from the HADS data collected in the baricitinib phase III studies compared to the EQ-5D data collected in the Adelphi DSP study.

Table 74. Cost-effectiveness results Scenario 5: Utilities based on HADS data from baricitinib phase III studies

	Total			Incremental			ICER versus baseline (£/QALY)
Technologies	**Costs (£)**	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
'Watch and wait'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	£90,482

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year; WTP, willingness to pay.

B.3.11.3.6 Scenario 6: Proportion of patients on BSC drugs based on clinical expert opinion

A summary of the clinical expert opinion estimates of the proportion of patients on BSC drugs is provided in Table 56. The cost-effectiveness outcomes of the scenario using clinical expert opinion to estimate the proportion of patients on BSC drugs are presented in Table 75 below. On average, a patient on baricitinib 4 mg accumulated [redacted] QALYs (discounted) compared to [redacted] QALYs on 'watch and wait', which amounts to an incremental difference of [redacted] QALYs. Total treatment costs were [redacted] for baricitinib 4mg and [redacted] in the 'watch and wait' arm of the model; an incremental difference of [redacted]. The ICER decreased from the base case analysis from £29,111/QALY to £10,378/QALY, indicating that baricitinib is more cost-effective when the proportion of patients on BSC is based on clinical expert opinion compared to the Adelphi DSP study.

Table 75. Cost-effectiveness results Scenario 6: Proportion of patients on BSC drugs based on clinical expert opinion

	Total			Incremental			ICER versus baseline (£/QALY)
Technologies	**Costs (£)**	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
'Watch and wait'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	£10,378

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year; WTP, willingness to pay.

B.3.11.4 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 probabilistic base case results showing that baricitinib is cost-effective versus 'watch and wait' and indicating it to be a cost-effective use of resources in the NHS. As demonstrated by the DSA (with PAS), the three most influential parameters driving the model the HSUV for $SALT_{75}$, the proportion of patients on DPCP treatment in the BSC state and the discontinuation rate for baricitinib 4mg in the Maintenance phase. Another driver of the model was the response rate at 36 weeks using $SALT_{50}$ for baricitinib 4mg and the resource frequency for the monitoring cost of DPCP treatment in the BSC state. Limited variation was observed in the majority of changes to the modelling approach that were explored in the scenario analyses: across the majority of scenarios conducted, baricitinib was associated with ICERs (with PAS) of less than £30,000 per QALY gained.

The utility scenario analyses based on the HRQoL data gathered from the baricitinib phase III studies are not relevant to patients with AA in the UK and should therefore be interpreted with caution. This is because the baricitinib Phase III studies had strict entry criteria that excluded patients with neuropsychiatric disorders who were likely experiencing the greatest HRQoL impairment due to AA. As a result, the HRQoL data collected from the BRAVE-AA trials do not adequately capture the HRQoL benefit associated with an improvement in SALT score following response to baricitinib treatment. Therefore, HRQoL data in the base case were taken from the

Adelphi DSP study that did not exclude patients suffering from most severe impacts of AA. Altogether, these results demonstrate the robustness of the model to uncertainty.

B.3.12 Subgroup analysis

Subgroups of patients with a SALT score of 50–94 (severe AA) or SALT 95–100 (very severe AA) were explored in scenario analyses. Results from these subgroup analyses are presented in Section B.3.11.3.1 and Section B.3.11.3.2, respectively.

B.3.13 Benefits not captured in the QALY calculation

Severe AA is associated with a significant HRQoL and psychological burden, which negatively impacts many aspects of a patient's life. Despite this, there are limited treatment options available, many of which are not supported by robust evidence and are associated with unpleasant or uncomfortable side effects.

Baricitinib is a clinically effective treatment for patients with severe AA and addresses the substantial unmet need for an evidence-based, effective and well-tolerated medication in this indication, by providing the patients who respond to treatment with a significant increase in hair growth and improved health-related quality of life (see Section B.2.6). However, as noted in Section B.2.6.3 and B.3.7, the EQ-5D instrument preferred by NICE may lack content validity in AA and therefore the QALY calculation may not fully capture the benefit of treatment response. Additionally, as noted in Section B.3.2.2, no mortality benefit was included in the model, despite literature evidence presented in Section B.1.3.2 that AA is associated with suicidal ideation and elevated mortality risk.

As the first evidence-based treatment specific to AA, baricitinib would allow patients to benefit from improved outcomes compared with current management, reducing the significant and negative burden of the disease on patients' lives. These improved outcomes for patients with severe AA also align with one of the focus areas outlined in the NHS long term plan, by contributing to reducing the demand on adult mental health services.[112] Baricitinib therefore is an innovative therapy that represents the first evidence-based treatment in this indication, thus representing an important step-change in the treatment of severe AA.

B.3.14 Validation

B.3.14.1 Validation of cost-effectiveness analysis

Clinical validity

Expert clinical input was sought during the development of the cost-effectiveness model to ensure that the inputs and assumptions used in the analysis were relevant to UK clinical practice. Feedback was obtained through individual teleconferences with a total of three clinical experts.

Internal model validity

The model programming was checked by an analyst who was not involved in the original development of the model using a validation checklist similar that reported in the published literature.[113] This involved a quality control check of the formulae used in the model and stress testing of the model to ensure that it behaves as expected when extreme values are used.[114]

B.3.15 Interpretation and conclusions of economic evidence

Summary of cost-effectiveness evidence

The cost-effectiveness of baricitinib in severe AA was evaluated versus 'watch and wait', the most clinically relevant comparator for this population, due to a lack of high-quality clinical data for another comparator. In the probabilistic base case, baricitinib was associated with higher costs (£[redacted] per patient over a lifetime horizon) and higher benefits ([redacted] QALYs per patient over a lifetime horizon) compared with 'watch and wait'. The base case probabilistic ICER was £29,111 per QALY gained and did not differ meaningfully from the deterministic ICER (£29,395 per QALY gained). In absolute terms, base case probabilistic results suggested that baricitinib 4mg was associated with a total cost of £[redacted], of which £[redacted] related to drug acquisition. These were partially compensated by cost savings due to reduced disease management costs, psychological burden costs and BSC drug monitoring costs. The PSA results indicated that baricitinib was [redacted]% likely to be cost-effective at a cost-effectiveness threshold of £30,000 per QALY gained.

Overall, the results indicate that baricitinib to be a cost-effective option for the treatment of severe AA within the NHS versus 'watch and wait'.

Strengths

The cost-effectiveness model developed for this submission has several strengths. The efficacy of baricitinib is based largely on robust phase III clinical trial data derived from two clinical trials. The model was built to align with the NICE reference case, adopting an NHS and PSS perspective, a lifetime time horizon to fully capture all costs and QALY gains associated with the interventions, and discount rates for costs and benefits of 3.5%. Utility data were elicited directly from AA patients in Europe, including the UK and were cross-walked from the EQ-5D-5L to the EQ-5D-3L, in line with the latest NICE guidance.

Limitations

There are some limitations to the current modelling approach that should be considered. Firstly, there was no previously published economic model for AA and therefore the current model structure was based on similar models in other dermatological disorders. The model structure and the key assumptions have been tested with clinical experts to ensure the model adequately captures disease progression and the main clinical and economics aspects of the disease, however, there remains an extent of uncertainty as to how well the model predicts the cost-effectiveness of baricitinib for the treatment of AA. Second, discontinuation rates for baricitinib and 'watch and wait' were based on data from the phase III studies with follow-up until Week 52 and Week 36 respectively. Given that these rates in the maintenance phase were used from Week 36 and beyond to estimate the probability of remaining in maintenance, there is inherent uncertainty on long-term discontinuation rates and the predictions they make. Third, the BSC health state assumes that patients would not benefit from treatment despite receiving a basket of various off-label therapies for the management of AA. This approach can underestimate the benefit that these patients exhibit from the BSC treatments, however, given the lack of comparative efficacy data for these treatments it was not possible to estimate their impact in efficacy and HRQoL terms. In addition to this, given the frequency of AEs and high relapse rates associated with such therapies, the extent of benefit of BSC for patients with AA is questionable.

Conclusion

Notwithstanding the limitations mentioned above, based on the currently available evidence, the results of the analyses conducted show, with a good degree of certainty, that baricitinib 4 mg is a clinically- and cost-effective alternative to 'watch and wait' in patients with severe AA.

B.4 References

1. Messenger AG, McKillop J, Farrant P, et al. British Association of Dermatologists' guidelines for the management of alopecia areata 2012. British Journal of Dermatology 2012;166:916-926.
2. Pratt CH, King LE, Jr., Messenger AG, et al. Alopecia areata. Nat Rev Dis Primers 2017;3:17011.
3. Palanivel JA, Macbeth AE, Chetty NC, et al. An insight into JAK-STAT signalling in dermatology. Clinical and Experimental Dermatology 2014;39:513-518.
4. Olamiju B, Friedmann A, King B. Treatment of severe alopecia areata with baricitinib. JAAD Case Rep 2019;5:892-894.
5. Eli Lilly. How does Olumiant work? 2022. Available at: https://www.olumiant.com/hcp/moa [Date accessed: February 2022].
6. 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: February 2022].
7. National Institute for Health and Care Excellence (NICE). Baricitinib for moderate to severe atopic dermatitis: Technology appraisal guidance [TA681]. Available at: https://www.nice.org.uk/guidance/ta681. [Date accessed: June 2020].
8. European Medicines Agency (EMA). Olumiant (baricitinib) SmPC. Available at: https://www.ema.europa.eu/en/documents/product-information/olumiant-epar-product-information_en.pdf. [Date accessed: August 2022].
9. British National Formulary. Baricitinib. Available at: https://bnf.nice.org.uk/medicinal-forms/baricitinib.html [Last accessed: February 2022].
10. Villasante Fricke AC, Miteva M. Epidemiology and burden of alopecia areata: a systematic review. Clin Cosmet Investig Dermatol 2015;8:397-403.
11. Montgomery K, White C, Thompson A. A mixed methods survey of social anxiety, anxiety, depression and wig use in alopecia. BMJ Open 2017;7:e015468.
12. Lintzeri DA, Constantinou A, Hillmann K, et al. Alopecia areata – Current understanding and management. JDDG: Journal der Deutschen Dermatologischen Gesellschaft 2022;20:59-90.
13. Macbeth AE, Holmes S, Harries M, et al. The associated burden of mental health conditions in alopecia areata: A population-based study in UK primary care. British Journal of Dermatology 2022;n/a.
14. Ikeda T. A new classification of alopecia areata. Dermatologica 1965;131:421-45.
15. Tosti A, Bellavista S, Iorizzo M. Alopecia areata: a long term follow-up study of 191 patients. J Am Acad Dermatol 2006;55:438-41.
16. Delamere FM, Sladden MM, Dobbins HM, et al. Interventions for alopecia areata. Cochrane Database Syst Rev 2008:Cd004413.
17. Alkhalifah A. Alopecia areata update. Dermatol Clin 2013;31:93-108.
18. Darwin E, Hirt PA, Fertig R, et al. Alopecia Areata: Review of Epidemiology, Clinical Features, Pathogenesis, and New Treatment Options. Int J Trichology 2018;10:51-60.
19. Triyangkulsri K, Suchonwanit P. Role of janus kinase inhibitors in the treatment of alopecia areata. Drug design, development and therapy 2018;12:2323-2335.
20. Xing L, Dai Z, Jabbari A, et al. Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. Nature medicine 2014;20:1043-1049.
21. Fukuyama M, Ito T, Ohyama M. Alopecia areata: Current understanding of the pathophysiology and update on therapeutic approaches, featuring the Japanese Dermatological Association guidelines. J Dermatol 2022;49:19-36.
22. National Institute for Health and Care Excellence (NICE). Alopecia areata: When should I suspect alopecia areata? Available at: https://cks.nice.org.uk/topics/alopecia-areata/diagnosis/diagnosis/. [Date accessed: February 2022].
23. Phillips TG, Slomiany WP, Allison R. Hair loss: Common causes and treatment. Am Fam Physician 2017;96:371-378.
24. Spano F, Donovan JC. Alopecia areata: Part 1: pathogenesis, diagnosis, and prognosis. Can Fam Physician 2015;61:751-5.
25. Food and Drug Administration (FDA). The voice of the patient: Alopecia areata. 2018. Available at: https://www.fda.gov/media/112100/download. [Date accessed: March 2022].

26. Olsen EA, Hordinsky MK, Price VH, et al. Alopecia areata investigational assessment guidelines--Part II. National Alopecia Areata Foundation. J Am Acad Dermatol 2004;51:440-7.
27. Lee S, Lee H, Lee CH, et al. Comorbidities in alopecia areata: A systematic review and meta-analysis. J Am Acad Dermatol 2019a;80:466-477.e16.
28. Olsen EA. Investigative guidelines for alopecia areata. Dermatologic Therapy 2011;24:311-319.
29. National Institute for Health and Care Excellence (NICE). Alopecia areata: Management of alopecia areata. Available at: https://cks.nice.org.uk/topics/alopecia-areata/management/management/. [Date accessed: February 2022].
30. Eli Lilly. Clinical Study Protocol: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA1). June 2021.
31. Eli Lilly. Clinical Study Protocol: AA Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata(BRAVE-AA2). June 2021.
32. Lyakhovitsky A, Aronovich A, Gilboa S, et al. Alopecia areata: a long-term follow-up study of 104 patients. Journal of the European Academy of Dermatology and Venereology 2019;33:1602-1609.
33. Meah N, Wall D, York K, et al. The Alopecia Areata Consensus of Experts (ACE) study part II: Results of an international expert opinion on diagnosis and laboratory evaluation for alopecia areata. J Am Acad Dermatol 2021;84:1594-1601.
34. Burroway B, Griggs J, Tosti A. Alopecia totalis and universalis long-term outcomes: a review. Journal of the European Academy of Dermatology and Venereology 2020;34:709-715.
35. National Institute for Health and Care Excellence (NICE). Alopecia areata: Summary. Available at: https://cks.nice.org.uk/topics/alopecia-areata/. [Date accessed: February 2022].
36. Harries M, Macbeth AE, Holmes S, et al. The epidemiology of alopecia areata: a population-based cohort study in UK primary care*. British Journal of Dermatology 2022;186:257-265.
37. You HR, Kim SJ. Factors associated with severity of alopecia areata. Ann Dermatol 2017;29:565-570.
38. FDA. The Voice of the Patient: Alopecia Areata. Public Meeting: September 11, 2017. Report Date: March 2018. Available at: https://www.fda.gov/files/about%20fda/published/Alopecia-Areata--The-Voice-of-the-Patient.pdf [Last accessed: February 2022].
39. Titeca G, Goudetsidis L, Francq B, et al. 'The psychosocial burden of alopecia areata and androgenetica': a cross-sectional multicentre study among dermatological out-patients in 13 European countries. Journal of the European Academy of Dermatology and Venereology 2020;34:406-411.
40. Aldhouse NVJ, Kitchen H, Knight S, et al. "'You lose your hair, what's the big deal?' I was so embarrassed, I was so self-conscious, I was so depressed:" a qualitative interview study to understand the psychosocial burden of alopecia areata. J Patient Rep Outcomes 2020;4:76.
41. Murphrey MB, Agarwal S, Zito PM. Anatomy, Hair. StatPearls. Treasure Island (FL), 2022.
42. Davey L, Clarke V, Jenkinson E. Living with alopecia areata: an online qualitative survey study. Br J Dermatol 2019;180:1377-1389.
43. Creadore A, Manjaly P, Li SJ, et al. Evaluation of Stigma Toward Individuals With Alopecia. JAMA Dermatol 2021;157:392-398.
44. Harries MJ, Sun J, Paus R, et al. Management of alopecia areata. BMJ 2010;341:c3671.
45. Delamere FM, Sladden MJ, Dobbins HM, et al. Interventions for alopecia areata. Cochrane Database of Systematic Reviews 2008.
46. Shi Q, Duvic M, Osei JS, et al. Health-Related Quality of Life (HRQoL) in alopecia areata patients-a secondary analysis of the National Alopecia Areata Registry Data. J Investig Dermatol Symp Proc 2013;16:S49-50.
47. Rencz F, Gulácsi L, Péntek M, et al. Alopecia areata and health-related quality of life: a systematic review and meta-analysis. Br J Dermatol 2016;175:561-71.
48. Liu LY, King BA, Craiglow BG. Health-related quality of life (HRQoL) among patients with alopecia areata (AA): A systematic review. J Am Acad Dermatol 2016;75:806-812.e3.
49. Abedini R, Hallaji Z, Lajevardi V, et al. Quality of life in mild and severe alopecia areata patients. Int J Womens Dermatol 2018;4:91-94.
50. Balieva F, Kupfer J, Lien L, et al. The burden of common skin diseases assessed with the EQ5D™: a European multicentre study in 13 countries. Br J Dermatol 2017;176:1170-1178.

51. Dalgard FJ, Bewley A, Evers AW, et al. Stigmatisation and body image impairment in dermatological patients: protocol for an observational multicentre study in 16 European countries. BMJ Open 2018;8:e024877.
52. Lee S, Lee YB, Kim BJ, et al. All-Cause and Cause-Specific Mortality Risks Associated With Alopecia Areata: A Korean Nationwide Population-Based Study. JAMA Dermatol 2019b;155:922-928.
53. Toussi A, Barton VR, Le ST, et al. Psychosocial and psychiatric comorbidities and health-related quality of life in alopecia areata: A systematic review. J Am Acad Dermatol 2021;85:162-175.
54. Mesinkovska N, King B, Mirmirani P, et al. Burden of Illness in Alopecia Areata: A Cross-Sectional Online Survey Study. J Investig Dermatol Symp Proc 2020;20:S62-s68.
55. European Medicines Agency (EMA). Olumiant (baricitinib). Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/olumiant. [Date accessed: August 2022].
56. Food and Drug Administration (FDA). FDA Approves First Systemic Treatment for Alopecia Areata. Available at: https://www.fda.gov/news-events/press-announcements/fda-approves-first-systemic-treatment-alopecia-areata. [Date accessed: August 2022].
57. Alopecia UK. A Guide to Wigs. 2018. Available at: https://www.alopecia.org.uk/wigs. [Date accessed: February 2022].
58. Gholizadeh N, Sadrzadeh-Afshar M-S, Sheykhbahaei N. Intralesional corticosteroid injection as an effective treatment method for oral lesions: a meta-analysis. Brazilian Journal of Pharmaceutical Sciences 2020;56.
59. Olsen EA, Canfield D. SALT II: A new take on the Severity of Alopecia Tool (SALT) for determining percentage scalp hair loss. J Am Acad Dermatol 2016;75:1268-1270.
60. Kar BR, Handa S, Dogra S, et al. Placebo-controlled oral pulse prednisolone therapy in alopecia areata. J Am Acad Dermatol 2005;52:287-90.
61. Yasir M, Goyal A, Sonthalia S. Corticosteroid Adverse Effects. StatPearls. Treasure Island (FL): StatPearls Publishing

62. Rokhsar CK, Shupack JL, Vafai JJ, et al. Efficacy of topical sensitizers in the treatment of alopecia areata. J Am Acad Dermatol 1998;39:751-61.
63. Lee S, Kim BJ, Lee YB, et al. Hair Regrowth Outcomes of Contact Immunotherapy for Patients With Alopecia Areata: A Systematic Review and Meta-analysis. JAMA Dermatology 2018;154:1145-1151.
64. King B, Ko J, Forman S, et al. Efficacy and safety of the oral Janus kinase inhibitor baricitinib in the treatment of adults with alopecia areata: Phase 2 results from a randomized controlled study. J Am Acad Dermatol 2021;85:847-853.
65. Eli Lilly. Clinical Study Report: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA1). 2022.
66. Eli Lilly. Clinical Study Report: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA2). 2022.
67. King B, Ohyama M, Kwon O, et al. Two Phase 3 Trials of Baricitinib for Alopecia Areata. N Engl J Med 2022.
68. EMA. Olumiant (baricitinib) Assessment Report. Available at: https://www.ema.europa.eu/en/documents/variation-report/olumiant-h-c-4085-ii-0029-g-epar-assessment-report-variation_en.pdf. Last accessed: August 2022.
69. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
70. Centre for Reviews and Dissemination. Systematic Reviews: CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination, University of York, 2008.
71. Wyrwich KW, Kitchen H, Knight S, et al. Development of the Scalp Hair Assessment PRO™ measure for alopecia areata. British Journal of Dermatology 2020;183:1065-1072.
72. National Institute for Health and Care Excellence (NICE). Alopecia areata. Available at: https://cks.nice.org.uk/topics/alopecia-areata/. [Date accessed: February 2022].
73. Titeca G, Goudetsidis L, Francq B, Sampogna F, Gieler U, Tomas-Aragones L et al. The psychosocial burden of alopecia areata and androgenetica": A cross-sectional multicenter

study among dermatological out-patients in 13 European Countries. *J Eur Acad Dermatol Venereol.* 2019; 34(2): 406-411. .
74. EuroQoL. 2018. Available from: https://euroqol.org/. [Date accessed: Februrary 2022].
75. Batchelor JM, Thomas KS, Akram P, et al. Home-based narrowband UVB, topical corticosteroid or combination for children and adults with vitiligo: HI-Light Vitiligo three-arm RCT. 2020;24:64.
76. Lai VWY, Chen G, Gin D, et al. Cyclosporine for moderate-to-severe alopecia areata: A double-blind, randomized, placebo-controlled clinical trial of efficacy and safety. Journal of the American Academy of Dermatology 2019;81:694-701.
77. Asilian A, Fatemi F, Ganjei Z, et al. Oral pulse betamethasone, methotrexate, and combination therapy to treat severe alopecia areata: A randomized, double-blind, placebo-controlled, clinical trial. Iranian Journal of Pharmaceutical Research 2021;20:267-273.
78. Ghandi N, Daneshmand R, Hatami P, et al. A randomized trial of diphenylcyclopropenone (DPCP) combined with anthralin versus DPCP alone for treating moderate to severe alopecia areata. Int Immunopharmacol 2021;99:107971.
79. Kurosawa M, Nakagawa S, Mizuashi M, et al. A comparison of the efficacy, relapse rate and side effects among three modalities of systemic corticosteroid therapy for alopecia areata. Dermatology 2006;212:361-365.
80. Thuangtong R, Varothai S, Triwongwaranat D, et al. Multi-concentration level patch test guided diphenyl cyclopropenone (DPCP) treatment in alopecia totalis or alopecia universalis. Journal of the Medical Association of Thailand 2017;100:86-92.
81. Shapiro J, Tan J, Ho V, et al. Treatment of chronic severe alopecia areata with topical diphenylcyclopropenone and 5% minoxidil: a clinical and immunopathologic evaluation. J Am Acad Dermatol 1993;29:729-35.
82. Tiwary AK, Mishra DK, Chaudhary SS. Comparative study of efficacy and safety of topical squaric acid dibutylester and diphenylcyclopropenone for the treatment of alopecia areata. North American Journal of Medical Sciences 2016;8:237-242.
83. Al Bazzal A, Hatami P, Abedini R, et al. A prospective comparative study of two regimens of diphenylcyclopropenone (DPCP) in the treatment of alopecia areata. Int Immunopharmacol 2021;101:108186.
84. Rocha VB, Kakizaki P, Donati A, et al. Randomized controlled study comparing the use of diphencyprone and anthralin in the treatment of extensive chronic alopecia areata. An Bras Dermatol 2021;96:372-376.
85. Alsufyani HS, Rawas WA, Alsufyani SS, et al. The effect of methotrexate in the treatment of alopecia areata. The Egyptian Journal of Hospital Medicine 2017;67:599-604.
86. English J, Heinisch S. Methotrexate treatment for alopecia areata with greater than 50% involvement. Hair Ther Transplant 2015;5:2.
87. Ferrando J, Grimalt R. Partial response of severe alopecia areata to cyclosporine A. Dermatology 1999;199:67-9.
88. Firooz A, Fouladi D. Methotrexate plus prednisolone in severe alopecia areata. American Journal of Drug Discovery and Development 2013;3:188-193.
89. Maryam A, Hassan S, Farshad F, et al. The efficacy of topical diphencyprone in the treatment of alopecia areata. Indian J Dermatol 2009;54:88-9.
90. Joly P. The use of methotrexate alone or in combination with low doses of oral corticosteroids in the treatment of alopecia totalis or universalis. J Am Acad Dermatol 2006;55:632-6.
91. Shapiro J, Lui H, Tron V, et al. Systemic cyclosporine and low-dose prednisone in the treatment of chronic severe alopecia areata: a clinical and immunopathologic evaluation. Journal of the American Academy of Dermatology 1997;36:114-117.
92. Gupta AK, Ellis CN, Cooper KD, et al. Oral cyclosporine for the treatment of alopecia areata: a clinical and immunohistochemical analysis. Journal of the American Academy of Dermatology 1990;22:242-250.
93. Jang YH, Kim SL, Lee KC, et al. A comparative study of oral cyclosporine and betamethasone minipulse therapy in the treatment of alopecia areata. Annals of Dermatology 2016;28:569-574.
94. Shin J-W, Huh C-H, Kim M-W, et al. Comparison of the treatment outcome of oral tofacitinib with other conventional therapies in refractory alopecia totalis and universalis: A retrospective study. Acta Dermato-Venereologica 2018;99:41-46.

95. Sriphojanart T, Khunkhet S, Suchonwanit P. A retrospective comparative study of the efficacy and safety of two regimens of diphenylcyclopropenone in the treatment of recalcitrant alopecia areata. Dermatology Reports 2017;9:55-58.
96. Vañó-Galván S, Hermosa-Gelbard Á, Sánchez-Neila N, et al. Pulse corticosteroid therapy with oral dexamethasone for the treatment of adult alopecia totalis and universalis. J Am Acad Dermatol 2016;74:1005-7.
97. Vañó-Galván S, Hermosa-Gelbard Á, Sánchez-Neila N, et al. Treatment of recalcitrant adult alopecia areata universalis with oral azathioprine. J Am Acad Dermatol 2016;74:1007-8.
98. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. January 2022. Available at: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741. Last accessed: April 2022.
99. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results From an English Population Study [EEPRU Report]. 2020. University of Sheffield & University of York.
100. Office for National Statistics. Population Estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2020. June 2021. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2020. [Last accessed: May 2020]. .
101. National Health Service (NHS). 2019/20 National Cost Collection data. National Cost Collection for the NHS. 2020. Available at: https://www.england.nhs.uk/national-cost-collection/#ncc1819. [Date accessed: May 2022].
102. National Health Service (NHS) Business Services Authority. Drug Tariff (April 2022). 2022. Available at: https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/drug-tariff. [Date accessed: May 2022].
103. National Health Service (NHS). Wigs and fabric supports on the NHS. 2020. Available at: https://www.nhs.uk/nhs-services/help-with-health-costs/wigs-and-fabric-supports-on-the-nhs/. [Lats accessed: May 2022].
104. PSSRU: Unit Costs of Health and Social Care (2021). Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-of-health-and-social-care-2021/. [Last accessed: May 2022].
105. National Institute for Health and Care Excellence (NICE). Depression in adults: recognition and management [CG90]. 2022. Available at: https://www.nice.org.uk/guidance/cg90. [Last accessed: May 2022].
106. National Institute for Health and Care Excellence (NICE). Dupilumab for treating moderate to severe atopic dermatitis [TA534]. Available at: https://www.nice.org.uk/guidance/ta534. [Date accessed: May 2022].
107. Hernández Alava M, Pudney S, Wailoo A. Methods for mapping between the EQ-5D-5L and the 3L for technology appraisal: Report by the Decision Support Unit. 2017. Decision Support Unit, University of Sheffield.
108. Brazier J, Connell J, Papaioannou D, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess 2014;18.
109. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by age and sex for the UK: Report by the Decision Support Unit. 2022. Decision Support Unit, University of Sheffield. .
110. Thermo Fisher Scientific. Treatment fees for DPCP hospital treatment (Diphenylcyclopropenone, 98%, Thermo Scientific™). 2022. Available at: https://www.fishersci.co.uk/shop/products/diphenylcyclopropenone-98-thermo-scientific/10317740. Last accessed: June 2022.
111. National Health Service (NHS). 2019/20 National Cost Collection data. National Cost Collection for the NHS. Volume February, 2022]. NHS: NHS, 2020.
112. National Health Service (NHS). NHS Long Term Plan: Adult Mental Health Services. Available at: https://www.longtermplan.nhs.uk/online-version/chapter-3-further-progress-on-care-quality-and-outcomes/better-care-for-major-health-conditions/adult-mental-health-services/. [Date accessed: July 2022].
113. Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, et al. TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. Pharmacoeconomics 2019;37:1391-1408.

114. Eli Lilly. NICE Cost-Effectiveness Model for Baricitinib in AA: Model Sanity Checklist. August 2022.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Clarification questions

September 2022

File name	Version	Contains confidential information	Date
ID3979_Baricitinib in AA_Clarification Questions [ACIC REDACTED]_16September2022	FINAL	Yes	16th September 2022

Section A: Clarification on effectiveness data

Study design and reporting

A1. Priority question. Please clarify the most recent data cut available and which data cuts have been used in the company submission (CS).

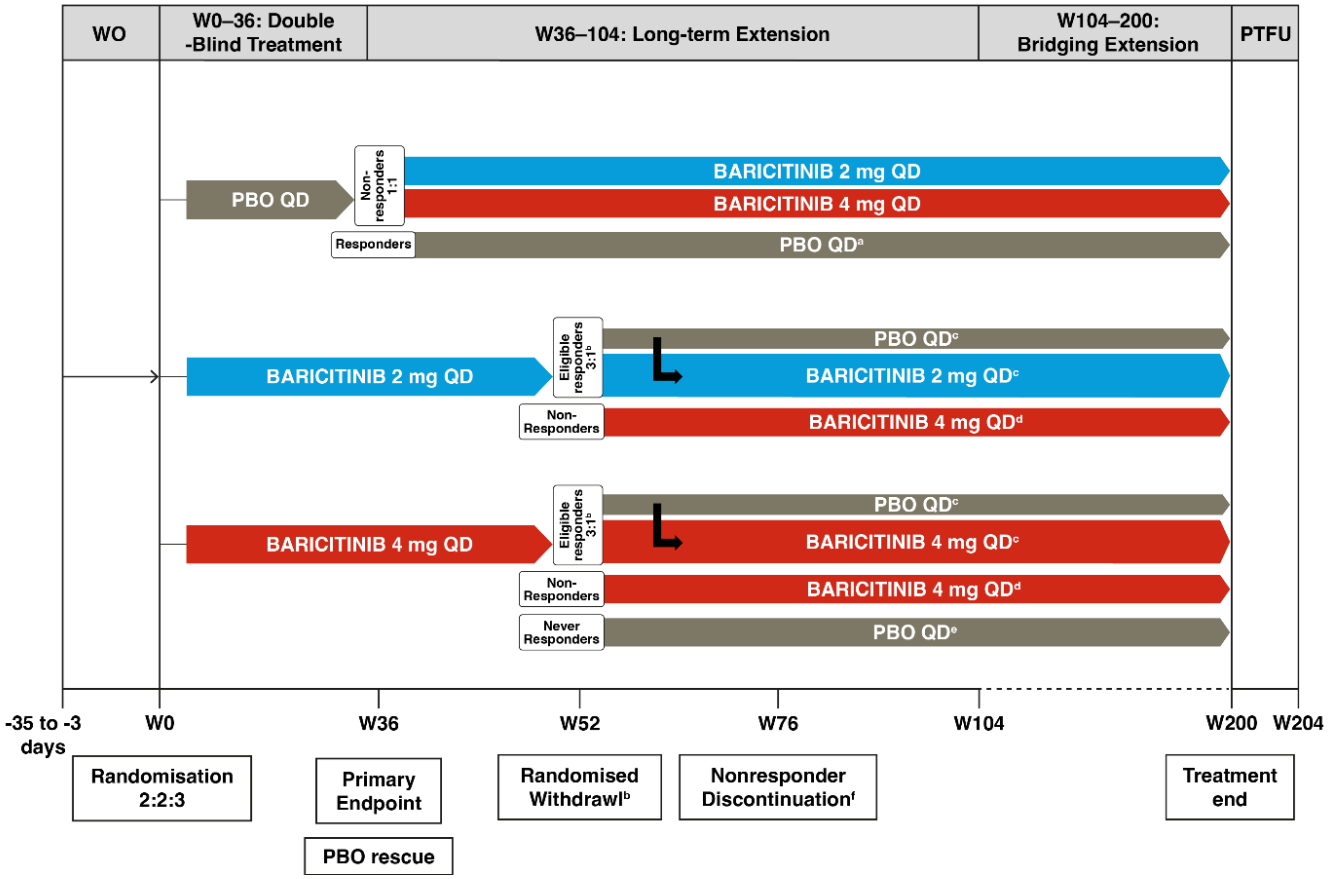
a) **Please provide any clinical study reports (CSRs) that are available for data cuts beyond Week 36.**

Clinical study report (CSR) addendums are available for Week 76, and these have been uploaded alongside responses to these clarification questions.

A2. Priority question. CS, Document B, section B.2.3.1, Figures 3 and 4; Eli Lilly 2022 BRAVE-AA1 CSR; Eli Lilly 2021 BRAVE-AA1 Protocol and Eli Lilly 2021 BRAVE-AA2 Protocol. Please clarify whether the information in Figure 3 (Study design of BRAVE-AA1) and Figure 4 (Study design of BRAVE-AA2) is correct. For BRAVE-AA1, Figure 3 shows that people whose condition had responded to treatment at 52 weeks in the 4mg baricitinib arm were eligible to be re-randomised (3:1) to 2mg baricitinib or placebo. In the protocol and CSR figures (for example, Figure JAHO.2), they are randomised to [redacted] or placebo. For BRAVE-AA2, Figure 4 shows that people whose condition had responded to treatment at 52 weeks in the 4mg baricitinib arm were eligible to be re-randomised (3:1) to 2mg or 4mg baricitinib. However, in the protocol (Figure JAIR.1), this ratio is [redacted]. If the information in Figures 3 and 4 are incorrect, please provide updated figures.

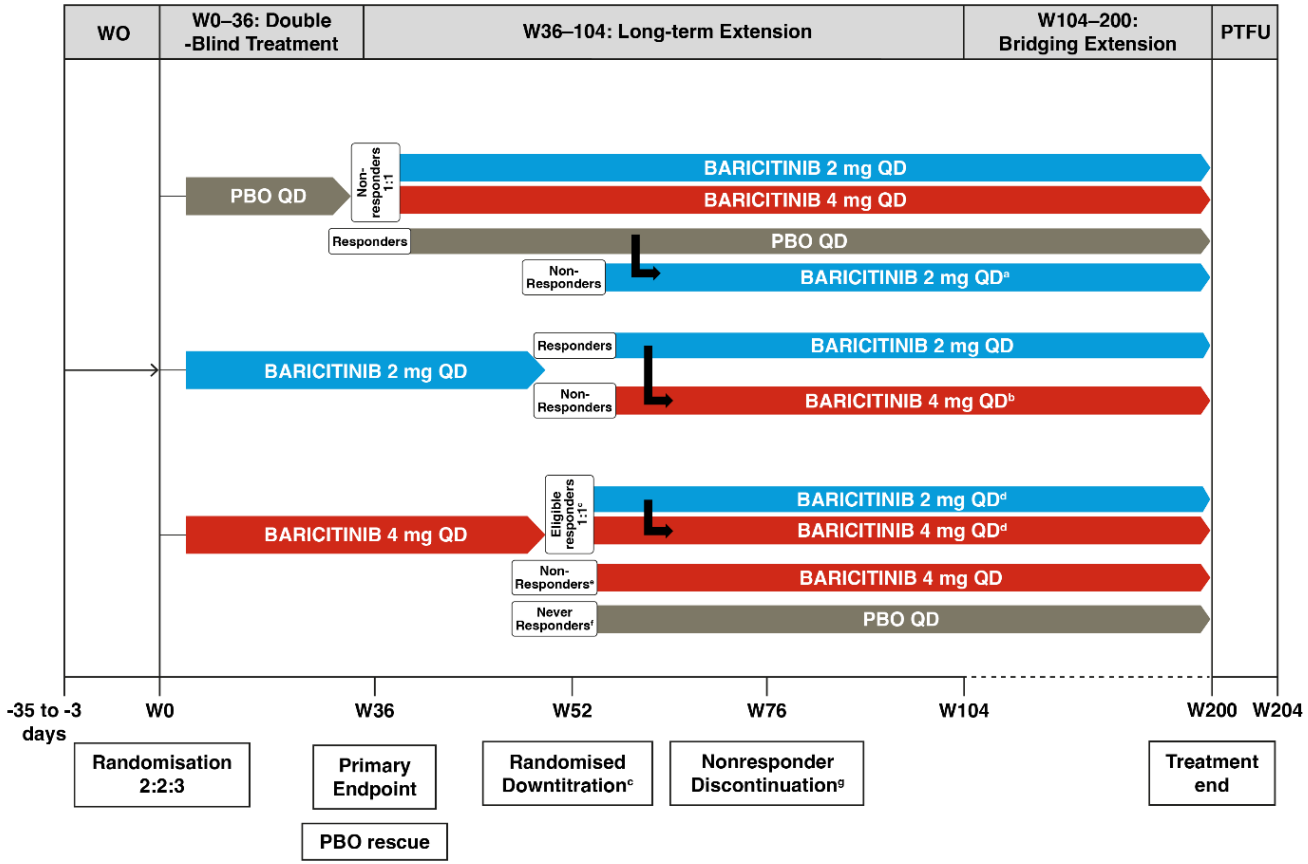
Figure 3 and Figure 4 in Document B of the company submission (CS) include typographical errors. These errors have now been corrected to align with the protocols for BRAVE-AA1 and BRAVE-AA2, and are presented in Figure 1 and Figure 2, respectively.

Figure 1. Study design of BRAVE-AA1



Footnotes: [a] Placebo responders stayed on placebo for remainder of the trial, even if relapse was observed later. [b] Patients with SALT ≤20 who stayed on the same dose of baricitinib from week 0 were randomised to stay on current baricitinib dose, or transitioned to placebo.[c] Responders participating in randomised withdrawal who experienced >20-point absolute worsening in total SALT score after week 52 were retreated with baricitinib dose to which they were originally randomised if they were randomised to placebo at week 52, OR continued to receive same dose of baricitinib if they were randomised to remain on baricitinib at week 52. [d] Non-responders at week 52 were rescued to baricitinib 4 mg if receiving baricitinib 2 mg from baseline, OR remained on baricitinib 4 mg if they were in the 4-mg group and achieved SALT ≤20 before week 52. [e] Never responders (never achieved SALT ≤20 by week 52 despite being in the baricitinib 4-mg group from baseline and had not experienced a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss) were automatically transitioned to placebo.[f] Non-responders at week 52 AND week 76 were automatically discontinued at week 76 unless they had a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss.
Abbreviations: PBO: placebo; PTFU: post-trial follow up; QD: once daily; W: week.
Source: BRAVE-AA1 Clinical Study Report.[1]

Figure 2. Study design of BRAVE-AA2



Footnotes: [a] Placebo-treated patients not eligible for rescue to baricitinib at week 36 (due to spontaneous remission) were rescued to baricitinib if they were non-responders at week 52, OR if they experienced loss of treatment benefit after week 52. [b] Patients randomised to baricitinib 2 mg at week 0 were rescued to the 4-mg dose if they were non-responders at week 52, OR were responders at week 52 but experienced a >20-point worsening in SALT score after week 52. [c] Responders in the baricitinib 4-mg group (SALT ≤20 who stayed on 4 mg from week 0) were randomised to either stay on 4 mg OR transition to 2 mg.[d] Responders participating in the randomised down-titration who experienced a loss of treatment benefit after week 52 were re-treated with baricitinib 4 mg if they were randomised to the 2-mg dose at week 52, OR continued to receive baricitinib 4 mg if they randomised to remain on the 4-mg dose at week 52. [e] At week 52, non-responders (SALT >20) in the baricitinib 4-mg group since baseline who achieved SALT ≤20 before week 52 remained on 4 mg. [f] Never responders (never achieved SALT ≤20 by week 52 despite being in the baricitinib 4-mg group from baseline and had not experienced a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss) were automatically transitioned to placebo. [g] Non-responders at week 52 AND week 76 were automatically discontinued at week 76 unless they had a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss.

Abbreviations: PBO: placebo; PTFU: post-trial follow up; QD: once daily; W: week.
Source: BRAVE-AA2 Clinical Study Report.[2]

Population

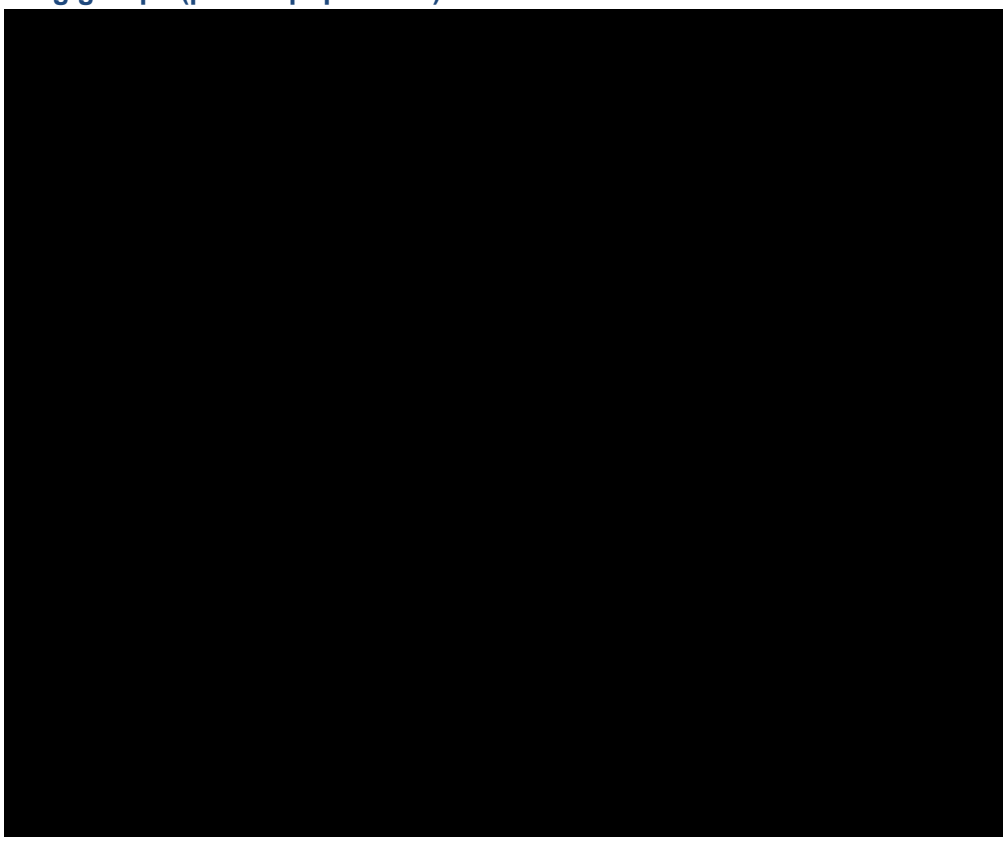
A3. Priority question. CS, Document B, section B.2.3.3, Tables 9 and 10. Please provide the baseline distribution of Severity of Alopecia Tool (SALT) scores in BRAVE-AA1 and BRAVE-AA2 for both the 4mg baricitinib and placebo arms using either, i) bins of width 5 (50-55, 55-60… 95-100), or ii) a continuous density distribution plot.

The baseline distributions of SALT scores for the pooled population from BRAVE-AA1 and BRAVE-AA2 are presented in Figure 3 and Figure 4 for the placebo arms and baricitinib 4mg arms, respectively, using bins of SALT5 overlayed with a continuous density distribution plot.

Figure 3. Baseline distribution of SALT scores in BRAVE-AA1 and BRAVE-AA2 placebo groups (pooled population)

Footnotes: One patient in the placebo arm of BRAVE-AA2 had a baseline score of SALT<50 (SALT32) due to a protocol violation. The x-axis is placed on the start of the bins of width.
Abbreviations: SALT: severity of alopecia tool.

Figure 4. Baseline distribution of SALT scores in BRAVE-AA1 and BRAVE-AA2 baricitinib 4 mg groups (pooled population)



Footnotes: The x-axis is placed on the start of the bins of width.
Abbreviations: SALT: severity of alopecia tool.

A4. Priority question. CS, Document B, section B.2.3.3, Tables 9 and 10. Please provide the baseline distribution of duration of current alopecia areata (AA) episode in BRAVE-AA1 and BRAVE-AA2 for both the 4 mg baricitinib and placebo arms using either, i) bins of width 6 months, or ii) a continuous density distribution plot.

The baseline distribution of current AA episode is presented in Figure 5 and Figure 6 for the placebo and baricitinib 4 mg arms of the BRAVE-AA1 trial, and in Figure 7 and Figure 8 placebo and baricitinib 4 mg arms of the BRAVE-AA2 trial. Figure 9 and Figure 10 present the baseline distribution of current AA episode in the pooled population for the placebo and baricitinib 4 mg arms, respectively. It should be noted that during the BRAVE-AA trials, the duration of current AA episode was measured in years, meaning that 6-month bins cannot be presented. As such, the graphs below present bins of 1 year, overlayed with a continuous density distribution plot.

Figure 5. Baseline distribution of duration of current AA episode in BRAVE-AA1 in the placebo arm

Footnotes: The x-axis is placed on the start of the bins of width
Abbreviations: AA: alopecia areata.

Figure 6. Baseline distribution of duration of current AA episode in BRAVE-AA1 in the baricitinib 4 mg arm

Footnotes: The x-axis is placed on the start of the bins of width
Abbreviations: AA: alopecia areata.

Figure 7. Baseline distribution of duration of current AA episode in BRAVE-AA2 in the placebo arm

Footnotes: The x-axis is placed on the start of the bins of width
Abbreviations: AA: alopecia areata.

Figure 8. Baseline distribution of duration of current AA episode in BRAVE-AA2 in the baricitinib 4 mg arm

Footnotes: The x-axis is placed on the start of the bins of width
Abbreviations: AA: alopecia areata.

Figure 9. Baseline distribution of duration of current AA episode in the pooled population from BRAVE-AA1 and BRAVE-AA2 in the placebo arm

Footnotes: The x-axis is placed on the start of the bins of width
Abbreviations: AA: alopecia areata.

Figure 10 Baseline distribution of duration of current AA episode in the pooled population from BRAVE-AA1 and BRAVE-AA2 in the baricitinib 4 mg arm

Footnotes: The x-axis is placed on the start of the bins of width
Abbreviations: AA: alopecia areata.

A5. Priority question. CS, Document B, section B.2.3.3, Tables 9 and 10. Please provide the median, interquartile range and a density distribution plot for the baseline age of participants in BRAVE-AA1 and BRAVE-AA2 for the 4mg baricitinib and placebo arms.

Table 1 provides a summary of the median and interquartile range for baseline patient age in the placebo and baricitinib 4 mg arms within the pooled population of the BRAVE-AA trials. Density distribution plots for the baseline age of participants in the pooled population of the BRAVE-AA1 and BRAVE-AA2 trials are presented in Figure 11 and Figure 12 for the placebo and baricitinib 4 mg arms, respectively.

Table 1. Summary of baseline age in BRAVE-AA1 and BRAVE-AA2 (pooled population)

	PBO (N=345)	Baricitinib 4 mg (N=515)
Median age	[redacted]	[redacted]
IQR	[redacted]	[redacted]

Abbreviations: IQR: interquartile range; PBO: placebo.

Figure 11. Baseline distribution of participant age in the placebo arm of BRAVE-AA1 and BRAVE-AA2 (pooled population)

Figure 12. Baseline distribution of participant age in the baricitinib 4 mg arm of BRAVE-AA1 and BRAVE-AA2 (pooled population)

A6. Priority question. Please provide scatter plots of baseline SALT score and duration of current AA episode for BRAVE-AA1, BRAVE-AA2 and the pooled population. Please provide the correlation coefficients.

A scatter plot of baseline SALT score and duration of current AA episode in the placebo and baricitinib 4 mg arms of the pooled population of BRAVE-AA1 and BRAVE-AA2 is presented in Figure 13 and Figure 14, respectively. The Pearson correlation coefficient is 0.087 in the placebo arm, and 0.039 in the baricitinib 4 mg arm, indicating no obvious relationship between baseline SALT score and the duration of current AA episode in either treatment arm.

Figure 13. Correlation between baseline SALT score and duration of current AA episode in the placebo arm of BRAVE-AA1 and BRAVE-AA2 (pooled population)

Footnotes: Pearson r=0.087. One patient in the placebo arm of BRAVE-AA2 had a baseline score of SALT<50 (SALT32) due to a protocol violation
Abbreviations: AA: alopecia areata; SALT: severity of alopecia tool.

Figure 14. Correlation between baseline SALT score and duration of current AA episode in the baricitinib 4 mg arm of BRAVE-AA1 and BRAVE-AA2 (pooled population)

Footnotes: Pearson r=0.039.
Abbreviations: AA: alopecia areata; SALT: severity of alopecia tool.

A7. Priority question. CS, Document B, section B.2.3.1, Table 6. Please comment on how the exclusion of patients with a current episode of severe or very severe AA of 8 years or more might affect the generalisability of the BRAVE-AA1 and BRAVE-AA2 trial populations to NHS clinical practice. If possible, please provide the number of patients at screening who had a current episode of severe or very severe alopecia for 8 years or more, and therefore, did not meet the inclusion criteria #3c in BRAVE-AA1 and BRAVE-AA2.

Based on clinical input, the Company understands that the proportion of patients in UK clinical practice with a current episode of severe or very severe AA lasting > 8 years who will seek or be offered treatment with baricitinib is relatively low. Data on the number of patients who failed screening due to having a current episode of severe or very severe AA of more than 8 years (criteria #3c) was not collected. However, across both BRAVE-AA1 and BRAVE-AA2, the proportion of patients who failed screening due to criteria #3 (inclusive of #3a, #3b, #3c) was approximately 5%. Furthermore, in line with the note attached to criteria #3c in the protocol, some patients with severe or very severe AA for ≥ 8 years may have been enrolled to the trial. The note specifies that patients with severe or very severe AA for ≥8 years who have demonstrated episodes of regrowth, spontaneous or under treatment, on the affected areas of the scalp over the past 8 years, may be enrolled.[3, 4] Therefore, it is unlikely that this inclusion criterion in BRAVE-AA1 and BRAVE-AA2 would impact the generalisability of the trial results to the population relevant to UK clinical practice.

A8. Priority question. CS, Document B, section B.2.6.3, p76; Eli Lilly 2022 BRAVE-AA2 CSR and Eli Lilly 2022 BRAVE-AA1 CSR. Table JAHO.8.1. (BRAVE-AA1 CSR) and Table JAIR.8.1. (BRAVE-AA2 CSR) suggest that [redacted] due to the presence of significant uncontrolled neuropsychiatric disorder (criteria #24). The company submission states that the exclusion of such patients may contribute to a ceiling effect in the trial EQ-5D data. Please confirm whether [redacted] because of the presence of significant uncontrolled neuropsychiatric disorder.

The Company confirms that zero patients failed screening due to the presence of significant uncontrolled neuropsychiatric disorder (criteria #24). However, it is possible that patients with significant uncontrolled neuropsychiatric disorder were not selected by trial investigators to enter the screening period for the trial. Consequently, these patients would not be captured in the clinical study reports. Furthermore, this does not negate the observed ceiling effect in the EQ-5D data generated, as discussed further in Question B7.

Treatment effectiveness

A9. Priority question. The EAG's clinical experts believe that some best supportive care (BSC) therapies, such as 2,3-diphenylcyclopropenone (DPCP), could be effective in a proportion of people with severe AA. This is supported by the DPCP studies included in the company's feasibility assessment, as well as a systematic review and meta-analysis in which complete hair regrowth was reported for 22% of alopecia totalis/universalis patients (Lee et al. 2018, doi:10.1001/jamadermatol.2018.2312). In the absence of any indirect treatment comparisons, please comment on the extent to which the placebo response might be smaller than response to BSC therapies.

While it cannot be excluded that BSC therapies could be more effective than a watch and wait approach for patients with severe AA, there is currently no robust evidence to support the efficacy of these treatments versus placebo in patients with severe AA, meaning that the Company does not consider this possibility to be of importance to the current appraisal nor to impact the cost-effectiveness outcomes of baricitinib in this indication.

To date, there have been no randomised control trials (RCTs) which have investigated the efficacy of DPCP versus placebo. In addition, as highlighted in the feasibility assessment presented in Document B of the CS Section B.2.9.2, most trials investigating BSC treatments include small and heterogenous patient populations. Studies of BSC treatments also frequently report relapse during/following treatment; in a study by Lamb *et al.* (2015) investigating DPCP, included in the meta-analysis by Lee et al. (2018) cited above, 13.5% of patients achieved initial hair regrowth but then lost hair during treatment.[5] As such, the quality and robustness of the published efficacy evidence for BSC treatments is generally low and lacks generalisability relative to the BRAVE-AA trials that form the evidence base of the current appraisal.

In addition to this, some studies do not account for non-completers and/or discontinuations in their efficacy results, which can produce artificially higher response rates in favour of the intervention. For instance, in a study by Durdu *et al.* (2015), only 25 patients (53.2%) on DPCP and anthralin were considered in the efficacy analysis, as the analysis did not consider the four patients (8.5%) who discontinued treatment or 18 patients (38.3%) who received treatment for less than 30 weeks.[6] Similarly, in the aforementioned study by Lamb *et al.* (2015), 43 courses of DPCP were not completed (21%) by enrolled patients, and these patients were subsequently excluded from the efficacy analysis.[5] Considering this, the 22% value reported in the above question for complete hair regrowth in alopecia totalis/universalis patients following DPCP treatment is likely to be an overestimation and should therefore be interpreted with caution. It should also be noted that while DPCP may be the most effective BSC treatment, its use is limited to selected specialist centres in the UK, which limits in role in AA management.[7]

For the reasons outlined above, the Company were not able to define a robust estimate of the efficacy of DPCP, or indeed any of the BSC treatments, for use in the model, and the CEM therefore instead relies on the efficacy estimates for placebo and baricitinib generated in the BRAVE-AA trials, which, unlike studies for BSC, provide robust and reliable evidence for the efficacy of these treatments in the population of relevance to this appraisal.

A10. Eli Lilly 2022 BRAVE-AA1 CSR; Eli Lilly 2022 BRAVE-AA2 CSR; Eli Lilly 2021 BRAVE-AA1 Protocol and Eli Lilly 2021 BRAVE-AA2 Protocol. The BRAVE-AA1 and BRAVE-AA2 CSRs and protocols suggest that [redacted]. Please:

a) **clarify the definition of relapse used in the trials and;**

In both BRAVE-AA trials, a relapse was measured after Week 52 and defined as a >20-point absolute worsening in total SALT score (e.g. from SALT 10 to SALT 30), in patients receiving baricitinib or placebo who had achieved a SALT≤20 at Week 52.

b) **provide data on the number of patients who experienced relapse by arm and week, if available.**

Since relapse was only measured from Week 52 in patients receiving baricitinib or placebo who had achieved a SALT≤20 at Week 52, we do not have these data for the primary study phase.

Statistical analyses

A11. CS, Document B, section B.2.6.1, Table 15 and section B.2.6.2, Tables 20 and 21. Please provide a version of Table 15 (Proportion of patients achieving SALT≤20 at Week 36), Table 20 (Proportion of patients achieving $SALT_{50}$ at Week 36) and Table 21 (Proportion of patients achieving $SALT_{75}$ at Week 36) in which the analyses have been performed using duration of current episode and baseline SALT scores as continuous predictors in the primary logistic regression analysis.

The proportion of patients achieving SALT≤20, $SALT_{50}$ and $SALT_{75}$ at Week 36 across BRAVE-AA1 and BRAVE-AA2, where the duration of current AA episode and baseline SALT scores have been used as continuous predictors in the primary logistic regression analysis, are presented in Table 2, Table 3 and Table 4, respectively.

Table 2. Proportion of patients achieving SALT≤20 at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	Placebo (N=189)	Baricitinib 2 mg (N=184)	Baricitinib 4 mg (N=281)	Placebo (N=156)	Baricitinib 2 mg (N=156)	Baricitinib 4 mg (N=234)
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Odds ratio (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, continuous duration of current episode at baseline and continuous baseline SALT score as factors.
Abbreviations: CI: confidence interval; FAS: full analysis set; NRI: non-responder imputation; SALT: severity of alopecia tool.

Table 3. Proportion of patients achieving $SALT_{50}$ at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	Placebo (N=189)	Baricitinib 2 mg (N=184)	Baricitinib 4 mg (N=281)	Placebo (N=156)	Baricitinib 2 mg (N=156)	Baricitinib 4 mg (N=234)
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, continuous duration of current episode at baseline and continuous baseline SALT score as factors.
Abbreviations: CI: confidence interval; FAS: full analysis set ; NRI: nonresponder imputation; SALT: severity of alopecia tool.

Table 4. Proportion of patients achieving $SALT_{75}$ at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS population; primary censoring rule [NRI])

	BRAVE-AA1			BRAVE-AA2		
	Placebo (N=189)	Baricitinib 2 mg (N=184)	Baricitinib 4 mg (N=281)	Placebo (N=156)	Baricitinib 2 mg (N=156)	Baricitinib 4 mg (N=234)
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, continuous duration of current episode at baseline and continuous baseline SALT score as factors.

Abbreviations: CI: confidence interval; FAS: full analysis set; NRI: nonresponder imputation; SALT: severity of alopecia tool.

A12. CS, Document B, section B.2.4.3. The company submission states that non-responder imputation and modified last observation carried forward were used to impute missing data in the analyses. In the primary publication of BRAVE-AA1 and BRAVE-AA2 (King et al. 2022, DOI: 10.1056/NEJMoa2110343), multiple imputation was used to align with journal requirements. Please clarify whether the analyses from the primary publication are otherwise identical to those presented in the CS, and the slightly higher SALT≤20 response rates in the publication for 4mg baricitinib at Week 36 (38.8% for BRAVE-AA1, 35.9% for BRAVE-AA2) are due to the imputation approach only.

The Company confirms that the above interpretation is correct; the analyses from the primary publication are otherwise identical to those presented in the CS, and the slightly higher SALT≤20 response rates in the publication for 4 mg baricitinib at Week 36 (38.8% for BRAVE-AA1, 35.9% for BRAVE-AA2) are due to the use of the less conservative multiple imputation approach only.

A13. Please provide the number of patients who were missing SALT data at Week 36 and at Week 52 for the 4 mg baricitinib and placebo arms of BRAVE-AA1 and BRAVE-AA2.

At Week 36, non-responders in the placebo arm (i.e., patients who did not achieve SALT≤20) were rescued and re-randomised in a 1:1 ratio to baricitinib 2 mg or baricitinib 4 mg. The concept of missing data does not therefore apply to almost all patients in the placebo arm beyond Week 36, and the few observations at Week 52 in the placebo arm were due to the few patients responding in the placebo arm. Table 5 therefore presents the number of patients who were missing SALT data for *all treatment arms at Week 36*, but only across the *2 mg and 4mg baricitinib treatment arms for Week 52*.

Table 5. Number of patients with observed SALT data at Week 36 and Week 52 in BRAVE-AA1 and BRAVE-AA2.

	BRAVE-AA1			BRAVE-AA2		
	Placebo (N=189)	Baricitinib 2 mg (N=184)	Baricitinib 4 mg (N=281)	Placebo (N=156)	Baricitinib 2 mg (N=156)	Baricitinib 4 mg (N=234)
Week 36						
Number of patients with observed SALT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Number of patients with missing SALT data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 52						

Number of patients with observed SALT[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Number of patients with missing SALT data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a] only patients still on placebo, baricitinib 2 mg or baricitinib 4 mg.
Abbreviations: NA: not applicable; SALT: severity of alopecia tool.

Subgroup analyses

A14. Priority question. CS, Appendices, Appendix E. Please provide a subgroup analysis similar to that presented in Appendix E for the probability of achieving SALT≤20 at Week 36 by atopic background status, covariate adjusted for baseline SALT score, in BRAVE-AA1 and BRAVE-AA2.

The proportion of patients in BRAVE-AA1 and BRAVE-AA2 achieving SALT≤20 at Week 36 according to atopic background status, using an adjusted covariate for baseline SALT score, is presented in Table 6 and Table 7.

Table 6. Proportion of patients with no atopic background in BRAVE-AA1 and BRAVE-AA2 achieving SALT ≤20 at Week 36 (FAS population)

	BRAVE-AA1			BRAVE-AA2		
	Placebo (N=116)	Baricitinib 2 mg (N=117)	Baricitinib 4 mg (N=184)	Placebo (N=89)	Baricitinib 2 mg (N=93)	Baricitinib 4 mg (N=147)
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs. PBO[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Atopic background is defined as having an ongoing or a medical history of atopic dermatitis, allergic rhinitis, allergic conjunctivitis, or allergic asthma [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at Baseline (<4 years vs. >=4 years), and baseline SALT score as factors.
Abbreviations: CI: confidence interval; FAS: full analysis set; PBO: placebo; SALT: severity of alopecia tool.

Table 7. Proportion of patients with an atopic background in BRAVE-AA1 and BRAVE-AA2 achieving SALT ≤20 at Week 36 (FAS population)

	BRAVE-AA1			BRAVE-AA2		
	Placebo (N=73)	Baricitinib 2 mg (N=67)	Baricitinib 4 mg (N=97)	Placebo (N=67)	Baricitinib 2 mg (N=63)	Baricitinib 4 mg (N=87)
Response, n (%) (95% CI) [a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio (95% CI) vs PBO	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs. PBO [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Atopic background is defined as having an ongoing or a medical history of atopic dermatitis, allergic rhinitis, allergic conjunctivitis, or allergic asthma [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at Baseline (<4 years vs. >=4 years), and baseline SALT score as factors.
Abbreviations: CI: confidence interval; FAS: full analysis set; PBO: placebo; SALT: severity of alopecia tool.

Literature searching

A15. CS, Document B, section B.2.9.1. Please provide a list of the 22 studies not included in the network meta-analysis (NMA) feasibility assessment along with the study specific reasons why they were excluded. Please include which primary efficacy endpoints were reported for those excluded for a lack of similar endpoints.

The 22 studies not included in the NMA feasibility assessment and reasons for exclusion are presented in Table 8.

Table 8. Studies not included in the feasibility assessment and reasons for exclusion

Study	Intervention	Reason for exclusion
King 2021 (ALLEGRO)	Ritlecitinib; brepocitinib	Not submitted to regulatory approval and/or not an established option for clinical management.
Almutairi 2019	Ruxolitinib; tofacitinib	
AlMarzoug 2021	Tofacitinib	
Chen 2019	Tofacitinib	
Cheng 2018	Tofacitinib	
Crispin 2016	Tofacitinib	
Dincer 2021	Tofacitinib	
Hogan 2019	Tofacitinib	
Ibrahim 2017	Tofacitinib	

Jabbari 2018	Tofacitinib	
Kerkemeyer 2020	Tildrakizumab	
Liu 2017	Tofacitinib	
Mackay-Wiggan 2016	Ruxolitinib	
Park 2017	Tofacitinib	
Serdaroğlu 2019	Tofacitinib	
Wambier 2021	Tofacitinib + minoxidil	
Avgerinou 2008	DCP	Reported endpoints not comparable *Endpoints reported: Time to response and grade of response (MacDonald Hull and Norris 4 category grading system). No fixed timepoint specified.*
Cotellessa 2001	DPCP	Reported endpoints not comparable *Endpoints reported: Grade of response (MacDonald Hull and Norris 4 category grading system) at 6 months.*
Case 1984	SADBE	Reported endpoints not comparable *Endpoints reported: Complete or cosmetically acceptable regrowth. No fixed timepoint specified.*
Dehghan 2013	Prednisolone	Reported endpoints not comparable *Endpoints reported: Grades of improvement (4 categories) at 1,3,6,12 months.*
Hull 1988	DCP	Reported endpoints not comparable *Endpoints reported: (MacDonald Hull and Norris 4 category grading system) at 8 months.*
Yoshimasu 2016	Methylprednisolone	Reported endpoints not comparable *Endpoints reported: Clinical response (development of vellus hair) <6 months and ≥6 months by category of baseline severity.*

Abbreviations: DCP: diphencyprone; DPCP: 2,3-diphenylcyclopropenone; NMA: network meta-analysis; SADBE: squaric acid dibutylester.

A16. CS, Appendices, Appendix D. In Section D.1.1. of the appendix, please confirm whether the British Association of Dermatologists (BAD) conference in 2021 was hand searched.

Table 4 within Appendix D of the CS contained a typographical error; this error has been corrected in Table 9 below. As such, the Company confirms that the British Association of Dermatologists conference was hand searched.

Table 9. Details of conference search methods

Conference Name	Year 2019	Year 2020	Year 2021
American Academy of Dermatology	Embase	Embase	Hand search
Annual Alopecia Areata Conference (National Alopecia Areata Foundation)	Hand search	Hand search	Hand search
British Association of Dermatologists	Embase	Embase	Hand search

European Academy of Dermatology and Venereology Congress	Hand search	Hand search	Hand search
European Society for Dermatological Research	Hand search	Hand search	Hand search
Society of Investigative Dermatology	Embase	Embase	Hand search

A17. CS, Document B, section B.2.1 and Appendices, Appendix D. The company submission states that, "*In total, the updated SLR identified 47 unique studies, with evidence generated from 13 RCTs and 34 observational studies*" (p28). However, Figure 2 in Appendix D suggests that there were only 45 studies from 47 records included. In addition, Table 12 in Appendix D only includes 12 randomised controlled trials (RCTs).

a) **Please clarify the total number of studies that were included from the original and updated clinical effectiveness systematic literature reviews (SLRs). Please provide details of how many were RCTs and whether the number of included studies comprises BRAVE-AA2.**

The SLR included 45 unique studies, from 47 records; Lai (2020)[8] and Lai (2019)[9] report interim and final results from the same study, respectively, while Liu (2017) reports an update of an original publication by Crispin (2016).

Of the 45 unique studies included in the SLR, 12 of them were RCTs, as outlined in Figure 2 in Appendix D of the CS. However, it should be noted that the 12 RCTs included in the feasibility assessment and the SLR differ by one study; the feasibility assessment includes BRAVE-AA2, which was not captured in the SLR (King 2022 was published after the SLR update was conducted) but excludes King 2021 (ALLEGRO).[10]

b) **Please update Tables 12 and 13 (if necessary) in Appendix D3 to provide quality assessments for all included studies.**

As explained in subsection (a) above, 45 studies were included in the SLR. Quality assessments have been performed for all 45 studies in Tables 12 and 13 in Appendix D3 of the CS, and so no updates are required. Furthermore, a quality assessment of BRAVE-AA2 is presented in Question A18.

c) **Please clarify which of the King 2021 studies are being referred to in rows 6 and 7 of Table 12 in Appendix D3.**

Full details of the King (2021) studies in Table 12 in Appendix D3 of the CS are provided below in Table 10.

Table 10. King 2021 studies clarification

Row	Study
6	Ritlecitinib and brepocitinib trial[10]
7	BRAVE-AA1 trial[11]

A18. CS, Document B, section B.2.5. Please provide justification for each of the risk of bias judgements in Table 14 of the critical appraisal of BRAVE-AA1 and BRAVE-AA2.

Table 11 presents full justifications for the risk of bias judgements for the critical appraisal of BRAVE-AA1 and BRAVE-AA2.

Table 11. Justifications for the risk of bias judgements for BRAVE-AA1 and BRAVE-AA2

Question	Risk of bias justification	
	BRAVE-AA1	BRAVE-AA2
Was the method used to generate random allocations adequate?	Assignment to treatment groups was determined by a computer-generated random sequence using an interactive web-response system (IWRS).	
Was the allocation adequately concealed?	Patients and site personnel remained blinded to treatment allocation.	
Were the groups similar at the outset of the study in terms of prognostic factors?	Demographic characteristics and disease history were overall well balanced among treatment groups. Furthermore, among the observed laboratory abnormalities, the data showed no clinically concerning imbalances between treatment groups.	
Were the care providers, participants, and outcome assessors blind to treatment allocation?	The investigator, site personnel performing assessments and patient were blinded to treatment allocations.	
Were there any unexpected imbalances in dropouts between groups?	There were few dropouts across all treatment arms overall, and there were no unexpected imbalances due to these dropouts.	
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Outcomes introduced in methods are also reported as results.	
Did the analysis include an intention–to–treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The analysis included an intention-to-treat analysis in the FAS. The efficacy analysis of the primary and key secondary endpoints was conducted in the FAS population (essentially the ITT population; includes all randomised patients, analysed according to study intervention to which they were randomised at baseline). Data collected remotely due the COVID-19 pandemic were included in secondary censoring rule analyses. The hybrid imputation was implemented to handle missing data due to the COVID-19 pandemic by multiple imputation or other missing data not due to COVID-19 by nonresponder imputation or mLOCF. Protocol deviations and missing data did not have an impact on the primary or secondary endpoints of this study.	

Abbreviations: FAS: full analysis set; ITT: intention-to-treat; mLOCF: modified last observation carried forward.
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report; BRAVE-AA1 Study protocol; BRAVE-AA2 study protocol.[1-4, 12]

A19. CS, Appendices, Appendix D.1.3. Please provide a list of references for the studies data extracted following identification in the clinical SLR for the:

a) 5 records included in Category 1

b) 42 records included in Category 2.

The full list of references included in Category 1 and 2 is presented in Table 12.

Table 12. List of records included in Category 1 and 2 of the clinical SLR

Category 1	
Kar 2005	Kar BR, Handa S, Dogra S, et al. Placebo-controlled oral pulse prednisolone therapy in alopecia areata. J Am Acad Dermatol 2005;52:287-90.
King 2021 (BRAVE-AA1 trial)	King B, Ko J, Forman S, et al. Efficacy and safety of the oral Janus kinase inhibitor baricitinib in the treatment of adults with alopecia areata: Phase 2 results from a randomized controlled study. Journal of the American Academy of Dermatology 2021.
King 2021 (ALLEGRO trial)	King B, Guttman-Yassky E, Peeva E, et al. A phase 2a randomized, placebo-controlled study to evaluate the efficacy and safety of the oral Janus kinase inhibitors ritlecitinib and brepocitinib in alopecia areata: 24-week results. Journal of the American Academy of Dermatology 2021.
Lai 2020 (interim)	Lai V, Chen G, Gin D, et al. Cyclosporine for moderate to severe alopecia areata: Interim analysis of a double-blind, randomised, placebocontrolled clinical trial of efficacy and safety. Journal of the Dermatology Nurses' Association 2020;12.
Lai 2019 (final)	Lai VWY, Chen G, Gin D, et al. Cyclosporine for moderate-to-severe alopecia areata: A double-blind, randomized, placebo-controlled clinical trial of efficacy and safety. Journal of the American Academy of Dermatology 2019;81:694-701.
Category 2	
Al Bazzal 2021	Al Bazzal A, Hatami P, Abedini R, et al. A prospective comparative study of two regimens of diphenylcyclopropenone (DPCP) in the treatment of alopecia areata. Int Immunopharmacol 2021;101:108186.
AlMarzoug 2021	AlMarzoug A, AlOrainy M, AlTawil L, et al. Alopecia areata and tofacitinib: a prospective multicenter study from a Saudi population. Int J Dermatol 2021.
Almutairi 2019	Almutairi N, Nour TM, Hussain NH. Janus Kinase Inhibitors for the Treatment of Severe

	Alopecia Areata: An Open-Label Comparative Study. Dermatology 2019;235:130-136.
Alsufyani 2017	Alsufyani HS, Rawas WA, Alsufyani SS, et al. The effect of mehotrexate in the treatment of alopecia areata. The Egyptian Journal of Hospital Medicine 2017;67:599-604.
Asilian 2021	Asilian A, Fatemi F, Ganjei Z, et al. Oral pulse betamethasone, methotrexate, and combination therapy to treat severe alopecia areata: A randomized, double-blind, placebo-controlled, clinical trial. Iranian Journal of Pharmaceutical Research 2021;20:267-273.
Avgerinou 2008	Avgerinou G, Gregoriou S, Rigopoulos D, et al. Alopecia areata: Topical immunotherapy treatment with diphencyprone. Journal of the European Academy of Dermatology and Venereology 2008;22:320-323.
Case 1984	Case PC, Mitchell AJ, Swanson NA, et al. Topical therapy of alopecia areata with squaric acid dibutylester. J Am Acad Dermatol 1984;10:447-50.
Chen 2019	Chen YY, Lin SY, Chen YC, et al. Low-dose tofacitinib for treating patients with severe alopecia areata: an efficient and cost-saving regimen. Eur J Dermatol 2019;29:667-669.
Cheng 2018	Cheng MW, Kehl A, Worswick S, et al. Successful Treatment of Severe Alopecia Areata With Oral or Topical Tofacitinib. J Drugs Dermatol 2018;17:800-803.
Cotellessa 2001	Cotellessa C, Peris K, Caracciolo E, et al. The use of topical diphenylcyclopropenone for the treatment of extensive alopecia areata. Journal of the American Academy of Dermatology 2001;44:73-76.
Crispin 2016	Kennedy Crispin M, Ko JM, Craiglow BG, et al. Safety and efficacy of the JAK inhibitor tofacitinib citrate in patients with alopecia areata. JCI insight 2016;1:e89776.
Dehghan 2013	Dehghan M, Alborzi A, Shahini N. Comparison of oral prednisolone pulse therapy with intravenous methylprednisolone pulse therapy in severe alopecia areata. Journal of Pakistan Association of Dermatologists 2013;23:159-162.
Dincer 2021	Dincer Rota D, Emeksiz MAC, Erdogan FG, et al. Experience with oral tofacitinib in severe alopecia areata with different clinical responses. J Cosmet Dermatol 2021;20:3026-3033.
English 2015	English J, Heinisch S. Methotrexate treatment for alopecia areata with greater than 50% involvement. Hair Ther Transplant 2015;5:2.
Ferrando 1999	Ferrando J, Grimalt R. Partial response of severe alopecia areata to cyclosporine A. Dermatology 1999;199:67-9.
Firooz 2013	Firooz A, Fouladi D. Methotrexate plus prednisolone in severe alopecia areata.

	American Journal of Drug Discovery and Development 2013;3:188-193.
Ghandi 2021	Ghandi N, Daneshmand R, Hatami P, et al. A randomized trial of diphenylcyclopropenone (DPCP) combined with anthralin versus DPCP alone for treating moderate to severe alopecia areata. Int Immunopharmacol 2021;99:107971.
Gupta 1990	Gupta AK, Ellis CN, Cooper KD, et al. Oral cyclosporine for the treatment of alopecia areata: a clinical and immunohistochemical analysis. Journal of the American Academy of Dermatology 1990;22:242-250.
Hogan 2019	Hogan S, Wang S, Ibrahim O, et al. LONG-TERM TREATMENT WITH TOFACITINIB IN SEVERE ALOPECIA AREATA: AN UPDATE. The Journal of clinical and aesthetic dermatology 2019;12:12-14.
Hull 1988	Hull SM, Norris JF. Diphencyprone in the treatment of long-standing alopecia areata. The British journal of dermatology 1988;119:367-74.
Ibrahim 2017	Ibrahim O, Bayart CB, Hogan S, et al. Treatment of Alopecia Areata With Tofacitinib. JAMA Dermatol 2017;153:600-602.
Jabbari 2018	Jabbari A, Sansaricq F, Cerise J, et al. An Open-Label Pilot Study to Evaluate the Efficacy of Tofacitinib in Moderate to Severe Patch-Type Alopecia Areata, Totalis, and Universalis. Journal of Investigative Dermatology 2018;138:1539-1545.
Jang 2016	Jang YH, Kim SL, Lee KC, et al. A comparative study of oral cyclosporine and betamethasone minipulse therapy in the treatment of alopecia areata. Annals of Dermatology 2016;28:569-574.
Joly 2006	Joly P. The use of methotrexate alone or in combination with low doses of oral corticosteroids in the treatment of alopecia totalis or universalis. Journal of the American Academy of Dermatology 2006;55:632-636.
Kerkemeyer 2020	Kerkemeyer KLS, Sinclair R. Treatment of chronic alopecia areata with tildrakizumab: an open-label pilot study. Int J Dermatol 2020;59:e136-e137.
Kurosawa 2006	Kurosawa M, Nakagawa S, Mizuashi M, et al. A comparison of the efficacy, relapse rate and side effects among three modalities of systemic corticosteroid therapy for alopecia areata. Dermatology 2006;212:361-365.
Liu 2017	Liu LY, Craiglow BG, Dai F, et al. Tofacitinib for the treatment of severe alopecia areata and variants: A study of 90 patients. Journal of the American Academy of Dermatology 2017;76:22-28.
Mackay-Wiggan 2016	Mackay-Wiggan J, Jabbari A, Nguyen N, et al. Oral ruxolitinib induces hair regrowth in patients

	with moderate-to-severe alopecia areata. JCI insight 2016;1:e89790.
Maryam 2009	Maryam A, Hassan S, Farshad F, et al. The efficacy of topical diphencyprone in the treatment of alopecia areata. Indian journal of dermatology 2009;54:88-89.
Park 2017	Park HS, Kim MW, Lee JS, et al. Oral tofacitinib monotherapy in Korean patients with refractory moderate-to-severe alopecia areata: A case series. J Am Acad Dermatol 2017;77:978-980.
Rocha 2021	Rocha VB, Kakizaki P, Donati A, et al. Randomized controlled study comparing the use of diphencyprone and anthralin in the treatment of extensive chronic alopecia areata. Anais brasileiros de dermatologia 2021;96:372-376.
Serdaroğlu 2019	Serdaroğlu S, Engin B, Çelik U, et al. Clinical experiences on alopecia areata treatment with tofacitinib: A study of 63 patients. Dermatol Ther 2019;32:e12844.
Shapiro 1997	Shapiro J, Lui H, Tron V, et al. Systemic cyclosporine and low-dose prednisone in the treatment of chronic severe alopecia areata: a clinical and immunopathologic evaluation. Journal of the American Academy of Dermatology 1997;36:114-117.
Shapiro 1993	Shapiro J, Tan J, Ho V, et al. Treatment of chronic severe alopecia areata with topical diphenylcyclopropenone and 5% minoxidil: a clinical and immunopathologic evaluation. J Am Acad Dermatol 1993;29:729-35.
Shin 2018	Shin J-W, Huh C-H, Kim M-W, et al. Comparison of the treatment outcome of oral tofacitinib with other conventional therapies in refractory alopecia totalis and universalis: A retrospective study. Acta Dermato-Venereologica 2018;99:41-46.
Sriphojanart 2017	Sriphojanart T, Khunkhet S, Suchonwanit P. A retrospective comparative study of the efficacy and safety of two regimens of diphenylcyclopropenone in the treatment of recalcitrant alopecia areata. Dermatology Reports 2017;9:55-58.
Thuangtong 2017	Thuangtong R, Varothai S, Triwongwaranat D, et al. Multi-concentration level patch test guided diphenyl cyclopropenone (DPCP) treatment in alopecia totalis or alopecia universalis. Journal of the Medical Association of Thailand 2017;100:86-92.
Tiwary 2016	Tiwary AK, Mishra DK, Chaudhary SS. Comparative study of efficacy and safety of topical squaric acid dibutylester and diphenylcyclopropenone for the treatment of alopecia areata. North American Journal of Medical Sciences 2016;8:237-242.
Vano-Galvan 2016	Vañó-Galván S, Hermosa-Gelbard Á, Sánchez-Neila N, et al. Pulse corticosteroid therapy with

	oral dexamethasone for the treatment of adult alopecia totalis and universalis. J Am Acad Dermatol 2016;74:1005-7.
Vano-Galvan 2016	Vañó-Galván S, Hermosa-Gelbard Á, Sánchez-Neila N, et al. Treatment of recalcitrant adult alopecia areata universalis with oral azathioprine. J Am Acad Dermatol 2016;74:1007-8.
Wambier 2021	Wambier CG, Craiglow BG, King BA. Combination tofacitinib and oral minoxidil treatment for severe alopecia areata. J Am Acad Dermatol 2021;85:743-745.
Yoshimasu 2016	Yoshimasu T, Kanazawa N, Yamamoto Y, et al. Multiple courses of pulse corticosteroid therapy for alopecia areata. Journal of Dermatology 2016;43:1075-1077.

A20. CS, Appendices, Appendix D.1.3. Please clarify the definition of "mostly adults with mild-to-moderate form of AA" used in Category 3 (Table 9) for the categorisation of records identified by the SLR for data extraction.

The definition of Category 3 used for the categorisation of records in the SLR includes a typographical error and should instead read "adults with mostly mild-to-moderate forms of AA". Studies which had a majority of patients with mild-to-moderate forms of AA were therefore allocated to Category 3, as these studies were considered to be of less relevance to this appraisal, given that baricitinib is indicated for severe AA patients, and not those with mild-to-moderate forms of the disease.

A21. CS, Appendices, Appendix D.1.2. In the eligibility criteria (Table 8), non-English articles were listed as to be included in the SLR. However, in the PRISMA diagram, non-English articles were listed as being excluded. Please confirm the eligibility criterion used for article language.

The PRISMA diagram presented in Appendix D.1.2 of the CS includes a typographic error; non-English articles were indeed included in the SLR, but were excluded at a later screening stage for reasons other than language.

A22. CS, Appendices, Appendix D.1.2 and D.1.5. Please confirm the exact eligibility criterion used for patient age in the SLR. Currently, articles that had inclusion criteria containing patients aged <18 have both been included (for example, Asilian 2021) and excluded (for example, Zerbinati 2018) from the SLR.

The exact eligibility criteria used for patients age in the SLR was individuals ≥16 years of age, as this aligned with the definition of adults that was across many studies in AA. Using this definition, Asilian 2021 was correctly included (age range: 16–60 years), while Zerbinati 2018 was excluded (age range: 5–61 years old).[13, 14]

Section B: Clarification on cost-effectiveness data

For any scenarios requested in Section B, please ensure these are implemented as user selectable options in the economic model ("Main" tab). If scenarios cannot be implemented as user selectable options, please supply instructions on how to replicate the scenario. Furthermore, if the company chooses to update its base case analysis, please ensure that cost-effectiveness results, sensitivity and scenario analyses incorporating the revised base case assumptions are provided with the response along with a log of changes made to the company base case.

The Company base case has been updated based on the final analysis of the Adelphi DSP (only the interim analysis was available at the time of submission), updated NHS reference costs, and errors/updates suggested by the EAG in Questions B14 and B15. Full details of all the changes made and cost-effectiveness outcomes are provided in the Appendix.

Treatment response

B1. Priority question. The EAG's clinical experts advised that an absolute reduction in SALT scores is a more clinically meaningful estimate of response than a relative improvement from baseline SALT scores. Please provide a scenario where treatment response in the model is based on the primary endpoint of achieving SALT≤20 from the BRAVE-AA1 and BRAVE-AA2 trials. Please combine this scenario with the utility values requested in question B8.

A relative improvement from baseline SALT score ($SALT_{50}$ and $SALT_{75}$) was used as the definition of response in the CEM rather than the achievement of an absolute SALT score (such as SALT≤10 or SALT≤20) for several reasons. Firstly, the use of an improvement from baseline response measure is aligned with modelling precedents in other dermatology indications, including atopic dermatitis ([TA681] [TA466]) and psoriasis [TA350].[15-17] In addition, the $SALT_{75}$ response is used in the model when $SALT_{50}$ (the base case analysis) is selected in order to obtain a more granular calculation of the total QALYs. The $SALT_{75}$ response is relatively aligned with the SALT≤20 endpoint from the BRAVE-AA trials. This is demonstrated in Table 13, which shows the greatest absolute SALT score (i.e., the poorest hair regrowth) that would be achieved after a $SALT_{50}$ and $SALT_{75}$ response (improvement in SALT score of 50% and 75%, respectively), based on the mean baseline SALT scores from the BRAVE-AA1 and BRAVE-AA2 trials.

It would therefore be expected that the use of SALT≤20 in the CEM would be relatively aligned with the scenario analyses presented in Document B Section B.3.11.1.2 in the CS, where $SALT_{75}$ is used as the definition of response in the model. It would further be expected that the use of SALT≤20 in the model would be associated with similar challenges to the $SALT_{75}$ response, which the Company considered to result in this definition of response being overly restrictive and

failing to adequately capture sufficient clinical benefit to justify continuing treatment after the trial induction period than using $SALT_{50}$. These issues include the gradual nature of hair regrowth and the subsequent continuous improvement in hair regrowth beyond Week 36 in some patients, as well as the possibility that concomitant pattern baldness may lead to a ceiling on the maximum possible response, meaning that some patients may not be able to achieve SALT≤20 within 36 weeks, or indeed at all. As such, the use of SALT≤20 as the response definition in the CEM would deem these patients non-responders, despite the fact they may either achieved their maximum possible hair regrowth or may achieve their maximum hair regrowth if treatment is continued.

Table 13. SALT score after achievement of $SALT_{50}$ and $SALT_{75}$ based on the mean baseline SALT score from BRAVE-AA1 and BRAVE-AA2

	BRAVE-AA1			BRAVE-AA2		
	PBO	Baricitinib 2 mg	Baricitinib 4 mg	PBO	Baricitinib 2 mg	Baricitinib 4 mg
Mean baseline SALT score (FAS population)	84.7	86.8	85.3	85	85.6	84.8
Maximum absolute SALT score after achievement of $SALT_{50}$	42.4	43.4	42.7	42.5	42.8	42.4
Maximum absolute SALT score after achievement of $SALT_{75}$	21.2	21.7	21.3	21.3	21.4	21.2

Abbreviations: FAS: full analysis set; SALT: severity of alopecia tool.
Source: BRAVE-AA1 Clinical Study Report; King et al. 2022; EPAR.[1, 12, 18]

Nonetheless, for transparency, the requested analyses are presented below in Table 14, which use the SALT≤20 response rates in the pooled population presented in Table 15.

Table 14. Proportion of patients achieving SALT≤20 at Week 36 (pooled FAS population; primary censoring rule [NRI])

	Placebo (N=345)	Baricitinib 2 mg (N=340)	Baricitinib 4 mg (N=515)
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction.
[b] Logistic regression analysis with treatment group, study, geographic region, duration of current episode at Baseline (<4 vs. ≥4 years), and baseline total SALT score as factors.
Abbreviations: FAS: full analysis set; NA: not applicable; SALT: severity of alopecia areata Tool.

The cost-effectiveness outcomes of the scenario in which an endpoint of SALT≤20 is used as the definition of response in the model are presented in Table 15. Utility values from the BRAVE-AA trials have not been used in this analysis, for reasons outlined in Question B9 (d)(i).

Table 15: Cost-effectiveness outcomes using SALT≤20 (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,312

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

a) Please run an alternative scenario exploring the outcome of SALT≤10 from the BRAVE-AA1 and BRAVE-AA2 trials. Please combine this scenario with the utility values requested in question B8.

The Company does not consider the use of SALT≤10 to be as clinically meaningful as the current definition of response in the CEM, as concomitant pattern baldness may lead to a ceiling on the maximum possible response, and this effect is more likely to be observed with more stringent endpoints such as SALT≤10. Furthermore, as discussed above in Question B1, the Company considers the use of the achievement of an improvement from baseline response ($SALT_{50}$ and $SALT_{75}$) to be consistent with previous appraisals in the field of dermatology.[15, 17, 19] Nonetheless, for transparency, the requested analyses are presented below.

The proportion of patients achieving SALT≤10 at Week 36 across the BRAVE-AA trials in the pooled population is summarised in Table 16 and the cost-effectiveness outcomes of the scenario using a response level of SALT≤10 are presented in Table 17. Utility values from the BRAVE-AA trials have not been used in this analysis, for reasons outlined in Question B9(d)(i).

Table 16. Proportion of patients achieving SALT≤10 at Week 36 (pooled FAS population; primary censoring rule [NRI])

	Placebo	Baricitinib 2 mg	Baricitinib 4 mg
Week 36			
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: FAS: full analysis set; SALT: Severity of Alopecia Areata Tool; NA: not applicable.

Table 17: Cost-effectiveness outcomes using SALT≤10 (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£21,254

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

b) Please run the SALT≤20 scenario for the severe and very severe subgroups.

The proportion of patients achieving SALT≤20 at Week 36 in the severe and very severe populations in the pooled population is summarised in Table 18.

Table 18. Proportion of patients achieving SALT≤20 at Week 36 in severe and very severe subgroups (pooled population; primary censoring rule [NRI])

SALT 50–94 patients (severe population)			
Intervention	**Placebo** [redacted]	**Baricitinib 2 mg** [redacted]	**Baricitinib 4 mg** [redacted]
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]
SALT 95–100 patients (very severe population)			
Intervention	**Placebo** [redacted]	**Baricitinib 2 mg** [redacted]	**Baricitinib 4 mg** [redacted]
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: SALT: Severity of Alopecia Areata Tool; NA: not applicable.

The cost-effectiveness outcome of the scenario using a response level of SALT≤20 in the severe and very severe subgroups are presented in Table 19 and Table 20, respectively. Utility values from the BRAVE-AA trials have not been used in this analysis, for reasons outlined in Question B9(d)(i).

Table 19: Cost-effectiveness outcomes using SALT≤20 in the severe population (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,598

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 20: Cost-effectiveness outcomes using SALT≤20 in the very severe population (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,196

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B2. Priority question. Please provide subgroup analysis based on duration of current AA episode, using the prespecified stratification of ≤4 years and >4

years in BRAVE-AA1 and BRAVE-AA2. Please ensure that the outcome of SALT≤20 is used in the subgroup analysis.

The proportion of patients achieving SALT≤20 at Week 36 in the pooled population, stratified by duration of current AA episode (≤4 years and >4 years), is summarised in Table 21.

Table 21. Proportion of patients achieving SALT≤20 at Week 36 according to duration of current AA episode (pooled population; primary censoring rule [NRI])

≤4 years			
Intervention	**Placebo (N=228)**	**Baricitinib 2 mg (N=230)**	**Baricitinib 4 mg (N=329)**
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]
>4 years			
Intervention	**Placebo (N=117)**	**Baricitinib 2 mg (N=110)**	**Baricitinib 4 mg (N=186)**
Response, n (%) (95% CI)[a]	[redacted]	[redacted]	[redacted]
p-Value vs. PBO[b]	[redacted]	[redacted]	[redacted]

Footnotes: [a] Difference confidence intervals are constructed using Newcombe-Wilson method, without continuity correction. Response confidence intervals are constructed using Wilson method, without continuity correction. [b] Logistic regression analysis with treatment group, geographic region, duration of current episode at baseline (<4 years vs. ≥4 years), and baseline SALT score as factors.
Abbreviations: SALT: Severity of Alopecia Areata Tool; NA: not applicable.

The cost-effectiveness outcomes of the scenario using a response level of SALT≤20 and a duration of episode <4 years is presented in Table 22. Table 23 presents the cost-effectiveness outcome of the scenario using a response level of SALT≤20 and a duration of episode >4 years; all other inputs remained consistent with the base case analysis.

Table 22: Cost-effectiveness outcomes using SALT≤20 and duration of episode <4 years (deterministic)

	Total cost	**Total QALYs**	**Incremental Cost**	**Incremental QALYs**	**ICER**
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£16,154

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 23: Cost-effectiveness outcomes using SALT≤20 and duration of episode >4 years (deterministic)

	Total cost	**Total QALYs**	**Incremental Cost**	**Incremental QALYs**	**ICER**
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,982

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B3. Priority question. CS, Document B, section B.2.8.2. The summary of product characteristics (SmPC) states that, "*a dose of 2mg once daily may be appropriate for patients such as those aged ≥ 75 years and for patients with a history of chronic or recurrent infections. A dose of 2mg once daily may also be considered for patients who have achieved sustained control of disease activity with 4mg once daily and are eligible for dose tapering*".

a) **Please provide a rationale for not exploring dose tapering to 2mg baricitinib for patients who achieved a sustained response on 4mg baricitinib.**

Baricitinib 2 mg does not incur a lower cost to the NHS than baricitinib 4 mg. As such, modelling of dose tapering in patients exhibiting sustained response is not expected to affect the ICER as neither cost nor response would be affected; at an individual patient level, any patient whose response began to fail due to dose tapering would be expected to resume the 4 mg dose to restore the previously sustained response.

b) **Please provide a scenario where patients who achieved a sustained response on 4mg baricitinib are down titrated to 2mg, applying the appropriate SALT≤20 efficacy data associated with down titration from the BRAVE extension studies (Figure 24).**

As explained in the answer to subsection (a) above, this analysis is not expected to be informative as any indication of incipient loss of response due to dose tapering would be immediately addressed through resumption of the standard dose, which cannot be modelled from the extension study data where patients remained on their re-randomised dose, while costs remain unchanged in any case.

c) **Please provide a scenario where the time horizon is capped at 75 years of age.**

As age is a protected characteristic under the Equality Act, the Company understand that the requested analysis could not form the basis for a recommendation and have therefore chosen not to present this; it may be noted that the effect of shorter model time horizons, not directly linked to patient age, is explored in the response to Question B20.

Furthermore, as noted above, any use of the 2 mg dose due to reduced renal function in older patients will not affect the costs in the model. Given that usage of the 2 mg dose in such patients would most likely be affected by reduced renal function leading to increased exposure to baricitinib in such patients, the data from the 2 mg arm of the clinical trials may not be applicable and the base case model time horizon remains the most informative analysis.

B4. CS, Document B, section B.3.3.3. For the discontinuation data used in the model, please clarify the definition used for lack of efficacy.

Lack of treatment effect was not defined in the study BRAVE-AA1 and BRAVE-AA2 study protocols and as such was recorded following the investigators decision. Up to week 36, if a patient discontinued treatment the investigator was invited to select a reason for discontinuation from a predefined list. Among the possible reasons was lack of efficacy, which was not defined and thus left to investigator judgment. Other possible reasons included withdrawal by subject, pregnancy, protocol deviation, lost to follow up, and other.

B5. Priority question. CS, Document B, section B.2.3.1 and B.3.3.3. The company submission states that, "*due to a lack of data beyond Week 36 for placebo, the Week 0-36 all-cause discontinuation rate is used to estimate the annual discontinuation rate*" (p115). Additionally, for the baricitinib arm, Week 0-52 all cause discontinuation is used in the maintenance period.

a) **Please clarify why longer-term discontinuation data from the BRAVE extension studies are not available for placebo given that responders can remain on placebo in each study beyond 36 weeks (Figure 3 and Figure 4)?**

As discussed in Question A13, patients in the placebo arm of the BRAVE-AA trials were rescued to either baricitinib 2 mg or baricitinib 4 mg. Beyond 36 weeks, the placebo arms of the trials therefore contain selected placebo responders only; this is a biased subgroup and is no longer representative of the placebo arm of the blinded period of the trial up to 36 weeks. Furthermore, beyond 36 weeks, the placebo arm in the pooled population of BRAVE-AA1 and BRAVE-AA2 includes only [redacted], further limiting the reliability and generalisability of these data.

i) **If discontinuation data for up to Week 52 are available for placebo, please explore this in a scenario.**

As discussed above in subsection (a), the Company does not consider these data to be relevant to decision-making, so the Company have chosen not to conduct these analyses.

b) **Please clarify why all-cause discontinuation data for Week 36-52 (or beyond) was not explored for use in the Maintenance health state?**

i) **Please explore a scenario where all-cause discontinuation data for Week 36-52 (or beyond if available) is applied to the Maintenance health state for both the baricitinib and placebo arms (if available).**

Of the [redacted] patients who remained on baricitinib 4 mg at Week 36, [redacted] patients subsequently discontinued treatment, such that [redacted] patients remained on baricitinib 4 mg at Week 52.

Assuming that discontinuation occurs at a constant rate over 52 weeks, 42.5 patients are assumed to discontinue over a year, giving an annual discontinuation rate in the maintenance state of [redacted]. This rate is used in the scenario analysis shown in Table 24. As discussed above in subsection (a), no discontinuation data beyond 36 weeks for placebo are relevant so only the baricitinib rate is updated in this scenario analysis.

Table 24. Cost-effectiveness outcomes using baricitinib discontinuation in the maintenance state based on Week 36–52 discontinuation data (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£16,369

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Health-related quality of life

B6. Priority question. CS, Document B, section B.3.4.1. The references for the Adelphi study provided no detail on the estimation of the utility gain for patients achieving $SALT_{50}$ and $SALT_{75}$. Please provide the company's internal report for the Adelphi study.

In the absence of an internal report for the Adelphi DSP study, we have provided the study protocol, the questionnaire answered by clinicians and patients, and the relevant sections of the data file containing the results.

Patients in the Adelphi DSP were categorised as having very severe, severe, moderate, or mild AA, and utility values were derived for each group. The baseline utility values in the model are based on patients with "Severe + Very Severe" AA (n=184). Utility values from patients with moderate AA (n=275) and mild AA (n = 97) are also available in the Adelphi DSP and the difference between these values is used as a proxy to represent $SALT_{50}$ and $SALT_{75}$ response levels, respectively.

B7. Priority question. CS, Document B, section B.3.4.1. Please explain why the EQ-5D estimates from the Adelphi study are more reliable than the EQ-5D estimates from BRAVE-AA1 and BRAVE-AA2, given the main criticism for the trial estimates was a lack of content validity associated with the EQ-5D instrument.

The lack of content validity of the EQ-5D instrument in this indication affects the HRQoL data generated from both the Adelphi DSP and BRAVE-AA1 and BRAVE-AA2 trials, likely resulting in an underestimation of the HRQoL gain associated with a response to treatment. However, this lack of content validity was one of two main criticisms of the EQ-5D data generated in BRAVE-AA1 and BRAVE-AA2. The second key criticism of these data, which was overcome by using the Adelphi DSP, was the observed ceiling effect associated with the EQ-5D data generated from the BRAVE-AA trials, whereby a substantial proportion ([redacted]) of participants in the BRAVE-AA1 and BRAVE-AA2 trials reported "perfect health" on the EQ-5D instrument at baseline (i.e. a score of

11111), as shown in Table 25. These patients would have been unable to report an improvement in HRQoL due to a response based on SALT score, leading to an underestimation of the HRQoL gain associated with response to baricitinib treatment. This ceiling was not observed to as great an extent in the Adelphi DSP EQ-5D data, though it is still somewhat evident, as shown in Table 26; of note, a much greater proportion of patients in the Adelphi DSP report Level 2 or 3 for Pain/Discomfort and Anxiety/Depression compared with the baseline BRAVE-AA trial data. Therefore, while the Adelphi DSP data remains affected by the lack of content validity of the EQ-5D instrument, these data are likely to more accurately capture the HRQoL benefits associated with a response to treatment.

Table 25. Summary of EQ-5D domain scores at baseline in BRAVE-AA1 and BRAVE-AA2 (pooled FAS population)

Levels	Mobility, n (%)	Self-care, n (%)	Usual activities, n (%)	Pain/ Discomfort, n (%)	Anxiety/ Depression, n (%)	Participants in health state 11111 (%)
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	

Abbreviations: EQ-5D: European Quality of Life-5 Dimensions; FAS: full analysis set

Table 26. Summary of EQ-5D domain scores at baseline in the severe and very severe group in the Adelphi DSP

Levels	Mobility, n (%)	Self-care, n (%)	Usual activities, n (%)	Pain/ Discomfort, n (%)	Anxiety/ Depression, n (%)
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: EQ-5D: European Quality of Life-5 Dimensions

B8. Priority question. CS, Document B, section B.3.4.5. Please clarify whether utility values associated with achieving SALT≤20 were estimated in the Adelphi study. If so, please provide these data and explore this in a combined scenario linked to Question B1.

As discussed in Question B6, utility values are not available from the Adelphi DSP for patients achieving SALT≤20. Instead, patients in the Adelphi DSP were categorised as having very severe, severe, moderate, or mild AA, and utility values were derived for each group. Utility values for patients in the 'mild' AA group in the DSP were subsequently used as a proxy for patients achieving SALT≤20 or $SALT_{75}$.

B9. Priority question. CS, Document B, section B.2.6.3 and B.3.4.1. There is a lack of detail around the EQ-5D data collected in BRAVE-AA1 and BRAVE-AA2.

a) **Please provide details on how EQ-5D was measured in the trials (timepoints of measurement, number of responses at each time point, length of follow up, etc.) along with the mean EQ-5D values at each timepoint.**

During the double-blinded period of BRAVE-AA1 and BRAVE-AA2, EQ-5D-5L was measured at the following timepoints:

- Weeks 0, 12, 24, and 36

During the long-term extension phase of the BRAVE-AA1 and BRAVE-AA2 trials, EQ-5D-5L was measured at the following timepoints:

- Weeks 52, 64, and 76

Patients were followed up for 200 weeks. Mean values for EQ-5D VAS scores and the number of responses at each timepoint are presented below in Table 27.

Table 27. Mean values for EQ-5D VAS at each timepoint in all treatment arms (pooled population; primary censoring rule [mLOCF])

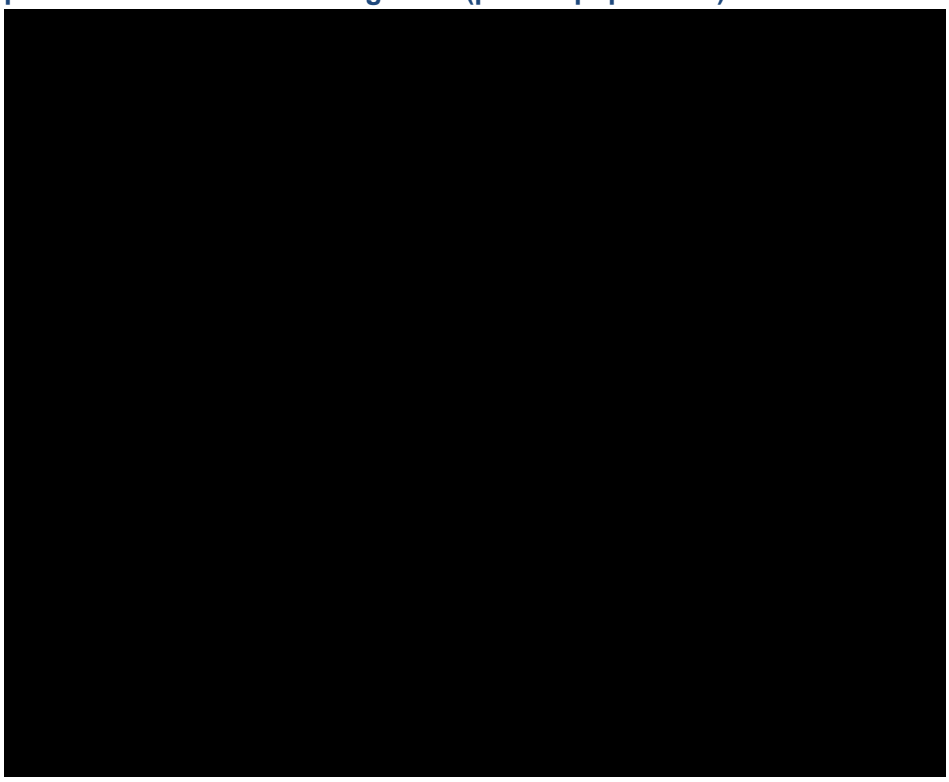
	Placebo	Baricitinib 2 mg	Baricitinib 4 mg
Week 0			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 12			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 24			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 36			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 52			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]

Abbreviations: EQ-5D: European Quality of Life-5 Dimensions; HADS: hospital anxiety and depression scale; mLOCF: modified last observation carried forward; LSM, least squares mean; SE: standard error.

b) **Please provide a scatter plot and the correlation, or provide a suitable regression analysis, between SALT score (continuous) and EQ-5D at baseline and at Week 36.**

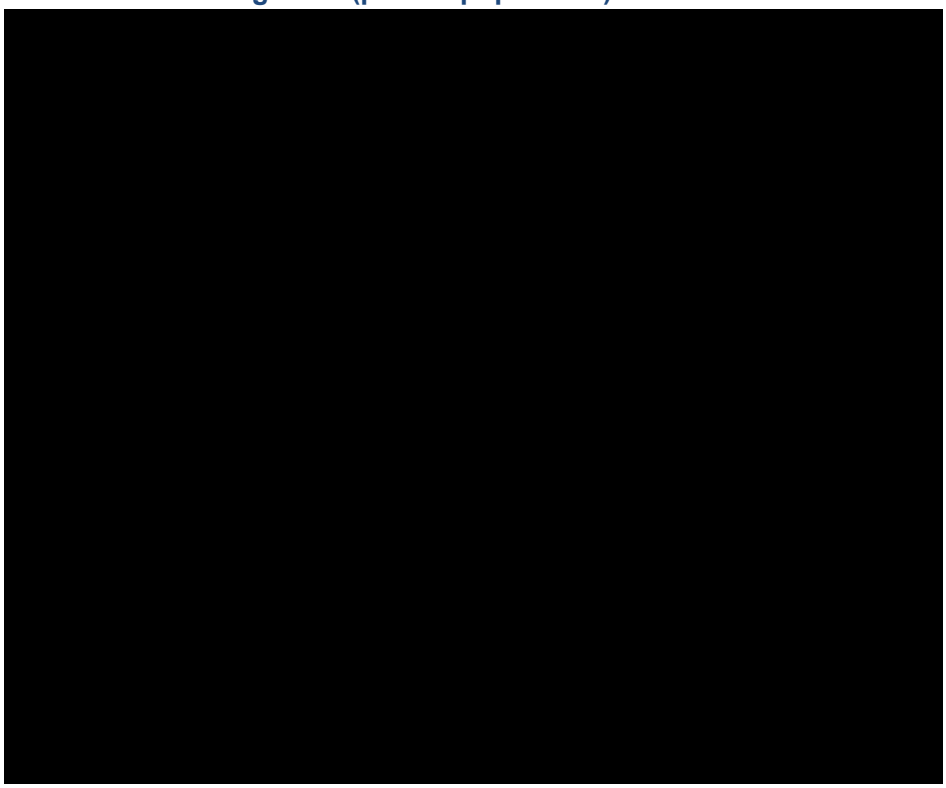
The correlation between SALT (continuous) scores and the EQ-5D-5L data (mapped into EQ-5D-3L using the cross-walk algorithm by Hernandez *et al*.[20]) in the pooled population at baseline and at Week 36 for the placebo and baricitinib 4 mg arms is shown in Figure 15 and Figure 16, respectively. Pearson's coefficient at baseline and Week 36 are 0.031 and −0.022, respectively, indicating no clear correlation between SALT scores and EQ-5D at either time point.

Figure 15. Correlation between SALT scores and EQ-5D-3L scores at baseline for both placebo and baricitinib 4 mg arms (pooled population)



Footnotes: Pearson r=0.031. One patient in the placebo arm of the BRAVE-AA2 trial had baseline SALT<40, but their age was not recorded so this patient was excluded from this analysis. EQ-5D-5L mapped into EQ-5D-3L using cross-walk algorithm by Hernandez *et al*.[20]
Abbreviations: SALT: severity of alopecia areata tool.

Figure 16. Correlation between SALT scores and EQ-5D-3L scores at Week 36 for placebo and baricitinib 4 mg arms (pooled population)



Footnotes: Pearson r=−0.022; EQ-5D-5L mapped into EQ-5D-3L using cross-walk algorithm by Hernandez *et al*.[20]
Abbreviations: SALT: severity of alopecia areata tool.

c) Please clarify whether regression modelling was performed to analyse the EQ-5D data. If so, please provide details of the models.

The following linear model was used to analyse the EQ-5D data:

cEQ5D = a0 + a1 EQ5Dbl + a2 SALTcat + a3 AGE

Abbreviations: cEQ5D = Change in EQ5D; EQ5Dbl = Baseline EQ5D; SALTcat = SALT improvement categories at Week 36 (<50%, ≥50% to <75%, ≥75%); AGE = Age in years

d) Please provide baseline and change from baseline values for SALT≤20 and SALT≤10 based on EQ-5D data from BRAVE-AA1 and BRAVE-AA2.

Baseline and change from baseline utility values for patients achieving a SALT≤10 and SALT≤20 response are presented in Table 28, derived from the EQ-5D data generated from the baricitinib 4 mg and placebo arms in the BRAVE-AA trials and the regression modelling approach noted above in subsection (c).

Table 28. Baseline and change from baseline utility values for patients achieving a SALT≤10 and SALT≤20 response at Week 36 derived from the placebo and baricitinib 4 mg BRAVE-AA trial EQ-5D data

	SALT response				
	Overall population [redacted]	SALT>20 [redacted]	SALT≤20 [redacted]	SALT>10 [redacted]	SALT≤10 [redacted]
Mean (SE) baseline score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LSM (SE) change from baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AA: alopecia areata; NA: not applicable; LSM, least squares mean; Skindex-16 AA: Skindex-16 Adapted for Alopecia areata; SD: standard deviation..

i) Please use these values in the scenario requested in Question B1.

The utility values presented in subsection (d) above are subject to the same issues highlighted in the CS and in Question B7 regarding the utility values derived from HRQoL data from the BRAVE-AA trials, including the observed ceiling effect and lack of content validity of the EQ-5D instrument in this indication. The Company have therefore chosen not to combine these utility values with the scenario request in Question B1.

B10. Priority question. CS, Document B, section B.2.6.3. The EAG's clinical experts suggested that any benefit in health-related quality of life (HRQoL) may lag behind any clinical response. Please provide absolute and change from baseline values in EQ-5D-5L, HADS-A and HADS-D at Week 52 for BRAVE-AA1, BRAVE-AA2 and the pooled population for the 4mg baricitinib arm.

Absolute and change from baseline values at Week 52 for EQ-5D VAS scores, HADS-A and HADS-D in the pooled population in the 4 mg baricitinib arm are presented in Table 29. These data demonstrate minimal difference in HRQoL data from the BRAVE-AA trials from baseline to Week 36 versus baseline to Week 52.

Table 29. Absolute and change from baseline values for EQ-5D VAS score, HADS-A and HADS-D for the 4 mg baricitinib arm (pooled population; primary censoring rule [mLOCF])

Week 52	EQ-5D [redacted]	HADS-A [redacted]	HADS-D [redacted]
Mean (SD) baseline score	[redacted]	[redacted]	[redacted]
Mean change from baseline (SD)	[redacted]	[redacted]	[redacted]
Absolute values (SD)	[redacted]	[redacted]	[redacted]

Abbreviations: EQ-5D: European Quality of Life-5 Dimensions; HADS: hospital anxiety and depression scale; mLOCF: modified last observation carried forward; LSM, least squares mean; SD: standard deviation.

a) **Please also provide these data at each timepoint up to Week 76 for the patients who are on 4mg baricitinib for the whole duration of the study.**

EQ-5D and HADS data are not yet available beyond Week 52.

B11. Priority question, CS, Document B, section B.2.6.3 and B.3.4.2. There is a lack of detail around the HADS data collected in BRAVE-AA1 and BRAVE-AA2.

a) **Please provide details on how HADS was measured in the trials (timepoints of measurement, number of responses at each time point, length of follow up, etc).**

During the double-blinded period of BRAVE-AA1 and BRAVE-AA2, HADS was measured at the following timepoints:

- Weeks 0, 12, 24, and 36

During the long-term extension phase of the BRAVE-AA1 and BRAVE-AA2 trials, HADS was measured at the following timepoints:

- Weeks 52 and 76

Patients were followed up for 200 weeks. Mean values for HADS-A and HADS-D scores and the number of responses at each timepoint are presented below in Table 30 and Table 31, respectively.

Table 30. Mean values for HADS-A at each timepoint in all treatment arms (pooled population; primary censoring rule [mLOCF])

	Placebo	Baricitinib 2 mg	Baricitinib 4 mg
Week 0			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 12			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 24			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 36			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 52			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]

Abbreviations: HADS: hospital anxiety and depression scale; mLOCF: modified last observation carried forward; SD: standard deviation.

Table 31. Mean values for HADS-D at each timepoint in all treatment arms (pooled population; primary censoring rule [mLOCF])

	Placebo	Baricitinib 2 mg	Baricitinib 4 mg
Week 0			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 12			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 24			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 36			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Week 52			
Number of patients	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]

Abbreviations: HADS: hospital anxiety and depression scale; mLOCF: modified last observation carried forward; SD: standard deviation.

b) Please clarify the mapping algorithm that was used to estimate EQ-5D-3L scores from the HADS data from BRAVE-AA1 and BRAVE-AA2.

The applied mapping algorithm that was used to estimate EQ-5D-3L score is based on HADS data (Total Score) and on age from the BRAVE-AA trials. The respective regression coefficients are reported in Table 48 (Page 166) in the publication by Brazier (2014).[21]

c) Please estimate EQ-5D baseline and change from baseline utility values based on the HADS data from BRAVE-AA1 and BRAVE-AA2 for patients with SALT≤20 and SALT≤10 and explore these data in the scenarios requested in Question B1.

Baseline and change from baseline utility values for patients achieving a SALT≤10 and SALT≤20 response are presented in Table 32, derived from the HADS data generated from the BRAVE-AA trials and the mapping approach noted above in subsection (b). It should be noted that these utility values are subject to the same issues highlighted in the CS and in Question B7 regarding the utility values derived from HRQoL data from the BRAVE-AA trials, including the high baseline HADS values and the resultant ceiling effect. The Company have therefore chosen not to combine these utility values with the scenario request in Question B1.

Table 32. Utility values for patients achieving a SALT≤10 and SALT≤20 response at Week 36 derived from BRAVE-AA trial HADS data

	SALT response				
	Overall population ([redacted])	SALT>20 ([redacted])	SALT≤20 ([redacted])	SALT>10 ([redacted])	SALT≤10 ([redacted])
Mean (SE) baseline score	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LSM (SE) change from baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: N numbers refer to baseline values.
Abbreviations: SALT: severity of alopecia tool; SE: standard error; SD: standard deviation

Resource use and costs

B12. Priority question. CS, Document B, section B.3.5.1. The references for the Adelphi study provided limited detail on the resource use estimation used in the model. As requested in question B6, please provide the company's internal report for the Adelphi study.

In the absence of an internal report for the Adelphi DSP study we have provided the study protocol, the questionnaire answered by clinicians and patients, and the relevant sections of the data file containing the results.

Resource use in the model is based on treatment patterns data from UK patients with AA (n=130), adjusted for patients treated with JAK-inhibitors off-label, and for treatments listed as 'Other'. These are split proportionally across all other BSC treatments included in the model.

B13. Priority question. CS, Document B, section B.3.5.4. The EAG's clinical experts advised that in UK clinical practice, patients who are not on an active treatment for AA (the company has considered this to be 'Watch and Wait') are unlikely to be monitored. As such, please provide a scenario where monitoring costs for 'Watch and Wait' patients in the induction and maintenance phases of the model are removed.

The cost-effectiveness outcomes of the scenario in which the monitoring costs for 'watch and wait' is removed is presented in Table 34, and instructions on running this scenario are provided in Table 33.

Table 33. Instructions to run "Removal of Monitoring Costs in Watch and Wait"

Sheet	Parameter	Cells	Justification
Treatment_Costs	Placebo monitoring costs	N48:O49	Removal of routine monitoring costs for 'Watch and Wait' patients

Table 34. Removal of monitoring costs from Watch and Wait (deterministic)

Technologies	Total discounted costs (£)	Total discounted QALYs	Incremental discounted costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib 4 mg	[redacted]	[redacted]	[redacted]	[redacted]	£20,887

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year.

B14. Priority question. CS, Document B, section B.3.5.1.2. Please provide a rationale for assuming no blood tests are conducted in the induction phase for patients on baricitinib. Please provide a scenario where 3-monthly blood tests are included for baricitinib in the induction phase (in line with the assumption made for the maintenance health state).

The Company has updated the base case to include 3-monthly blood tests for baricitinib during the Induction phase; further details of this change are presented in the Appendix.

B15. Priority question. CS, Document B, section B.3.5.1.1. Please clarify the following assumptions around BSC and make amendments where necessary:

- **Azathioprine: the dose presented by the company is similar to that for Crohn's disease but higher than for autoimmune diseases. Please clarify the basis for assuming a dose of 2 mg/kg QD for 1 year.**

Since Azathioprine is used off-label for AA, there is no stated dose specifically for this condition. However, the BNF website states that for autoimmune conditions a dose of 1–3 mg/kg daily should be initiated and adjusted according to response.[22] We assumed the mid-point of this range for the economic model.

- **Prednisolone: according to the British National Formulary (BNF), 5mg tablets are cheaper than 2.5mg tablets. Please update the model using the cheaper price.**

Based on the NHS Drug Tariff, we agree that the cost of 5 mg tablets is lower than the cost of 2.5 mg tablets. This has been implemented in the model (see Appendix) and incorporated into the updated base case.

- **Mometasone ointment: the EAG's clinical expert considered that a dose of 32mg once daily is quite high. The SmPC recommends a thin film (1g) to applied to the affected area. Additionally, the cost of the scalp lotion may be more applicable (£4.36 per 30ml according to the BNF). Please provide a rationale for choosing 32mg once daily for the model. Please**

provide a scenario exploring a lower dose and associated costs of scalp lotion.

The Company agrees that scalp lotion formulation of mometasone ointment is more appropriate. We have updated the base case to reflect this cost (see Appendix), based on the NHS Drug Tariff. For mometasone scalp lotion dosage, we have applied a dosage of 2ml per day based on clinical feedback for facial vitiligo, and assuming that the area the lotion is applied to would be similar in both conditions.

- **Minoxidil 5% foam (topical): the dose for women and men has been swapped. For women, the dose is 1g once daily for 24 weeks and for men, it is 1g twice daily for 16 weeks. Please update the model.**

We agree that this error was present in the original model and has now been corrected (see Appendix).

- **Mycophenolate mofetil: the company submission states 50mg tablets, but 500 mg tablets are only available according to the BNF. Please clarify whether this is an error.**

This was a typographical error within the CS; the cost of 500 mg tablets has been applied correctly within the model (see Appendix).

B16. Priority question. The EAG's clinical experts advised that there is significant variation in care for patients with AA and it is likely that if patients do not respond to treatment, they will not engage with further treatment or will not be followed up (effectively patients are discharged). Additionally, the EAG's clinical experts considered that if a patient's condition does not adequately respond to a Janus Kinase (JAK) inhibitor, they would not be given any further treatment. As such, please provide a scenario where disease management (not including wig use and orthotics) and drug acquisition and monitoring costs in the BSC health state are removed.

The EAG's clinical expert feedback suggests that BSC costs should be removed from the baricitinib arm of the model as it is not anticipated that patients who do not adequately respond to treatment with a JAK inhibitor will receive any further treatment. Results of this scenario are presented in Table 35.

Table 35. Removal of BSC costs following treatment with baricitinib (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life-years; QALY, quality-adjusted life year.

Although it is potentially true that differing approaches to BSC costs may be taken following treatment with baricitinib versus a watch and wait approach, the Company considers that the extent of this difference is currently uncertain. Therefore, functionality has been added to the model to allow exploration of the level of BSC costs following treatment with baricitinib or with watch and wait. Further input from clinical experts is likely needed to define the difference in subsequent treatment approach between baricitinib and watch and wait.

B17. Priority question. CS, Document B, section B.3.5.4. The EAG's clinical experts advised that in the NHS, provision of psychological support for patients with AA is limited.

a) **Please provide evidence to support the company base case assumptions for the provision of psychological support to patients with AA in the NHS.**

b) **Please provide a scenario removing the costs of psychological support in both the induction and BSC health states.**

Clinical expert feedback sought by the Company ahead of the submission suggested that some psychological support would be available to patient suffering the most severe psychological impact of AA. However, it was noted that this would be limited only to the patients that most need it based on capacity and that many more patients with AA could potentially benefit.

It is assumed based on clinical input that 5% of patient would see a psychiatrist through the NHS and 10% of patients would see a psychologist, often via self-referral through the Improving Access to Psychological Therapies (IAPT) scheme. Therefore, non-pharmacological treatments were adjusted based on the 15% of patients who would be able to access them.

Furthermore, proportions of patients receiving pharmacological treatments were also estimated via clinical input. It was assumed that pharmacological treatments for depression/anxiety in patients with AA would be prescribed by the patient's GP, so these proportions were not adjusted for those patients receiving a referral for secondary care.

Nonetheless, for transparency, the cost-effectiveness outcome of the scenario in which psychological support is removed in both the induction and BSC health states is presented in Table 36, and instructions on running this scenario are provided in Table 36.

Table 36. Instructions to run "Removal of Psychological Care Costs"

Sheet	Parameter	Cells	Justification
Treatment_Costs	Psychological Care Costs	J79:S97	Removal of costs associated with psychological care

Table 37. Cost-effectiveness outcomes removing psychological care costs (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			

Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£19,243

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B18. CS, Document B, section B.3.5.1.2. The EAG's clinical experts advised that wigs are predominantly used by female patients only. As such, please provide a scenario where the percentage of wig use in all health states is reflective of the percentage of females at baseline in the model (60.7%).

The cost-effectiveness outcomes of the scenario in which the percentage of wig use in all health states is aligned with the percentage of females at baseline in the model (60.7%) is presented in Table 39, and instructions on running this scenario are provided in Table 38.

Table 38. Instructions to run "Female patients only incurring wig costs"

Sheet	Parameter	Cells	Justification
Treatment_Costs	Percentage of wig use and orthotics in all health states	C70:E71 G70:G71	Reducing to the percentage of females at baseline in the model (60.67%)

Table 39: Cost-effectiveness outcomes using female-only incurring wig costs (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,732

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B19. CS, Document B, section B.3.5.1.2. The EAG's clinical experts advised that in England, up to 2 wigs are provided per year on the NHS. In the induction phase, it is assumed that 2 wigs would be provided in 36 weeks and that 2 wigs per year are provided in the BSC health state, which is inconsistent. Please provide a scenario where only one wig is provided in the induction phase.

The cost-effectiveness outcomes of the scenario in which only one wig is provided in the induction phase is presented in Table 41, and instructions on running this scenario are provided in Table 40.

Table 40. Instructions to run "Only one wig provided during induction phase"

Sheet	Parameter	Cells	Justification
Treatment_Costs	Resource use of wig use in induction	J70, L70, N70	Reducing to only 1 wig in the induction phase

Table 41: Cost-effectiveness outcomes using the provision of one wig (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,068

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B20. As BSC costs make up the bulk of total costs for both arms of the model, please explore the following shorter time horizons as scenarios:

- **5 years**
- **10 years**
- **20 years**

The cost-effectiveness outcomes of the scenarios in which a time horizon of 5 years, 10 years and 20 years are presented in Table 42, Table 43, and Table 44, respectively.

Table 42: Cost-effectiveness outcomes using a time horizon of 5 years (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£34,902

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 43: Cost-effectiveness outcomes using a time horizon of 10 years (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£23,827

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 44: Cost-effectiveness outcomes using a time horizon of 20 years (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£19,256

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B21. Please provide a scenario where costs of adverse events are included in the model.

The cost-effectiveness outcomes of the scenario in which adverse events are included in the model is presented in Table 46, and instructions on running this scenario are provided in Table 45. The Company have chosen not to include adverse events in the base case since the

inclusion of costs associated with AEs do not have a significant impact on the ICER given that AEs observed in the BRAVE-AA trials were mostly mild, with [redacted] of patients experiencing serious adverse events (SAEs) across trial arms in BRAVE-AA1 and BRAVE-AA2.

Table 45. Instructions to run "Inclusion of Adverse Events"

Sheet	Parameter	Cells	Justification
Treatment costs	Adverse event unit costs	C125:C128	The following unit costs were added: URTI £39.0, Nasopharyngitis £39.0, Headache £206.34, Acne £171.53
Main	Dropdown menu for inclusion of adverse events	D28	Select 'Yes' to allow the inclusion of AE costs in the model calculations

Table 46: Cost-effectiveness outcomes under inclusion of adverse events (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,348

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

B22. The SmPC states that "*it is recommended to continue treatment for at least several months, in order to avoid relapse*". Please comment on the likely maximum time on treatment with baricitinib. Currently in the model, treatment duration for patients who achieve $SALT_{50}$ and $SALT_{75}$ is approximately 15 months and 37 months, respectively.

The duration of treatment with baricitinib in practice is expected to be an individual decision made between the patient and their treating clinician, guided by a number of factors including patient response, the guidance in the SmPC and emerging clinical experience. Given this, the definition of a maximum treatment duration for any given individual is considered to be of less relevance than the population-level average duration. The Company consider the approach taken in the model base case to be the best reflection of the available population-level evidence at this time, and a suitable basis for decision-making.

Section C: Textual clarification and additional points

C1. Priority question. CS, Document B, section B.1.2. Please confirm the PAS discount of [redacted] is correct. Based on the list price of £805.56 and applying the PAS discount, the EAG calculates the PAS price of baricitinib to be [redacted] not [redacted] as stated in the CS and applied in the model.

The legal construction of the PAS price of baricitinib is as a fixed pack price of [redacted] The fixed price is not dependent on a percentage discount. The [redacted] discount cited was illustrative only

and subject to rounding error; and the fixed unit price specified should continue to be used in the appraisal.

C2. Please clarify why a standard error (SE) of 10% has been used in the PSA when SE from the data are unavailable? Please consider exploring the use of a SE of 20% when SE from the data are unavailable.

An assumption of the SE being 10% of the mean in situations where the SE was unavailable is common practice in NICE Technology Appraisals, as well as in analyses undertaken by NICE in preparing Clinical Guidelines. Given that upper and lower 95% confidence limits may be estimated as 95% CL = mean ± (SE × 1.96), the commonly used assumption of SE being 10% of the mean tests an approximate uncertainty of ±20%, whereas the suggestion to use a SE of 20% of the mean would represent an extreme assumption of uncertainty being ±39%.

C3. Priority question. CS, Document B, section B.3.11.3. The EAG was not able to replicate the company's scenario analyses. Please clarify whether the results are deterministic or probabilistic. For reference, the EAG's deterministic results for each scenario are presented below.

- **Scenario 1 (Table 70) - the EAG produces an ICER of £36,064**
- **Scenario 2 (Table 71) - the EAG produces an ICER of £24,502**
- **Scenario 3 (Table 72) - the EAG produces an ICER of £26,681**
- **Scenario 4 (Table 73) - the EAG produces an ICER of £274,051**
- **Scenario 5 (Table 74) - the EAG produces an ICER of £87,163**
- **Scenario 6 (Table 75) - the EAG produces an ICER of £10,762**

Scenario analyses 1–6 were performed using probabilistic sensitivity analyses, as per the methods outlined by NICE.

References

1. Eli Lilly. Clinical Study Report: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA1). 2022.
2. Eli Lilly. Clinical Study Report: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA2). 2022.
3. Eli Lilly. Clinical Study Protocol: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA1). June 2021.
4. Eli Lilly. Clinical Study Protocol: AA Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata(BRAVE-AA2). June 2021.
5. Lamb RC, Young D, Holmes S. Retrospective review of diphencyprone in the treatment of alopecia areata. Clin Exp Dermatol 2016;41:352-8.
6. Durdu M, Özcan D, Baba M, et al. Efficacy and safety of diphenylcyclopropenone alone or in combination with anthralin in the treatment of chronic extensive alopecia areata: a retrospective case series. J Am Acad Dermatol 2015;72:640-50.
7. Messenger AG, McKillop J, Farrant P, et al. British Association of Dermatologists' guidelines for the management of alopecia areata 2012. British Journal of Dermatology 2012;166:916-926.
8. Lai V, Chen G, Gin D, et al. Cyclosporine for moderate to severe alopecia areata: Interim analysis of a double-blind, randomised, placebocontrolled clinical trial of efficacy and safety. Journal of the Dermatology Nurses' Association 2020;12.
9. Lai VWY, Chen G, Gin D, et al. Cyclosporine for moderate-to-severe alopecia areata: A double-blind, randomized, placebo-controlled clinical trial of efficacy and safety. Journal of the American Academy of Dermatology 2019;81:694-701.
10. King B, Guttman-Yassky E, Peeva E, et al. A phase 2a randomized, placebo-controlled study to evaluate the efficacy and safety of the oral Janus kinase inhibitors ritlecitinib and brepocitinib in alopecia areata: 24-week results. J Am Acad Dermatol 2021;85:379-387.
11. King B, Ko J, Forman S, et al. Efficacy and safety of the oral Janus kinase inhibitor baricitinib in the treatment of adults with alopecia areata: Phase 2 results from a randomized controlled study. J Am Acad Dermatol 2021;85:847-853.
12. King B, Ohyama M, Kwon O, et al. Two Phase 3 Trials of Baricitinib for Alopecia Areata. N Engl J Med 2022.
13. Asilian A, Fatemi F, Ganjei Z, et al. Oral pulse betamethasone, methotrexate, and combination therapy to treat severe alopecia areata: A randomized, double-blind, placebo-controlled, clinical trial. Iranian Journal of Pharmaceutical Research 2021;20:267-273.
14. Zerbinati N, Esposito C, D'Este E, et al. Topical Immunotherapy of Alopecia Areata: A Large Retrospective Study. Dermatol Ther (Heidelb) 2018;8:101-110.
15. National Institute for Health and Care Excellence (NICE). Baricitinib for moderate to severe atopic dermatitis: Technology appraisal guidance [TA681]. Available at: https://www.nice.org.uk/guidance/ta681. [Date accessed: June 2020].

16. 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: February 2022].
17. National Institute for Health and Care Excellence (NICE). Secukinumab for treating moderate to severe plaque psoriasis: Technology appraisal guidance [TA350]. Available at: https://www.nice.org.uk/guidance/ta350 [Date accessed: September 2022].
18. EMA. Olumiant (baricitinib) Assessment Report. Available at: https://www.ema.europa.eu/en/documents/variation-report/olumiant-h-c-4085-ii-0029-g-epar-assessment-report-variation_en.pdf. Last accessed: August 2022.
19. National Institute for Health and Care Excellence (NICE). Dupilumab for treating moderate to severe atopic dermatitis [TA534]. Available at: https://www.nice.org.uk/guidance/ta534. [Date accessed: May 2022].
20. Hernández Alava M, Pudney S, Wailoo A. Methods for mapping between the EQ-5D-5L and the 3L for technology appraisal: Report by the Decision Support Unit. 2017. Decision Support Unit, University of Sheffield.
21. Brazier J, Connell J, Papaioannou D, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess 2014;18.
22. British National Formulary (BNF). Azathioprine. Available at: https://bnf.nice.org.uk/drugs/azathioprine/ [Last accessed: 8th September 2022].

Appendix: Revised Base Case

The Company base case has been updated based on the final analysis of the Adelphi DSP (only the interim analysis was available at the time of submission), updated NHS reference costs, and errors/updates suggested by the EAG in Questions B14 and B15. A change log and a summary of the updated data to inform the revised Company base case are presented in Table 47 and Table 48, respectively. The cost-effectiveness results of the revised company base case are presented in Table 49.

Abbreviations: BSC: best supportive care; DSP: disease specific program; NHS: national health service.

Table 47. Log of changes made to the Company base case

Sheet	Parameter	Cells	Justification
Default_pop tabs	Utilities	J56:K72	Revision to reflect final analysis set from Adelphi DSP
Default_pop tabs	BSC proportions of pts	D185:D196, H95, H96	Revision to reflect final analysis set from Adelphi DSP
Default_pop tabs	NHS reference costs	J95, J105: J112, J118, D141:D151	Revision in unit costs to reflect the most up-to-date source of NHS reference costs (20/21)
Default_pop tabs	Blood tests in induction	K108, M108	Revision to reflect suggested change in question B14
Default_pop tabs	Prednisolone cost	D174:E174	Revision to reflect suggested changes in question B15
Default_pop tabs	Mometasone dosage	G176	Revision to reflect suggested changes in question B15
Default_pop tabs	Mometasone (topicals)	D175, E175, G175	Revision to reflect suggested changes in question B15

Abbreviations: BSC: best supportive care; DSP: disease specific program; NHS: national health service.

Table 48. Updated data to inform the revised Company base case

Input Category	Parameter	Original Value	Updated Value	Source
DSP Utilities	Baseline Utility	[redacted]	[redacted]	Adelphi DSP (Final Analysis Set)
	$SALT_{50}$ CFB	[redacted]	[redacted]	
	$SALT_{75}$ CFB	[redacted]	[redacted]	
DSP Resource Use	Ciclosporin	12.39%	13.72%	Adelphi DSP (Final Analysis Set)
	Methotrexate	14.29%	12.86%	
	Azathioprine	2.86%	2.57%	
	Intralesional steroids	9.53%	9.43%	
	DPCP treatment	20.96%	21.63%	
	Prednisolone	17.15%	17.15%	
	Topicals (TCS)	24.77%	24.77%	
	Minoxidil (topical) foam	5.72%	5.72%	

	Minoxidil tablets	0.00%	0.00%	
	Mycophenolate Mofetil	2.86%	2.86%	
	Anthralin 0.1% cream	5.72%	5.72%	
	Patients not currently on treatment	13.00%	12.00%	
NHS Reference Costs 20–21	Dermatology Outpatient	£124.79	£171.53	Weighted average of WF01A–D and WF02A–D - Dermatology
	Thyroid function	£2.55	£3.63	DAPS05 - Haematology
	Vitamin D	£2.55	£3.63	DAPS05 - Haematology
	Ferritin	£2.55	£3.63	DAPS05 - Haematology
	Full blood count	£2.55	£3.63	DAPS05 - Haematology
	Liver function	£1.20	£1.85	DAPS04 - Clinical biochemistry
	Renal function	£1.20	£1.85	DAPS04 - Clinical biochemistry
	Tuberculosis	£8.15	£10.18	DAPS07 - Microbiology
	Lipids	£2.55	£3.63	DAPS05 - Haematology
	Orthotics	£132.00	£220.46	Service Code 658 - Total Outpatient Attendances
	Minor Skin Procedures	£156.56	£250.70	Outpatient Procedures – Service Code 330
Relating to B14	Blood tests in induction	0	3	To align with assumption in maintenance state
Relating to B15	Prednisolone cost	£3.91	£0.79	NHS Drug Tarriff
	Mometasone dosage	32mg	2ml	Assumption
	Mometasone (topicals)	£7.33	£4.36	NHS Drug Tariff

Abbreviations: BSC: best supportive care; DSP: disease specific program; NHS: national health service, SALT: severity of alopecia tool.

Table 49. Revised Company base case cost-effectiveness results (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			

Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,942

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Probabilistic sensitivity analysis (PSA)

A probabilistic sensitivity analysis (PSA) was run with 1,000 Monte Carlo simulations in order to assess the uncertainty associated with model input parameters. Use of 1,000 iterations was deemed appropriate based on the results of an ICER convergence test, shown in Figure 17.

Figure 17. Convergence plot for NMB

Abbreviations: CI: confidence interval; NMB, net monetary benefit; PSA: probabilistic sensitivity analysis; QALY: quality-adjusted life year.

A visual representation of the PSA results comparing baricitinib 4mg and placebo is provided in the cost-effectiveness plane (see Figure 18 below). Each dot represents one Monte Carlo simulation where the input parameters are sampled from the distributions in a total of 1,000 loops. The results of the cost-effectiveness plane show moderate uncertainty with regards to the extent of the additional costs for baricitinib 4 mg compared with 'watch and wait', but little uncertainly with regards to the existence of these additional costs, as most dots fall on the North quadrants of the plane. On the effectiveness side, PSA results show moderate uncertainty with regards to the extent of additional benefits.

Figure 18. Cost-effectiveness plane for baricitinib 4 mg compared with 'watch and wait'

Abbreviations: QALY, quality-adjusted life year.

The cost-effectiveness acceptability curve shows a [redacted]% probability of baricitinib 4 mg being cost-effective compared with 'watch and wait' at a cost-effectiveness threshold of £30,000/QALY (Figure 19), with the probability increasing with higher thresholds.

Figure 19. Cost-effectiveness acceptability curve for baricitinib 4mg compared with placebo

Footnotes: "placebo" in the figure represents "watch and wait"

Deterministic sensitivity analysis (DSA)

A deterministic one-way sensitivity analysis (OWSA) has been performed and the ten most important drivers of the model have been plotted in a tornado diagram (Figure 20). The three most influential parameters in the model were the frequency of DPCP treatment in the BSC state, drug monitoring resource use of DCPC treatment in the BSC state, and the $SALT_{50}$ response for baricitinib 4 mg at 36 weeks.

Figure 20. Tornado diagram for baricitinib 4 mg compared with 'watch and wait' at PAS price

Footnotes: "placebo" in the figure represents "watch and wait"
Abbreviations: BSC: best supportive care; DPCP: diphenylcyclopropenone; HSUV: health state utility value; SALT: severity of alopecia tool.

Scenario analysis

A number of scenario analyses were explored in which model assumptions or parameters were altered; these are presented in Table 50–Table 55

Table 50. Cost-effectiveness results Scenario 1: Starting population with SALT 50–94 (severe population)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£25,154

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 51. Cost-effectiveness results Scenario 2: Starting population with SALT 95–100 (very severe population)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£12,685

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 52. Cost-effectiveness results Scenario 3: Response based on $SALT_{75}$

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£16,490

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 53. Cost-effectiveness results Scenario 4: Utilities based on EQ-5D data from the BRAVE-AA trials

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£174,446

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 54. Cost-effectiveness results Scenario 5: Utilities based on HADS data from the BRAVE-AA trials

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£55,483

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 55. Cost-effectiveness results Scenario 6: Proportion of patients on BSC drugs based on clinical expert opinion

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]			
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Follow Up on Deterministic Results Corresponding to Probabilistic Scenarios Provided in Clarification Questions Response

September 2022

File name	Version	Contains confidential information	Date
ID3979_Baricitinib in AA_Clarification Questions [ACIC]_16September2022	FINAL	Yes	28th September 2022

Section B: Clarification on cost-effectiveness data

Treatment response

B1. Priority question. The EAG's clinical experts advised that an absolute reduction in SALT scores is a more clinically meaningful estimate of response than a relative improvement from baseline SALT scores. Please provide a scenario where treatment response in the model is based on the primary endpoint of achieving SALT≤20 from the BRAVE-AA1 and BRAVE-AA2 trials. Please combine this scenario with the utility values requested in question B8.

Table 1: Cost-effectiveness outcomes using SALT≤20 (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,312

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 2: Cost-effectiveness outcomes using SALT≤20 (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,071

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

a) **Please run an alternative scenario exploring the outcome of SALT≤10 from the BRAVE-AA1 and BRAVE-AA2 trials. Please combine this scenario with the utility values requested in question B8.**

Table 3: Cost-effectiveness outcomes using SALT≤10 (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£21,254

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 4: Cost-effectiveness outcomes using SALT≤10 (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£20,782

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

b) Please run the SALT≤20 scenario for the severe and very severe subgroups.

Table 5: Cost-effectiveness outcomes using SALT≤20 in the severe population (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,598

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 6: Cost-effectiveness outcomes using SALT≤20 in the severe population (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,773

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 7: Cost-effectiveness outcomes using SALT≤20 in the very severe population (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,196

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 8: Cost-effectiveness outcomes using SALT≤20 in the very severe population (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£16,929

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Appendix: Revised Base Case

Table 9. Revised Company base case cost-effectiveness results (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£17,942

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

Table 10. Revised Company base case cost-effectiveness results (deterministic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
Watch and Wait	[redacted]	[redacted]	[redacted]	[redacted]	
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,072

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year; SALT: severity of alopecia tool.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Response to Additional Queries Received 7th October 2022

October 2022

File name	Version	Contains confidential information	Date
ID3979_Baricitinib in AA_Additional Responses_11October2022	FINAL	No	11th October 2022

Query 1: The technical team would like to check the setting in which baricitinib would be commissioned in NHS practice, because this setting helps inform the most appropriate source of price to use. Do you anticipate that baricitinib (and its comparators, and any subsequent and concomitant treatments) would be commissioned primarily in primary care, secondary care, or a mix of both? Please note that the methods guide (section 4.4.7) states "For medicines that are mainly prescribed in primary care, base prices on the drugs tariff". If, however the drugs are prescribed primarily in secondary care, the drug tariff price would likely not be the most appropriate price to use.

Baricitinib is anticipated to be commissioned in secondary care only; this assumption is based on the Company's understanding of the treatment pathway for patients with AA, in which patients with mild to moderate AA are treated in primary care and patients with severe AA are treated in secondary care, following referral to a specialist dermatologist. Given the population of relevance to this submission is adults with severe AA, secondary care is the most relevant setting for the dispensing of baricitinib.

For BSC treatments, the Company has assumed that the majority would be commissioned in primary care, as it is anticipated that patients would be discharged from secondary care following non-response to baricitinib or watch-and-wait. Costs for the majority of BSC treatments were therefore sourced from the NHS Drug Tariff. The only exceptions to this assumption were DPCP and intralesional corticosteroids, which require specialist administration and would therefore be anticipated to delivered in secondary care. However, neither of these treatments could be sourced from eMIT. As such, the cost of a 5g vial of DPCP was sourced from Fisher Scientific (available here), and the cost of intralesional corticosteroids (triamcinolone acetonide) was sourced from the Drug Tariff, as these were considered the most reliable alternative source of costs for these treatments.

Query 2: The EAG is trying to work out how the company calculated the 12-week drug-acquisition cost presented in Table 63 of the company submission. In the model, the figures are hardcoded. The EAG has tried to calculate the figures, but they can only get a figure close to the company's figure, and not exact. Please can you send over the methodology for this?

The 12-week drug acquisition costs presented within Table 63 of the Company Submission (CS) were not used directly in the model. Instead, the total costs of pharmacological treatment for the management of depression were used as inputs (found in 'Treatment Costs!I92:I97' in the model), which were taken directly from the NICE guideline in development for depression in adults (evidence review B). This approach was taken to avoid replicating the methodology and calculations used by NICE, which are also presented in the NICE guideline in development noted above.

However, it should be noted that since the selection of the model inputs and the development of the model, the NICE guideline in development for depression in adults has been updated and replaced with NICE guideline NG222 (available here). The costs of pharmacological treatment for the management of depression presented in Table 63 of the CS therefore no longer exactly match those presented in Table 83 within the guideline in development for depression in adults (available here) and in Table 86 of guideline NG222. However, given these updated costs are very similar to the previous costs used in the cost-effectiveness model, the Company anticipate this change to have limited impact on the cost-effectiveness results. Nevertheless, if it is

preferred, the pharmacological treatment costs in the model can be updated to align with those listed in the updated guideline NG222.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Patient Organisation Submission

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.

About you

1.Your name	**Mrs Lynn Wilks**
2. Name of organisation	Alopecia UK
3. Job title or position	Trustee and Volunteer
4a. Brief description of the organisation (including who funds it). How many members does it have?	Alopecia UK – Charity Number 1111304 Alopecia UK is the national alopecia charity, covering all 4 nations of the UK. The organisational aims are; • To support people affected by alopecia, we will provide impartial information, advice, and support to help people feel less isolated. • To raise awareness to the general public and healthcare professionals about alopecia and its psychological impact. • To provide hope and confidence to people with alopecia by funding research into its causes, with the aim of finding treatments, and ultimately, a cure. We are not a membership organisation, but our community includes over 11,000 people who engage with us for information and support. The majority of our income comes from individual funding from the people affected by alopecia in our community. We have recently received a grant of £9,250 from The National Lottery Community Fund, and as below we have received a grant from Pfizer to lead some independent research.

4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	No funding has been received from Eli Lilly. An independent research grant from Pfizer Inc. was won by Alopecia UK in 2021. Value £55,000. The scope of that piece of research is a survey to explore the social and economic impact of alopecia areata (including totalis & universalis). Research is being carried out in collaboration with the University of West of England.
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?	• Open dialogue from social media private support groups • Gathered from our private group face-to-face meetings & events • 1:1 telephone support calls and emails • Facilitate PPI (Public and Patient Involvement) meetings for alopecia related research • Own patient research questionnaires – findings published

Living with the condition

See next pages

6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	• Some real, and typical comments from our community's private Facebook group "I hate myself and this", "I can't cope anymore", "I am deciding what treatment to try next but finding myself full of shame and quite depressed, more than I expected.", "They tell me to avoid stress but I can't turn off my life, no matter how much I'd like to. This is the worst it's ever been", "I'm heartbroken, I've been given a scalp ointment which I've had before and has little chance of working", "My kids really want to go swimming, but I haven't been since losing all my hair. I don't know if I'm brave enough to go", "I started with Alopecia three months ago, with a tiny patch of hair loss, now I have lost about 75% of my hair. This really affected my mental health and had to take time away from work and family", "This morning I lost a clump of long hair that came off in my hands in the shower, yesterday one eyebrow fell out in a day, I am scared of washing.". • People with alopecia describe feelings of shock, trauma, and disrupted identity (Davey L et al, 2019). • Leads to depression, anxiety, isolation, and even suicidal thoughts. • Alopecia UK research 2017 – clinically significant levels of anxiety in 35.5% and depression in 29%. • 25% of people had been told by healthcare professionals it was 'just a cosmetic issue' – which fails to recognise the psychosocial impacts (Johnson A, Montgomery K, 2017). • Psychosocial impacts include not wanting to go out and mix in social settings (66.3% of AUK respondents would not go out without wearing a wig); this leads to absenteeism from work/college; feeling of visible difference and stigma leads to a person not being 'present' in a role and hence possibly being passed aside for promotion. Children and young people report being bullied at school for "not being normal". Adults feel they are less likely to succeed, as they 'look different'. Reinforcing anxiety, depression, and social isolation. • People feel 'hopeless' as alopecia areata is still poorly understood with no cure and no real effective treatments. The few treatments are general and not licensed for alopecia and have limited access on the NHS. • In our studies over 25% of people voiced that having hair loss had negatively affected their close, intimate relationships. • For men with alopecia areata there is social pressure that they accept their visible difference and 'put on a brave face', as many men suffer from androgenetic alopecia (baldness). We know they suffer the same feelings of anxiety, depression, and psychosocial impact. • Alopecia UK understands that approx. 40% of people with alopecia areata have other autoimmune conditions – such as lupus, thyroid conditions, and psoriasis. Hence these people are having to deal with associated co-morbidities.

	• People with alopecia totalis and universalis can struggle with temperature regulation and report being cold all the time (as no scalp, face, or body hair). With no eyelashes you often suffer watering or dust in the eyes. With no nasal hair you can suffer from an embarrassing runny nose where nasal secretions suddenly drip, as no hair to trap mucus. • The speed of hair loss differs widely, some people can lose their hair in days, and for others it can be far longer. The lack of predictability makes it difficult for people to come to terms with their visible difference, and people report feeling a loss of control and their identity. • Our community tells us they spend a significant amount of money on unfounded “miracle cures”, we know they are targeted by unscrupulous sales techniques aimed at vulnerable people. • In the early stages of alopecia people often experiment with legitimately prescribed treatments, seeing private consultants and trichologists in the hope that something will work. Some of those treatments are extremely uncomfortable, and contact immunotherapy is described as especially painful in our groups. • We know that many people will spend a significant amount of their disposable income on products (e.g., microblading, wigs, false eyelashes) to adjust their visible difference to feel more socially normal so that they can improve their quality of life. Many people tell us about the costs of paying for products and services related to hair loss which can create further challenges. Alopecia UK is currently leading some independent research on this topic.

Current treatment of the condition in the NHS

7. What do patients or carers think of current treatments and care available on the NHS?	• Often poor empathy and understanding from primary care. Patients are told to 'wait and see' if more hair sheds or it grows back. Very few treatments are offered from primary care – only topical steroids. • 1 in 3-4 people are referred to dermatology but people are frustrated that referral times for alopecia are often +1 year. • Treatments offered by dermatology are limited and vary depending on whether you are referred to a tertiary centre where a dermatologist has an interest in alopecia or standard secondary care dermatology. • Patients accept there is no cure but are frustrated and despair that limited treatments are available with limited success in terms of a) the number of patients who respond and b) % hair regrowth. • People are distressed that for most treatments, hair will re-shed when the treatment is stopped. • Patients feel marginalised, alopecia appears to have fewer clinical and patient care guidelines than other skin conditions.
8. Is there an unmet need for patients with this condition?	Yes – absolutely! There is currently no on-label product available for alopecia. This is the first much-needed treatment, and it will change lives. Enabling hair regrowth addresses the debilitating psychosocial impacts of hair loss and improves people's quality of life.

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	It works! From the clinical trials that have been made public, it is exciting to see the percentage of people who seem to respond to the treatment and the percentage of hair regrowth that is generated. It gives people the hope that will then be able to live a 'normal' life and the ability to participate and contribute to society.

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	Overall no disadvantages – it provides hope, where currently there is none. People have viewed the side effects and feel the benefits outweigh any side effect risks.

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.	Men may benefit more as they are often less likely to seek help for anxiety/depression and are expected to put on a brave face to cope with baldness. A recent population-based alopecia areata epidemiology study in UK primary care, (M. Harries et al., 2021) covering 4.16m patient records, found that alopecia areata is more common in people. • living in urban areas compared to rural areas. • living in socially deprived areas. • of non-white ethnicity compared to those of white ethnicity. It was three times as common in people of Asian ethnicity. People in these groups are likely to benefit proportionality more. In some communities alopecia areata is seen as a cultural weakness. Also, appropriate orthotics (wigs) are more difficult to source for diverse hair types i.e., hair style/texture

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Alopecia areata is a visible difference and often develops into a 'hidden disability', mental health issues and psychosocial impact. As the research on stigma highlighted, lay people would stigmatise bald images which could affect the quality of life of people with alopecia (Creadore, Andrew et al. JAMA Dermatology, 2021:157(4)392-398).

Other issues

13. Are there any other issues that you would like the committee to consider?	Access to treatments and expertise on alopecia areata is currently still a postcode lottery. Alopecia UK hope that the committee will consider how to ensure fair and equitable access to this treatment across England (& 4 nations) once this treatment is approved. The degree of psychosocial impact is probably more important than the percentage of hair loss for many patients.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission.	• Alopecia areata is NOT just cosmetic, it is an autoimmune condition. • Alopecia areata – it is not just about the degree of hair loss, please consider the impact on the quality of life lived with a non-curable and unpredictable visible difference. There are debilitating mental health conditions (depression, anxiety) and psychosocial impacts (isolation, panic, absenteeism, life outcomes). • The process of losing your hair can be traumatic, and like any other trauma, this can lead to unhealthy coping strategies and lasting effects on health, behaviours, and life potential. These are costly to the individual, society at large, and the NHS. • This treatment gives hope – there is no cure and very few effective treatments. Effective being number of people helped and % hair regrowth. • Approving this technology will not open the flood gates. Only 1 in 4 people are referred to secondary care, many have limited patchy hair loss and this will not be the preferred treatment, and even when a JAK is a potential treatment many people will choose not to take it with the risk of side effects.

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 select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Professional organisation submission

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] on behalf of the British Association of Dermatologists' Therapy & Guidelines sub-committee, and [redacted] on behalf of the BAD's guideline development group and [redacted] on behalf of the British Hair & Nail Society
2. Name of organisation	British Association of Dermatologists (the BAD)
3. Job title or position	Consultant dermatologists
4. Are you (please select Yes or No):	An employee or representative of a healthcare professional organisation that represents clinicians? Yes A specialist in the treatment of people with this condition? Yes A specialist in the clinical evidence base for this condition or technology? Yes Other (please specify):
5a. Brief description of the organisation (including who funds it).	The BAD is a not-for-profit organisation whose charitable objectives are the practice, teaching, training, and research of dermatology. It works with the Department of Health, patient bodies and commissioners across the UK, advising on best practice and the provision of dermatology services across all service settings. It is funded by the activities of its members.
5b. 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 purpose of funding.**	No.
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No.

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.)	To increase the probability of more significant hair regrowth in those with severe alopecia areata (AA), control/prevent progression and improve quality of life (QoL).
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.)	At least a 50% reduction in hair loss (i.e. SALT50, analogous to PASI90/75/50 in psoriasis), improvement in QoL and significant patient-rated hair growth (e.g. able to stop wearing a wig/camouflage).
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, there is an unmet need. Current quality of evidence for most AA treatments is poor with high relapse rates (Meah *et al.* https://pubmed.ncbi.nlm.nih.gov/32165196/) and come with significant adverse effects (e.g. oral corticosteroids and immunosuppressants). Baricitinib has now been approved by the United States Food and Drug Administration (FDA), the first ever systemic treatment approved for AA.

What is the expected place of the technology in current practice?

9. How is the condition currently treated in the NHS?	Primary care clinicians will treat many patients with mild disease with topical corticosteroids or observe those with limited disease. Secondary care dermatologists and paediatric dermatologists will treat the majority of individuals with severe disease, but referral rates are lower in those of lower socioeconomic status. There are also a limited number of tertiary care hair specialist dermatologists in the UK who will treat the full spectrum of extent of hair loss but will also be referred patients in whom there are complex issues or if available treatments have failed and specialist treatments are needed. Limiting the availability of the drug to

	those who have been reviewed by a tertiary specialist may lead to geographic inequalities in drug availability. Initiation of treatment varies. Current primary care guidance suggests that a "watch and wait" policy in recent-onset, limited patch AA is reasonable as spontaneous regrowth is common. When treatment is given in primary care this usually comprises a topical corticosteroid (see Harries *et al.* https://onlinelibrary.wiley.com/doi/10.1111/bjd.20628 for further information on issued prescriptions in this population). However, 1 in 5 people with limited disease will go on to develop extensive AA from which spontaneous regrowth, or response to treatment, is rare (Tosti *et al.* https://onlinelibrary.wiley.com/doi/10.1111/bjd.20628). Therefore, many hair specialists advocate earlier treatment to prevent progression to more extensive disease (Meah *et al.* https://pubmed.ncbi.nlm.nih.gov/32165196/).
9a. Are any clinical guidelines used in the treatment of the condition, and if so, which?	British Association of Dermatologists' guidelines for the management of AA 2012 https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2133.2012.10955.x. This guideline is currently being updated using the BAD's NICE-accredited guideline development process based on GRADE.
9b. 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.)	There are no licensed treatments specific for AA. Generally, janus kinase (JAK) inhibitors such as baricitinib would fit at the stage when topical contact immunotherapy (if available) is being considered, i.e. ≥50% hair loss that has not responded to topical +/- oral corticosteroids and intralesional corticosteroids (where appropriate). N.B. Topical contact immunotherapy can only treat *scalp* hair loss.
9c. What impact would the technology have on the current pathway of care?	It would provide an effective treatment option which can address the scalp, eyebrow/eyelash and body hair loss.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	The technology is new and will be helping address a significant unmet, clinical need for a safe, effective and approved medication for patients with moderate-to-severe AA. Despite low rates of success, commonly used systemic treatments for chronic AA include prednisolone (up to 50 mg daily), ciclosporin (up to 5 mg/kg), methotrexate (up to 25 mg weekly), azathioprine (up to 200 mg daily). Sulfasalazine is not as commonly prescribed. The treatment response to these immunosuppressants is very variable.

10a. How does healthcare resource use differ between the technology and current care?	Monitoring and assessment in clinics would be similar to other systemic agents. If using contact immunotherapy as a comparator, this technology would reduce outpatient attendances in many cases, as contact immunotherapy would require weekly dermatology outpatient attendances unless home treatment is offered. Home treatment is only offered in a few specific centres.
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Dermatology secondary care and tertiary specialist hair clinics. Restricting to tertiary clinics alone would lead to geographic inequality, due to the relatively small number of tertiary specialist hair clinics in the UK. Indirectly, this could also lead to exclusion of certain patient populations.
10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	No new facilities or equipment needed for this new oral medicine.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes, a greater proportion of patients have clinically meaningful hair regrowth, i.e. SALT50, or SALT scores of <20% at week 36 as per the BRAVE-AA1 and BRAVE-AA2 trials (https://www.nejm.org/doi/full/10.1056/NEJMoa2110343); however, some patients who are almost completely bald may find it more difficult to deal with patchy hair growth as they might choose to shave it off so that wigs fit better. It also provides a treatment option for those with eyebrow/eyelash and body hair loss which current treatments cannot address or are not as effective. The BRAVE-AA1 and BRAVE-AA2 trials showed 38.9% had full regrowth of eyebrows and 36.8% for eyelashes.
11a. Do you expect the technology to increase length of life more than current care?	Life expectancy is not a clinically relevant outcome in this condition. Quality of life is a more relevant outcome for patients with alopecia areata.
11b. Do you expect the technology to increase health-related quality of life more than current care?	Yes, as it also helps regrow eyebrows/eyelashes, although some patients experience patchy hair re-growth which *may* result in a reduced improvement in QoL; anecdotally, these patients have opted to continue the treatment for this reason (eyebrows/eyelashes regrowth) with significant improvement in quality of life, self-esteem/confidence which then also impacts on their relationships and careers. This requires more objective measures to be performed.

12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	We are unlikely to advocate the use of this technology for acute alopecia areata (AAA) which is defined by disease duration less than 6 months. AAA has a high rate of spontaneous remission and systemic therapy is rarely required.

The use of the technology

13. Will the technology be 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.)	Concomitant treatments may be needed to optimise baricitinib and achieve SALT score of 0% in some cases. JAK inhibitors are already used by dermatologists for eczema and therefore the blood test monitoring, etc would be in place already.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Certain criteria such as extent of hair loss, duration of disease (chronic AA), involvement of facial/body hair and psychosocial impact of disease may provide guidance in terms of setting the initiation criteria. Treatment is usually stopped if there is no hair regrowth after 12 months of treatment. Some patients can take 6-9 months to start demonstrating any hair growth.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that	Those with AA have a significant mental health burden associated with their disease and hopefully availability of evidence-based treatments will possibly improve this, although this is yet to be proven in

are unlikely to be included in the quality-adjusted life year (QALY) calculation?	clinical trials. AA is also associated with time away from work, which will have a significant economic impact on the wider population. It is difficult to truly capture the impact of treatments for AA using QALYs, as this may not take into account the domains relevant to our patient population; perhaps another measure may need to be considered. Some health-related QoL measures may not capture adequately the impact of living with health conditions in older people (questions about work, studying, sport) or those who are not in a relationship (question about sexual activity); additionally, they may not capture anxiety and depression across all groups – two parameters that are commonly and negatively influenced by AA. Additionally, they may discriminate against those who are non-native English speakers.
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?	Baricitinib is innovative as it has demonstrated the greatest efficacy out of all other systemics previously used such as ciclosporin, methotrexate and azathioprine, and with a better side effects profile and level of immunosuppression. Other current therapies include contact immunotherapy (diphencyprone) which is not readily available as only a few centres in the UK are able to deliver this service. Good-quality wigs are expensive and there is variability in access to/support for these across the UK; also, this intervention does not help with facial/body hair loss. Baricitinib will make a significant impact on health-related benefits as their AA can be better controlled, by extension their well-being will improve, thus having a positive impact on the psychosocial aspects of their life.
16a. Is the technology a 'step-change' in the management of the condition?	Yes, it would be a step change since there is no effective *and* safe systemic treatment for severe AA. Current available therapies for AA are often ineffective, and topical corticosteroids are usually ineffective in severe AA. Regular clinic visits, blood monitoring and drug costs, along with wig prescription and wider societal issues (e.g. unemployment) all contribute to the impact of AA on the individual, NHS and society more widely. Effective treatment options are needed urgently to prevent the longer term sequalae of ongoing AA (e.g. mental health issues).
16b. Does the use of the technology address any particular unmet need of the patient population?	There is no licensed systemic treatment for AA. The FDA recently approved baricitinib used to treat severe AA in the US. The expert paper (Meah *et al.* https://pubmed.ncbi.nlm.nih.gov/32165196/) reported that consensus was achieved for the following statement regarding preferred second-line agents for AA: "If all treatments were equally reimbursed, JAK inhibitors would be the ideal choice for systemic therapy in adults".

17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	If adverse effects occur, upper respiratory tract infections, cutaneous HSV/VZV and acne would affect the patient's QoL. Acne was more common in the baricitinib arm compared with the placebo arm, occurring in 16/280 patients (5.7%) with 4 mg baricitinib, 10/183 (5.5%) with 2 mg baricitinib and 1/189 (0.5%) with placebo in the BRAVE-AA1 trial and in 11/233 patients (4.7%), 9/155 (5.8%) and 3/154 (1.9%), respectively, in the BRAVE-AA2 trial. Herpes zoster infections was found in 2/280 patients (0.7%) with 4 mg baricitinib, 1/183 (0.5%) with 2 mg baricitinib and 1/189 (0.5%) with placebo in the BRAVE-AA1 trial and in 3/233 patients (1.3%), 3/155 (1.9%), and 1/154 (0.6%), respectively, in the BRAVE-AA2 trial – these were localised with no disseminated infections. There were no venous thromboembolic events, opportunistic infections or gastrointestinal perforations in either trial. The incidence of urinary tract infection was higher with those receiving baricitinib than placebo in the BRAVE-AA2 trial, with such infection occurring in 11/233 patients (4.7%) with 4 mg baricitinib, 12/155 (7.7%) with 2 mg baricitinib and 2/154 (1.3%) with placebo. The commonest biochemical abnormality was raised low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol, observed in approximately 25% and 40% of patients in the baricitinib groups, respectively.

Sources of evidence

18. Do the clinical trials on the technology reflect current UK clinical practice?	
18a. If not, how could the results be extrapolated to the UK setting?	

18b. What, in your view, are the most important outcomes, and were they measured in the trials?	At least a 50% reduction in hair loss (i.e. SALT50, analogous to PASI90/75/50 in psoriasis), improvement in QoL and significant patient-rated hair growth (e.g. able to stop wearing a wig/camouflage)
18c. If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Long-term outcomes in alopecia areata are unpredictable. Further long-term, real-world studies would be needed to assess long-term outcomes.
18d. Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not that we are aware of. In clinics, patients are demonstrating similar adverse effects to those observed in the trials with the main one being mild-to-moderate acne. Overall, no suggestion of serious infections and low rates of discontinuation due to adverse effects in clinic, reflecting clinical trial data.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	Not that we are aware of.
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance?	Not that we are aware of.
21. How do data on real-world experience compare with the trial data?	Anecdotally in clinics, patients demonstrate excellent response on the higher dose. It is very well tolerated and mild-to-moderate facial acne is a common side effect but easily managed with topical treatments. It is showing excellent and meaningful response in patients who would have been excluded from trials either due to their duration of disease being >8 years, their age being >65 years or their comorbidities such as immunodeficiencies, etc. It does require dose titrations and optimisation during their management.

Equality

22a. Are there any potential equality issues that should be taken into account when considering this treatment?	Having a disease duration cut-off of 8 years will indirectly lead to possible age-discrimination. Epidemiological data has shown that AA is more common in those of Asian background and those of lower socioeconomic status and urban location, but referral to secondary care is lower in these groups (Harries *et al.* https://onlinelibrary.wiley.com/doi/10.1111/bjd.20628). Inclusion of individuals with these characteristics is important in the clinical and cost-effectiveness data and in the patient representation in the consultation process. Beard hair loss can have some religious implications, e.g. some from the Sikh and Jewish faiths. Here, many standard treatments are more challenging for beard hair loss, where systemic medication is often required at an earlier stage. Including adolescents (age 12-17) with severe AA: treatment of children and young people with AA is very challenging and increasing available treatments would have a significant impact in this patient population. Although the peak incidence of AA onset is those aged 25-29 years (Harries *et al.* https://onlinelibrary.wiley.com/doi/10.1111/bjd.20628), a significant proportion of patients first experience AA in childhood or adolescent years. This group tends to have a worse prognosis, and visible hair loss can have a profound impact psychologically at this stage of development. Some health-related QoL measures may not capture adequately the impact of living with health conditions in older people (questions about work, studying, sport) or those who are not in a relationship (question about sexual activity); they may also not capture anxiety and depression across all groups – two parameters that are commonly and negatively influenced by AA. Additionally, they may discriminate against those who are non-native English speakers. Geographic variability in wig provision could mean that certain geographic locations are already disadvantaged financially, by having to buy their own wigs for camouflage. This could indirectly affect specific minority populations based on geography. Providing an effective systemic treatment for alopecia areata with geographic equity may seek to address this. It is therefore important that this treatment is not limited to provision at the small number of tertiary hair clinics and instead is available at all secondary care dermatology sites, provided clinical criteria are applied to ensure appropriate use of resources.

22b. Consider whether these issues are different from issues with current care and why.	

Key messages

23. In up to 5 bullet points, please summarise the key messages of your submission.	• Alopecia areata is a chronic, autoimmune disease with significant psychosocial implications including social isolation and withdrawal, work absenteeism, illness-induced career change, loss of income, loneliness, failure to establish relationships and relationship (including marriage) breakdown, anxiety, depression, suicidal ideation, attempted suicide and actual suicide. • There is a significant unmet, clinical need for a safe, effective and approved medication for people with moderate-to-severe AA. • Initial trial data to date indicate that this treatment is effective and with a good safety profile. • In the BRAVE-AA1 and BRAVE-AA2 studies, 38.8% and 35.9% of patients, respectively, who received baricitinib 4 mg daily, achieved clinically meaningful hair regrowth including scalp, eyebrow and eyelashes.

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 select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

Baricitinib for treating severe alopecia areata [ID 3979]

STA Report

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135648T.

Title:	Baricitinib for treating severe alopecia areata
Produced by:	BMJ Technology Assessment Group (BMJ TAG)
Authors:	Steve Edwards, Director of Health Technology Assessment, BMJ-TAG, London Ben Farrar, Senior Clinical Evidence Analyst, BMJ-TAG, London Tracey Jhita, Health Economist Manager, BMJ-TAG, London Victoria Wakefield, Principal Clinical Evidence Analyst, BMJ-TAG, London Archie Walters, Health Economist, BMJ-TAG, London
Correspondence to:	Steve Edwards, BMJ TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	13/10/20211
Source of funding:	This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135648T.
Declared competing interests of the authors	Ben Farrar was employed by Costello Medical between September 2020 to December 2020 and November 2021 to May 2022 but did not work on baricitinib or any of its comparators while employed there. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The EAG would like to thank Professor Andrew Messenger, Honorary Professor of Dermatology, Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield and Dr Susan Holmes, Consultant Dermatologist, Alan Lyell Centre for Dermatology, Glasgow Royal Infirmary and Honorary Associate Clinical Professor, University of Glasgow for providing clinical advice throughout the project. The EAG would also like to thank Dr Leila Asfour, Consultant Dermatologist, Department of Dermatology, Salford Royal Foundation Trust, and Dr Nekma Meah, Consultant Dermatologist, St Helens and Knowsley NHS Trust and Honorary Clinical Lecturer, University of Manchester for providing feedback on the clinical sections of the report.
Rider on responsibility for report:	The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.
Report reference:	Edwards SJ, Farrar B, Jhita T, Wakefield V and Walters A. Baricitinib for treating severe alopecia areata : A Single Technology Appraisal. BMJ Technology Assessment Group, 2022.

Contribution of authors:

Steve Edwards	Critical appraisal of the company's submission; validated the statistical analyses; provided feedback on all versions of the report. Guarantor of the report.
Ben Farrar	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; and drafted the summary, background and clinical results sections.
Victoria Wakefield	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and assisted with drafting the clinical results sections.
Tracey Jhita	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections
Archie Walters	Critical appraisal of the company's submission; critical appraisal of the economic model; and drafted the economic sections

All authors read and commented on draft versions of the EAG report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA	Alopecia areata
AE	Adverse event
AESI	Adverse event of special interest
ANCOVA	Analysis of covariance
AU	Alopecia universalis
BAD	British Association of Dermatologists
BARI	Baricitinib
BID	Twice per day
BSC	Best supportive care
CENTRAL	Cochrane Central Register of Controlled Trials
CDSR	Cochrane Database of Systematic Reviews
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
ClinRO	Clinician reported outcome
CMU	Confidential medicines unit
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
DLQI	Dermatology Life Quality Index
DPCP	Diphenylcyclopropenone
DSA	Deterministic sensitivity analysis
DSP	Disease Specific Programme
DSU	Decision Support Unit
EAG	Evidence assessment group
EB	Eyebrow
EL	Eyelash
EMA	European Medicines Agency
eMIT	Drugs and pharmaceutical electronic market information tool
EQ-5D	European Quality of Life-5 Dimensions
FAS	Ful analysis set
GP	General practitioner
HADS	Hospital Anxiety Depression Scale
HI	High intensity
HRQoL	Health-related quality of life
HSUV	Health state utility values
ICER	Incremental cost-effectiveness ratio
IL	Intralesional
IMT	Immunotherapy

ITC	Indirect treatment comparison
IWRS	Interactive web response system
JAK	Janus kinase
LSM	Least squares mean
LYG	Life-years gained
MACE	Major adverse cardiovascular event
MBCR	Mindfulness-based cognitive therapy
MHRA	Medicines and Healthcare products Regulatory Agency
mLOCF	Modified last observation carried forward
NHB	Net health benefit
NHS	National Health Service
NICE	National Institute for Health and Care Excellent
NMSC	Nonmelanoma skin cancer
NMA	Network meta-analysis
NMB	Net monetary benefit
NRI	Non-responder imputation
ONS	Office for National Statistics
OWSA	One-way sensitivity analysis
PAS	Patient Access Scheme
PBO	Placebo
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSSRU	Personal Social Services Research Unit
PWP	Psychological well-being practitioner
PTFU	Post-trial follow up
QALY	Quality-adjusted life year
QD	Once daily
RCT	Randomised controlled trial
SAE	Serious adverse event
SALT	Severity of Alopecia Tool
SD	Standard deviation
SE	Standard error
SF-36	Medical outcomes study 36-item short form health survey
SLR	Systematic literature review
STA	Single technology appraisal
STAT	Signal transducers and activators of transcription
TCS	Topical corticosteroids
TEAE	Treatment emergent adverse event
TSD	Technical Support Document

UK	United Kingdom
US	United States
VAS	Visual analogue scale
W	Week
WPAI	Work productivity and activity impairment
WTP	Willingness to pay

1 Executive summary

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

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

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

1.1 Overview of the EAG's key issues

Table 1 presents a summary of the EAG's key issues on the evidence submitted on the clinical and cost effectiveness baricitinib for treating adults with severe alopecia areata (AA).

Table 1. Summary of key issues

ID xxx	Summary of issue	Report sections
1	Definition of the comparator	2.2.1, 2.3.3, 3.4, 4.2.3
2	Definition of treatment response at Week 36	4.2.5
2	Source of utilities in the model	4.2.8
4	Disease monitoring costs for best supportive care	4.2.9

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are around the definition of the comparator and treatment response at Week 36, the source of utilities used in the model, and the assumptions of costs incurred in the best supportive care (BSC) health state. However, other secondary differences in the preferred assumptions between the company and EAG's approach include how long-term all-cause treatment discontinuation for baricitinib is calculated, inclusion of adverse events (AEs), removal of non-pharmacological psychological support costs and wig resource use in the induction phase of the model.

It should be noted that for AEs, the EAG was unable to verify the inputs used in the company's AE scenario provided in their clarification response and was unable to produce an alternative scenario due to a paucity of time. Nonetheless, the EAG requests that during technical engagement, the company provides a more thorough description and justification of their approach to the inclusion of AEs and assumed unit costs to treat each AE and update the scenario if necessary.

1.2 Overview of key model outcomes

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

Overall, the technology is modelled to affect QALYs by:

- Improving and maintaining scalp hair regrowth.

Overall, the technology is modelled to affect costs by:

- Its higher unit price than established clinical management.

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

- Changing the definition of the comparator to 'discharged from care' and removing all monitoring costs in the induction phase and Maintenance health state as a result.
- Using utilities in the model sourced from the key trials of baricitinib, BRAVE-AA1 and BRAVE-AA2.
- Removing the costs associated with disease management in the BSC health state.

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

Table 2. Issue 1: Definition of the comparator

Report section	2.2.1, 2.3.3, 3.4, 4.2.3
Description of issue and why the EAG has identified it as important	The company's comparator is "Watch and wait", which is defined as no active treatment but frequent monitoring. In contrast, the EAG's clinical experts advised that, although no active treatment is a common management strategy for adults with severe AA, the company's definition of "Watch and wait" did not capture this adequately. The EAG considered three alternative comparators: • Treatment with DPCP, the most effective treatment currently used to treat severe AA in adults, which is the only active treatment recommended by the British Association of Dermatologists Guidelines for treating severe AA; • Treatment from a basket of "low-effectiveness" non-DPCP therapies sometimes used to treat severe AA in adults, primarily systemic immunosuppressants and systemic corticosteroids; • No active treatment and discharge from care. The EAG concluded that: • DPCP is not a reasonable comparator for the appraisal as DPCP is only available to a minority of patients with inequitable access and is associated with severe adverse events and a high rate of relapse; • No active treatment and discharge from care is the most commonly used approach for the prevalent population of adults with severe AA who would be eligible to receive baricitinib at the point of approval in the UK; • While systemic immunosuppressants and systemic steroids could be considered appropriate comparators for newly diagnosed cases of severe AA, their use is too heterogenous and their effectiveness too limited to be considered an established standard of care for severe AA. In lieu of robust treatment pattern data or comparative effectiveness data with baricitinib, the EAG considers no active treatment and discharge from care to be an acceptable comparator for this population.
What alternative approach has the EAG suggested?	The EAG recommends no active treatment and discharge from care as the most appropriate comparator for this appraisal. As such, in the economic model, the EAG considers that monitoring costs included for the induction phase and maintenance health state should be removed.
What is the expected effect on the cost-effectiveness estimates?	The impact on the company's ICER post clarification when monitoring costs are removed in the induction phase and maintenance health state for 'Watch and wait' (which redefines the comparator to 'discharged from care'), increases from £18,072 to £20,887.
What additional evidence or analyses might help to resolve this key issue?	Comprehensive treatment pattern data for AA and severe AA from a range of care settings in the UK would help to resolve some of the uncertainty in the treatment pathway of AA, especially for newly diagnosed severe AA patients. The EAG's clinical experts highlighted that such data do not exist to their knowledge.

Abbreviations: AA: alopecia areata: DPCP: diphencyprone EAG: evidence assessment group; ICER; incremental cost-effectiveness ratio; ITC: indirect treatment comparison NMA: network meta-analysis

1.4 The clinical and cost-effectiveness evidence: summary of the EAG's key issues

Table 3. Issue 2: Definition of treatment response at Week 36

Report section	4.2.5
Description of issue and why the EAG has identified it as important	In the company's base case, the primary outcome in the model was $SALT_{50}$ (defined as at least a 50% improvement from baseline SALT score). In addition to the outcome of $SALT_{50}$, the company also included the outcome of $SALT_{75}$ (defined as at least a 75% improvement from baseline SALT score), as a way of capturing additional quality of life benefit associated with achieving an increased relative improvement in scalp hair growth. In the key trials of BRAVE-AA1 and BRAVE-AA2, the primary endpoint was the proportion of patients achieving SALT≤20 at Week 36. A response of SALT≤20 indicated scalp hair loss of less than 20% (or ≥80% scalp coverage with hair). The EAG considers using SALT≤20 to be a more clinically meaningful benefit for patients. The EAG's clinical experts noted that a relative benefit of $SALT_{50}$ or $SALT_{75}$, is unlikely to be meaningful to patients unless it results in a similar increase in coverage to SALT≤20.
What alternative approach has the EAG suggested?	The EAG's preferred approach is to use SALT≤20 as the definition of treatment response at Week 36.
What is the expected effect on the cost-effectiveness estimates?	By using the outcome of SALT≤20 at Week 36, the company's ICER post clarification reduced from £18,072 to £17,071 for the overall population. For the severe and very severe subgroups, the company's ICER post clarification changed from £25,154 (severe) and £12,685 (very severe) to £18,773 (severe) and £16,929 (very severe), respectively.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence required as the scenario resolves the issue.

Abbreviations: EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; SALT, Severity of Alopecia Tool

Table 4. Issue 3: Source of utilities in the model

Report section	4.2.8
Description of issue and why the EAG has identified it as important	The BRAVE-AA1 and BRAVE-AA2 trials collected EQ-5D data up to Week 36 directly from patients but the company stated that the values obtained from the trials were insensitive to changes in the severity of AA and lacked content validity as baseline values were almost the same as UK age- and sex-adjusted general population values. Additionally, the company stated that [redacted] of participants in the BRAVE-AA1 and BRAVE-AA2 trials reported a score of perfect health at baseline (score of 11111) and as such an improvement in HRQoL would not be obtained at Week 36 for these patients. Thus, the utility values informing the economic model were derived from a company sponsored Adelphi DSP study. The company explained that in the Adelphi DSP study, the ceiling effect was also observed, but not to the same extent. However, the company did not provide the overall proportion of patients reporting a score of perfect health from the Adelphi DSP study.

	The EAG considers the company's justification for not using pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials is a criticism of the EQ-5D tool and not the methods to obtain the data used in the trial and thus extends to the EQ-5D data obtained from the company sponsored Adelphi DSP study. Furthermore, the company hasn't supplied sufficient evidence to validate the lack of content validity with the EQ-5D nor has it demonstrated why patients should have a substantial change in their QoL. As recommended in the NICE methods guide, the reference case should report the measurement of changes in health-related quality of life directly from patients. As such, the EAG considers the pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials represents a more robust source of utility data that matches the NICE reference case and should be used in the cost-effectiveness analysis for the base case.
What alternative approach has the EAG suggested?	During the clarification stage, the EAG requested, and the company provided, change from baseline at Week 36 for patients achieving SALT≤20 based on pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials. Thus, when using the outcome of SALT≤20 for treatment response at Week 36, the baseline utility and change from baseline associated with achieving SALT≤20 should be used in the model.
What is the expected effect on the cost-effectiveness estimates?	When implementing the baseline utility and change from baseline utility associated with achieving SALT≤20 in combination with using SALT≤20 at Week 36, the company's ICER post clarification increases from £18,072 to £118,494 for the overall population. For the severe and very severe subgroups, the ICERs are £130,303 and £117,510, respectively.
What additional evidence or analyses might help to resolve this key issue?	The EAG's clinical experts advised that for most patients' HRQoL may only be mildly affected and thus may not be that different to the general population but equally HRQoL is severely affected for a few patients (primarily driven by adverse mental health). Additionally, the EAG's clinical experts advised that overtime, some patients may come to terms with their hair loss, while a few may remain distressed about their condition. Thus, the EAG acknowledges that there is a small, but heterogenous, patient population that is more adversely affected in terms of HRQoL but that the demographics of this population are difficult to identify clinically and consistently, and it is beyond the scope of this assessment to identify that group. Nonetheless, the EAG has estimated the QALY gain needed to reach the £20,000 and £30,000 cost-effectiveness thresholds and advises the committee to consider if the estimated QALY gain needed for baricitinib 4 mg is plausible for the condition under consideration.

Abbreviations: AA, alopecia areata; DSP, disease specific programme; EAG, Evidence Assessment Group; HRQoL, health-related quality of life; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; SALT, Severity of Alopecia Tool

Table 5. Issue 4: Disease monitoring costs for best supportive care

Report section	4.2.9.3 and 4.2.9.5
Description of issue and why the EAG has identified it as important	The EAG's clinical experts considered that it is likely that if response to treatment is not achieved, patients will not engage with further treatment and will not be followed up (effectively patients are discharged from care). The EAG considers that lack of engagement with treatment and being discharged from care has implications for the costs assumed in the BSC health state for both arms of the model, as patients transition to this health state upon loss of treatment response or treatment discontinuation for any other reason.
What alternative approach has the EAG suggested?	The EAG considers that disease management costs in the BSC health state should be excluded for both arms of the model.
What is the expected effect on the cost-effectiveness estimates?	Removal of disease monitoring costs in the BSC health state for both arms of the model increased the company's ICER post clarification from £18,072 to £63,941. However, when combined with a change to the treatment response definition (SALT≤20) and source of utilities, the ICER increases to £419,926.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence required as the scenario resolves the issue.

Abbreviations: BSC, best supportive care; EAG, Evidence Assessment Group; HRQoL, health-related quality of life; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; SALT, Severity of Alopecia Tool

1.5 Other key issues: summary of the EAG's view

- The EAG considers the BRAVE-AA trial populations to be narrower than the population of the final scope issued by NICE. Specifically, patients with current AA episodes >8 years and who had showed no sign of previous regrowth and patients >60 years (males) and >70 years (females) were excluded from the BRAVE-AA trials. These patients would be eligible to receive baricitinib as per the [redacted] marketing authorisation, but likely have a lower probability of response than the trial populations.
- The BRAVE-AA trial populations, having relatively long disease and episode durations at baseline and being treatment-experienced, are more similar to the prevalent population in the UK than to newly diagnosed patients severe AA. This may cause the trials to underestimate treatment effectiveness in newly diagnosed severe AA patients, as shorter current AA episodes are associated with favourable treatment response, and treatment inexperience may also be associated with favourable treatment response. The EAG notes, however, that the magnitude of any effect of treatment experience on response to baricitinib is uncertain because the mode of action of baricitinib is different to current therapies used to treat severe AA.
- Current AA episode duration and baseline SALT score are clinically meaningful variables that predict treatment response and vary substantially in the trial. Shorter AA episodes and lower

baseline SALT scores are associated with a higher probability of treatment response. Categorising AA episode duration and baseline SALT score could form clinically meaningful subgroups, however any categorisation of these continuous variables would be arbitrary.

1.6 Summary of EAG's preferred assumptions and resulting ICER

Table 6 presents the EAG preferred assumptions as well as the EAG deterministic and probabilistic base case ICER. Table 7 presents scenarios around the EAG base case.

Table 6. EAG preferred assumptions and base case ICER

Scenario	Incremental costs	Incremental QALYs	ICER (change from company base case
Company base case post clarification	[redacted]	[redacted]	18,072
SALT≤20 at Week 36	[redacted]	[redacted]	17,071
SALT≤20 at Week 36 + baseline and CFB utility from BRAVE trials	[redacted]	[redacted]	118,494
Long-term all-cause discontinuation based on Week 36-52 data for baricitinib 4 mg ([redacted])	[redacted]	[redacted]	107,217
No monitoring costs in the induction phase and Maintenance health state for 'Watch and wait' (comparator defined as 'discharged from care')	[redacted]	[redacted]	126,309
Removal of disease monitoring costs in the BSC health state for both arms of the model	[redacted]	[redacted]	419,926
Removal of non-pharmacological psychological support costs	[redacted]	[redacted]	423,809
One wig assumed in the induction phase for both arms of the model	[redacted]	[redacted]	423,775
EAG's preferred deterministic base case - combination of all scenarios	[redacted]	[redacted]	**423,775**
EAG's preferred probabilistic base case - combination of all scenarios	[redacted]	[redacted]	**379,030**

Abbreviations: BSC, best supportive care; CFB, change from baseline; EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; SALT, Severity of Alopecia Tool

*It should be noted that QALY gain in the probabilistic analysis is [redacted] higher than the deterministic analysis. However, given that the incremental costs and QALYs are relatively small, the ICERs are sensitive to very small changes.

Table 7. Deterministic scenarios around the EAG base case

	Results per patient	Baricitinib 4 mg	'Discharged from care'	Incremental value
0	**EAG base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			423,775
1	**Severe subgroup - baseline SALT 50-95**			

	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			407,212
2	**Very severe subgroup - baseline SALT 95-100**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			456,573

Abbreviations: EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

Note: the same baseline utility ([redacted]), change from baseline ([redacted]) and treatment discontinuation rate ([redacted]) have been used for the subgroups as for the base case as the relevant data were not available by severity.

For further details of the exploratory and sensitivity analyses done by the EAG, see Section 6.3.

2 Introduction and background

2.1 Introduction

This document contains the Evidence Assessment Group's (EAG's) critique of the clinical and cost-effectiveness evidence submitted for the Single Technology Appraisal (STA) of baricitinib (brand name Olumiant®; Eli Lilly and Company) in the treatment of severe alopecia areata (AA) in adults.

2.2 Background

Section B.1 of the company submission (CS) provides information on:

- AA, including its aetiology, burden of disease and current pathway of clinical care in the NHS, and;
- Baricitinib, including its mechanism of action, details of its pending marketing authorisation, its costs and its method of administration and dosage.

The EAG's clinical experts agreed that Section B.1.3 of the CS provides a reasonable overview of AA, its aetiology and burden of disease. AA is an autoimmune disease that leads to non-scarring loss of hair on a person's scalp, face or body. In 2018, 0.58% of UK adults who were registered in electronic primary care records had an active or historic diagnosis of AA.[1] While the exact aetiology of AA is unknown, a suite of genetic risk factors[2] and environmental stressors[3, 4] exist that heighten the risk of AA.

AA can vary in severity, which can be measured using The Severity of Alopecia Tool (SALT). The SALT score ranges from 0-100 and measures the severity of scalp hair loss, with 0 corresponding to complete loss of scalp hair and 100 corresponding to a full head of hair. In the CS, severe AA is defined as SALT 50–94 and very severe AA is defined as SALT ≥95. Severity is associated with prognosis: those who are missing more scalp hair are less likely to regrow hair, either spontaneously or through treatment.[4, 5] Similarly, the length of an AA episode is related to prognosis: the longer a patient has had an AA episode for, the less likely the patient is to regrow hair.[5]

AA is caused by the loss of immune privilege of hair follicles. This occurs due to the production of pro-inflammatory cytokines, such as interferon-gamma, causing the stimulation of natural killer cell receptors and subsequent activation of the janus kinase (JAK) signal transducer and activator of transcription (STAT) signalling pathway (JAK/STAT). The inflammation associated with JAK/STAT activation causes the early termination of the anagen phase in hair follicles, preventing hair growth.[6, 7] Drugs that inhibit JAK therefore have the potential to prevent and reverse autoimmune hair loss in

AA.[8] Baricitinib is one such JAK inhibitor that selectively and reversibly inhibits JAK1 and JAK2, and is expected to receive marketing authorisation from the Medicines and Healthcare products Regulatory Agency (MHRA) for treating severe AA in adults in [redacted]. In this STA, baricitinib is being evaluated as an oral JAK inhibitor taken as a 4 mg dose once daily, with a lower dose of 2 mg once daily potentially being more suitable for some subgroups, such as those >75 years or those with a history of chronic or recurrent infections.

2.2.1 Treatment pathway for severe AA

There are no approved treatments for severe AA, defined as SALT ≥50, in England and Wales, and there is a clear unmet need for these patients. Several JAK inhibitors are in development for AA, but baricitinib is the first to undergo an appraisal by NICE for this indication. In the CS, Baricitinib is positioned as: i) a first-line treatment for severe AA, and ii) a later-line treatment to treat patients with severe AA who do not respond to other treatment strategies. The EAG's clinical experts thought this positioning is an accurate reflection of where baricitinib would be used in the treatment pathway for severe AA. The first wave of eligible patients, the prevalent population, would likely be later-line patients who have failed on or were intolerant to pre-existing therapies, and after this, baricitinib would become a preferred first-line therapy for newly diagnosed severe AA.

In Section B.1.3.3, the company outline their interpretation of the current treatment pathway for severe AA: patients may initially be left untreated under a "Watch and wait" approach similar to that used in mild AA, or patients can be treated from a range of often off-label therapies that have limited effectiveness in severe AA. These treatments include topical, intralesional (IL) or oral corticosteroids, topical immunotherapy, immunosuppressives such as methotrexate, and minoxidil and calcineurin inhibitors. In the economic analysis, the company defines established clinical management as "Watch and wait" followed by best supportive care—which comprised of the range of off-label therapies and psychological support.

The EAG's clinical experts believed it reasonable that patients with severe AA may be untreated up to around six months, however noted that this reflects the wait period to see a dermatologist rather than necessarily a decision to "Watch and wait". The EAG's clinical experts believed that most patients would have used a potent topical steroid during milder disease, or IL steroids if a patient visited a dermatologist. For severe AA, topical immunotherapy, systemic steroids and systemic immunosuppressants may be offered, and a small number of patients may have received these with milder disease, too. The EAG's clinical experts highlighted that there is no clear single standard management for severe AA and highlighted how only topical immunotherapy and wigs are

recommended by the British Association for Dermatologists (BAD) guidelines for treating severe AA.[9] Treatment for severe AA varies based on setting (primary care, specialist dermatologist, specialist dermatologist with an interest in AA), treatment availability and patient preference. Not all dermatologists will offer or have access to the more effective best supportive care therapies and not all patients will opt to take them, instead managing the condition with wigs or head shaving, if required. While the 2012 BAD guidelines for the management of alopecia areata recommend topical immunotherapy, e.g., DPCP, for extensive patchy hair loss and alopecia totalis/universalis,[9] DPCP is not widely available across the NHS, can lead to potent allergic reactions in patients and staff, and has a high rate hair-loss recurrence. For example, a meta-analysis reported a recurrence rate of 38% in patients receiving maintenance treatment and treatment-emergent severe eczema in 31% of DPCP treated patients, although the EAG's clinical experts noted this may be an overestimation of the rate of severe eczema.[10]

While the EAG's clinical experts did not recognise "Watch and wait" as a standard option for treating severe AA, they did recognise a similar no active treatment, and ultimately discharge from care, as a common management strategy opted for by severe AA patients. This was especially the case for the long-term care for patients who do not respond to treatment. Such patients may be prescribed wigs, or shave their heads, but this would not require intensive follow-up. This absence of intensive follow-up is the key difference between the no active treatment the EAG considers a common management strategy in clinical practice and the company's definition of "Watch and wait", which involves intensive surveillance and support. The EAG's clinical experts advised that most of the prevalent population will have opted for no active treatment and discharge from care, however they highlighted how there is a lack of comprehensive treatment pattern data for AA patients in the UK.

Hence, The EAG therefore considers there to be no established or highly- effective standard clinical management of severe AA, with no active treatment and discharge from care being a common endpoint. The EAG's clinical experts further highlighted the near absence of high-quality randomised controlled trials for the treatment of severe AA, excluding recent trials on JAK inhibitors.

2.3 Critique of the company's definition of the decision problem

In Table 1 of the CS, the company outlines: i) the final scope issued by NICE, ii) the decision problem addressed in the CS and, iii) the company's justification for differences between them. The EAG considers the decision problem addressed by the company to largely match the final scope issued by NICE.[11] However, the EAG notes that the two trials informing the clinical effectiveness data analysis in the submission, BRAVE-AA1 and BRAVE-AA2, have patient populations that differ in several

regards to the patient population that would be eligible to receive baricitinib in England and Wales, and that specified in the NICE final scope. Overall, however, the EAG considers the BRAVE-AA trial data to be suitable to inform decision making. An overview of the EAG's critique of the company's definition of the decision problem and the relevance of the BRAVE-AA trial populations can be found in Table 8.

Table 8. EAG critique of the decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope	EAG comment
Population	Adults with severe alopecia areata	Adults with severe alopecia areata	NA	The decision problem matches that of the final scope issued by NICE: adults with severe alopecia areata. However, the BRAVE-AA trials: • excluded patients with baseline AA episodes >8 years; • excluded males >60 years and females >70 years. Hence, the BRAVE-AA trials provide data on a narrower population than those who could receive baricitinib in clinical practice, and has excluded some of the patients least likely to respond to treatment.
Intervention	Baricitinib	Baricitinib	NA	The intervention described in the CS, baricitinib, matches the intervention described in the final scope. Baricitinib is an oral JAK inhibitor that is expected to receive marketing authorisation for treating severe AA in adults in [redacted].
Comparator(s)	Established clinical management without baricitinib	Established clinical management without baricitinib, which may include supportive care	NA	As per the NICE final scope, the company has outlined what it believes to be established clinical management without baricitinib, informed by real-world dataset and three clinical experts. The EAG's clinical experts outline how there was no clear established clinical management for severe AA, with a large degree of variation between centres. The EAG notes that:

				• There is no clear standard clinical management for severe AA; • No active treatment or follow-up is currently a realistic end prospect for severe AA patients; • The company likely overestimates the amount of psychological support patients receive in the NHS. The EAG's clinical experts highlighted how the availability of psychological care is far below what is needed for severe AA patients.
Outcomes	The outcome measures to be considered include: • Disease severity e.g. Severity of Alopecia Tool (SALT); • Improvement in hair loss e.g. Scalp Hair Assessment Score, Measure for Eyebrow Hair Loss, Measure for Eyelash Hair Loss; • Adverse effects of treatment; • Health-related quality of life.	The outcome measures to be considered include: • Measures of disease severity and improvement in hair loss (including SALT, ClinRO for eyebrow hair loss and eyelash hair loss, PRO measures for scalp hair assessment, PRO measures for eyelashes and eyebrows); • Adverse effects of treatment (including AEs, SAEs, AESIs); • Health-related quality of life (including EQ-5D, Skindex-16 AA, HADS and SF-36).	NA	The outcomes in the company's submission match the outcomes described in the final scope.
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of	As per NICE final scope	NA	NA, as per NICE final scope

	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.			
Subgroups to be considered	Due to an assumed typographical error, the NICE final scope was ambiguous about which subgroups were to be considered, stating that: "If the evidence allows, the following subgroups based on severity and type of alopecia areata will be considered", but without specifying any subgroups.	The company provided subgroup analyses and a scenario analysis based on the baseline severity of alopecia areata (severe disease, SALT 50-95 and very severe disease, SALT 95-100) and current duration of AA at baseline. No scenario analyses were presented based on the type of alopecia areata.	NA	NA
Special considerations, including issues related to equity or equality	None identified.	None identified.	NA	The EAG's clinical experts highlighted for some cultures loss of beard hair can be an important issue.

Abbreviations: AA: alopecia areata; AE: adverse event; AESI: adverse event of special interest; CS: company submission; EAG: evidence assessment group; EQ-5D: the European Quality of Life-5 Dimensions; HADS: Hospital Anxiety Depression Scale; JAK: janus kinase; NICE: National Institute for Health and Care Excellence; PRO: patient reported outcome; SAE: serious adverse event; SALT: Severity of Alopecia Tool; SF-36: Short Form 36 Health Survey Questionnaire.

2.3.1 Population

Two Phase III trials of baricitinib in adults with severe AA, BRAVE-AA1 and BRAVE-AA2,[12] inform the clinical effectiveness evidence in the submission. Despite only including adults with severe AA, the populations of BRAVE-AA1 and BRAVE-AA2 are from a narrower population than that of the NICE final scope.[11] Specifically, the following patients were excluded from both BRAVE-AA trials:

- Patients with current AA episodes >8 years and who had showed no sign of previous regrowth;
- Male patients >60 years and female patients >70 years.

Such patients would be eligible to receive baricitinib per the [redacted] marketing authorisation but may be less likely to achieve hair regrowth. Patients with longer episodes of AA have a lower probability of hair regrowth and treatment response (Section 3.3.2.2), and a less effective 2 mg dose may be used in patients >75 years, leading to a lower probability of treatment response. In addition, patients with co-existing hair loss conditions, such as male patients >60 years who have male pattern baldness, the amount of scalp hair regrowth possible could be limited and difficult to determine. The EAG's clinical experts also noted that it is plausible that patients with long AA episodes who have disengaged will reengage with care to receive baricitinib, should it become available.

Any overestimation of the effectiveness of baricitinib because of the trial exclusion criteria may be balanced by the fact that around [redacted] of participants in the BRAVE-AA trials had received, and likely failed on, prior therapies, including [redacted] of participants having prior experience with therapies usually only given to patients with severe AA. The EAG's clinical experts highlighted that the level of treatment experience in the BRAVE-AA trials likely exceeds that seen in UK clinical practice for severe AA, both in terms of the percentage of patients receiving therapies such as contact immunotherapy and systemic immunosuppressants, but also that patients in the BRAVE-AA trials received therapies not currently used in the NHS, such as phototherapy, cryotherapy and platelet-rich-plasma injections. This may mean that the BRAVE-AA trials may underestimate of the effectiveness of baricitinib in the first-line setting, as patients who had succeeded on prior therapies would not have entered the trial. However, the EAG notes that the mode of action of baricitinib is different to the current therapies used to treat (severe) AA, and as such the magnitude of any effect of treatment experience is unknown.

Underestimation of the treatment effectiveness of baricitinib in the first-line setting in UK clinical practice is also likely due to the BRAVE-AA trial participants having varying, and often long, lengths of current AA episode at baseline, with mean durations ≥3.5 years for all arms. In contrast, patients presenting with newly diagnosed severe AAA in the first-line setting are likely to have shorter durations of AA episodes, and therefore a larger probability of treatment response.

Regarding the company's positioning of baricitinib both in the first-line setting and later-line settings for patients who have failed on previous treatment, the EAG considers that:

- The BRAVE-AA trial population is most similar to the prevalent population in clinical practice who would be eligible to receive baricitinib at the point of approval, i.e. a later-line treatment experienced population. The treatment effectiveness of baricitinib in this later-line population might be overestimated by the BRAVE-AA trial data due to the exclusion of patients with current AA episodes >8 years and who had showed no sign of previous regrowth and male patients >60 years and female patients >70 years;
- The BRAVE-AA trial data may underestimate the effectiveness of baricitinib in the first-line population, because of high rate of prior treatment with agents usually only given to patients with severe AA, and the presence of patients with relatively long baseline AA episode durations in the BRAVE-AA trial populations. In the company's Adelphi DSP study, [redacted] of severe/very severe AA patients were treatment experienced[13, 14], including both therapies given at milder stages of disease (e.g. topical corticosteroids), but also those primarily given to patients with severe disease (topical immunotherapy, systemic immunosuppressants and systemic steroids).

2.3.2 *Intervention*

The intervention under consideration is oral baricitinib 4 mg and matches the final scope issued by NICE.[11] Baricitinib is an oral JAK inhibitor that is expected to receive marketing authorisation for treating severe AA in adults in [redacted]. Further details of baricitinib for AA, including the method and administration and dosing can be found in Section B.1.2 of the CS. Baricitinib has previously been recommended in certain populations for treating moderate to severe atopic dermatitis in TA681,[15] and for treating severe rheumatoid arthritis in TA466.[16]

2.3.3 Comparators

The comparator in the final scope issued by NICE was, "established clinical management without baricitinib". To this, the company added, "which may include supportive care", which the EAG agrees is in-line with the final scope issued by NICE.[11] The EAG has discussed the current pathway of care for severe AA in Section 2.2.1 and notes that:

- There is no clear standard clinical management for severe AA and no single most suitable comparator therapy;
- Many severe AA patients may opt for or will end up receiving no active treatment, with only occasional, if any, follow-up, especially after disappointing results from available therapies.

The EAG is concerned with how the company defines of "Watch and wait" as established clinical management for adults with severe AA, which the company outline as involving continued monitoring. The EAG agrees with the company that the no active treatment component of "Watch and wait" is a common management strategy used for adults with severe AA, however the EAG's clinical experts highlighted that this would not require intensive follow-up as they would be discharged from care, or receive only occasional follow-up. In addition, the EAG's clinical experts highlighted how access to psychological support, while needed, is in practice minimal due to resource constraints.

Of the best supportive care options, the EAG considers there to be three candidates to provide comparative cost-effectiveness data against baricitinib in this submission:

- A comparison with DPCP, which the EAG's clinical experts highlighted might be the closest in effectiveness to baricitinib and the only active treatment recommended by the BAD Guidelines for treating severe AA in adults.[9] However, the EAG considers there to be no valid means of performing valid comparison between DPCP and baricitinib 4 mg in adult severe AA patients (see Section 3.4), and further notes that: i) DPCP is only available to a minority of patients with no equitable access and, ii) many patients discontinue treatment, and some will suffer strong adverse reactions to DPCP, iii) some patients will have already received or been eligible for DPCP for milder disease;
- A comparison with the systemic corticosteroids or systemic immunosuppressants currently used to treat severe AA, each at relatively low frequency. Again, the EAG does not consider there to be appropriate data available to perform a valid comparison between all or any of

these therapies and baricitinib 4 mg in adult severe AA patients (see Section 3.4), nor does the EAG consider any to be established standard of care. The EAG also notes that some of these treatments may have been given to patients when they had mild or moderate disease, and, because of the limited effectiveness of these treatments, the placebo group from the BRAVE-AA trials (i.e. no active treatment) may provide a reasonable approximation for the treatment effect;

- A comparison with no active treatment with discharge from care. The EAG considers this to be both the most appropriate comparison for treatment-experienced patients and for those newly diagnosed with severe AA who opt not to receive further treatment. Moreover, it is the only comparison for which robust comparative data are available with baricitinib 4 mg through the BRAVE-AA trials.

The EAG considers the comparison with no active treatment to be the most relevant for the current submission, as it reflects a viable treatment option across the prevalent and incident populations of adults with severe AA and is the only comparison for which high-quality data comparative effectiveness data are available. As outlined in Section 4.2.5.2, the EAG considers the placebo arm of the BRAVE-AA trials to provide a reasonable estimate of the treatment response a patient with no active treatment would receive.

The EAG considers there to be a distinction between the prevalent and incident populations of adults with severe AA when considering the most appropriate comparator for the appraisal, given the treatment experience of these patients may differ:

- Patients in the prevalent population are likely to have explored all treatment options available to them. Most of these patients will have opted for no active treatment and may manage their severe AA with head shaving or wigs. For this population — the population that would be treated at the point of approval in UK clinical practice — the EAG considers no active treatment and discharge from care to be the appropriate comparator;
- Patients in incident population will be less treatment-experienced than those in the prevalent population, and only a minority will have prior experience with topical immunotherapy (e.g. DPCP), systemic immunosuppressants or systemic corticosteroids. Upon progression to severe AA, they may opt to trial one or more of these therapies, assuming their dermatologist offers it to them. The EAG does not consider any specific systemic immunosuppressants or systemic corticosteroids to be an established standard of

care for these patients, and considers no active treatment to be a realistic endpoint for these patients. The EAG notes the absence of data available to permit meaningful comparisons of these aforementioned treatments with baricitinib 4 mg (see Section 3.4). The EAG agrees with the company that these therapies have very limited effectiveness for severe disease. The EAG highlights a lack of a: i) clear standard of care, ii) robust data on treatment patterns, and iii) robust comparative effectiveness of active treatments with baricitinib. Hence, the EAG considers a direct comparison between baricitinib and systemic immunosuppressants and/or systemic corticosteroids unlikely to adequately capture established clinical management in the UK until more data are available to demonstrate: i) that these therapies are frequently used by, and accessible to, most patients and, ii) to provide robust comparative effectiveness data with baricitinib. In lieu of such data, the EAG considers no active treatment and discharge from care is an acceptable comparator for adults newly diagnosed with severe AA. The EAG notes that at the point of approval this population will be small, but that baricitinib would become a preferred first-line therapy for newly diagnosed severe AA, if approved.

2.3.4 Subgroups

No subgroups were clearly identifiable from the NICE final scope, which stated, "If the evidence allows, the following subgroups based on severity and type of alopecia areata will be considered". The company provided subgroup analyses and a scenario analysis based on the baseline severity of alopecia areata (severe disease, SALT 50-95 and very severe disease, SALT 95-100). No scenario analyses were presented based on the type of alopecia areata, however, the EAG notes that all alopecia totalis/universalis patients would be included in the SALT 95-100 subgroup.

Subgroup efficacy analyses based on baseline length of current AA episode were available in the clinical study reports (CSRs), and, in response to clarification question B2, the company provided an economic scenario analysis for subgroups of ≤4 years and >4 years (Section 5.1.2.2). In response to clarification question A14, the company also provided a subgroup analysis based on atopic background status, a factor highlighted by the EAG's clinical expert as a potentially meaningful subgroup.

2.3.5 *Special conditions*

No special conditions were identified in the NICE final scope or by the company. The EAG's clinical experts highlighted how health related quality of life deficits may vary for different reasons across cultures – in line with the issues raised in the equality impact assessment.[17] For example, in some cultures loss of beard hair can be an important issue. However, the EAG's clinical experts highlighted that the negative consequences of AA can and do extend to patients of all demographics.

3 Clinical effectiveness

3.1 Critique of the methods review

The company conducted a systematic literature review (SLR) to identify randomised controlled trials (RCTs) and observational studies providing clinical efficacy and safety data for baricitinib for the treatment of alopecia areata (AA) comparators and supportive care therapies. The Evidence Assessment Group (EAG) notes that the original SLR was conducted in July 2021 and that it was subsequently updated in February 2022.

A total of 45 studies from 47 records were included from the SLR, including 12 RCTs. An overview of the methods used by the company for the SLR, together with the EAG's critique of the appropriateness of these methods, is presented in Table 9. In summary, the EAG considers the methods applied by the company to be adequate and likely to have identified most of the clinical evidence of relevance to the decision problem. One study of baricitinib was included from the SLR, King *et al.* 2021,[18] that reported on the Phase II portion of BRAVE-AA1. The company provided additional data on the Phase III portion of BRAVE-AA1 and from BRAVE-AA2 in the submission, as the primary publication for BRAVE-AA1 and BRAVE-AA2 was published after the SLR update search date.[12]

Table 9. Summary of EAG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	EAG assessment of robustness of methods
Data sources	Appendix D1.1.	**The EAG considers the sources and dates searched to be appropriate.** Databases searched: Embase, MEDLINE In-Process and the Cochrane Library (CENTRAL and CDSR). Additional sources: Hand-searching of conference proceedings (published in 2019 to 2021) and clinical trial registers. Latest search update: 4 February 2022.
Search strategies	Appendix D1.1	**The EAG is satisfied that the searches have identified all evidence relevant to the decision problem.** Search strategies for the literature review combined comprehensive terms for the population, interventions and study designs, using free-text and medical subject headings.
Inclusion criteria	Appendix D1.2	**The EAG considers it likely that no relevant evidence was excluded, although the EAG notes that as young adolescent patients are often treated as adults in AA, many studies that**

		included a small number of patients <16 years were excluded. The inclusion criteria of the SLR were in line with the NICE final scope, except for severity and age. For age, studies reporting on patients ≥16 years were included. For severity, studies containing patients with moderate AA were included throughout the SLR, meaning the studies identified contain a wider population than specified in the NICE final scope.[11] Full reference details are available in the CS Appendix for included studies and excluded studies at full text review.
Screening and data extraction	Appendix D.1.2 and D.1.3	**The EAG considers the methods for screening and data extraction to be robust.** Two reviewers independently screened titles and abstracts, and subsequently studies selected for full text appraisal, against predefined criteria, with a third reviewer consulted when consensus could not be reached. Results of the literature screening processes were summarised in PRISMA diagrams. Conference proceedings and clinical trial registries were searched by a single reviewer and checked by a second reviewer. Data extraction was carried out by one reviewer, with a second researcher independently quality checking the extracted data.
Tool for quality assessment of included study or studies	B.2.5 & Appendix D.1.4 and D.3	**The EAG agrees with the company's choice of quality assessment tool for assessing BRAVE-AA1 and BRAVE-AA2.** The company used the using the Appraisal of RCT checklist by Cochrane[19] for the quality assessment of the RCTs included in the SLR. The quality of the included observational studies was assessed using the quality assessment tool developed by the York University CRD.[20]

Abbreviations: AA: alopecia areata; CENTRAL: Cochrane Central Register of Controlled Trials; CDSR: Cochrane Database of Systematic Reviews; CS: company submission; EAG: Evidence Assessment Group; NICE: National Institute for Health and Care Excellence; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT: randomised controlled trial; SLR: systematic literature review.

3.1.1 SLR reporting quality

The EAG notes there were numerous reporting errors in the conduct of the SLR that raised concerns about the overall quality of the SLR, but that the company was able to clarify these adequately (clarification questions A16, A17, A21 and A22).

3.1.2 Age eligibility criterion

The EAG notes that studies reporting on patients ≥16 years were included in the SLR (Clarification question A22), a wider inclusion criterion than the BRAVE-AA trial inclusion criterion and the provisional marketing authorisation of baricitinib. The EAG considers including such studies to be

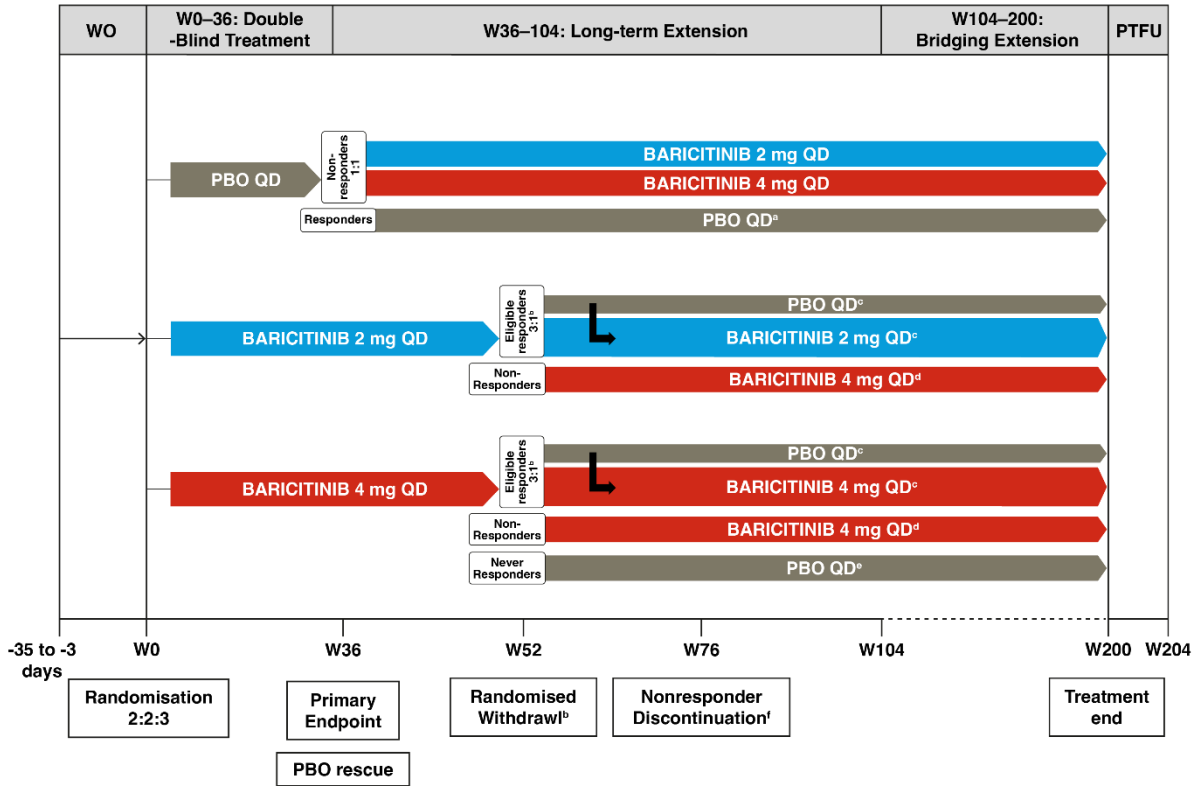
reasonable, as the EAG's clinical experts noted it was not uncommon for young adolescent AA patients be treated as adults, and that many studies identified in the SLR were single-centre reviews of all patients in a centre. Given the large number of records excluded from the SLR at the full text stage due to containing some patients <16 years (190 records vs 47 ultimately extracted in the SLR), the EAG believes that, if anything, the age edibility criterion of the SLR was too restrictive, and some relevant studies containing a small number of patients <16 years may have been excluded. Nevertheless, the EAG's clinical experts noted the absence of high-quality placebo controlled RCT data for treating severe AA, something also noted by other systematic reviews of the field.[21] Hence, the EAG considers it unlikely that any key studies have been excluded from the SLR due to containing some paediatric patients.

3.2 Critique of trials of the technology of interest

In this section, the EAG critiques the BRAVE-AA1 and BRAVE-AA2 trials that provide the key clinical effectiveness data used in the cost-effectiveness analysis of the CS.[12] Both BRAVE-AA1 and BRAVE-AA2 are international, double-blind, randomised, placebo-controlled trials comparing baricitinib 2 mg daily and baricitinib 4 mg daily with placebo. While BRAVE-AA1 is an adaptive Phase 2/3 trial, only the Phase 3 data are used in the submission and considered hereafter. As baricitinib 4 mg is the dosage under consideration in this STA, the data from the baricitinib 4 mg arms and placebo arms of BRAVE-AA1 and BRAVE-AA2 will be focused on in this critique. Supporting information will be cited from the baricitinib 2 mg data where appropriate, and the EAG notes that the results from the baricitinib 2 mg arms were consistent with the results of the baricitinib 4 mg throughout the results of the trial, albeit with a lower magnitude of benefit over placebo throughout.

The primary outcome of the BRAVE-AA trials was achieving an absolute SALT ≤20 at Week 36, i.e., at the end of the double-blind treatment stage. However, additional data up to Week 72 were provided for a randomised withdrawal sub study (BRAVE-AA1) and a randomised down-titration sub study (BRAVE-AA2). The EAG notes that because placebo non-responders at Week 36 were eligible for rescue therapy, robust comparative data between baricitinib and placebo are only available up to Week 36. The design of BRAVE-AA1 and BRAVE-AA2 are reproduced in Figure 1 and Figure 2.

Figure 1. Study design of BRAVE-AA1 (Reproduced from Figure 1, Clarification Response)

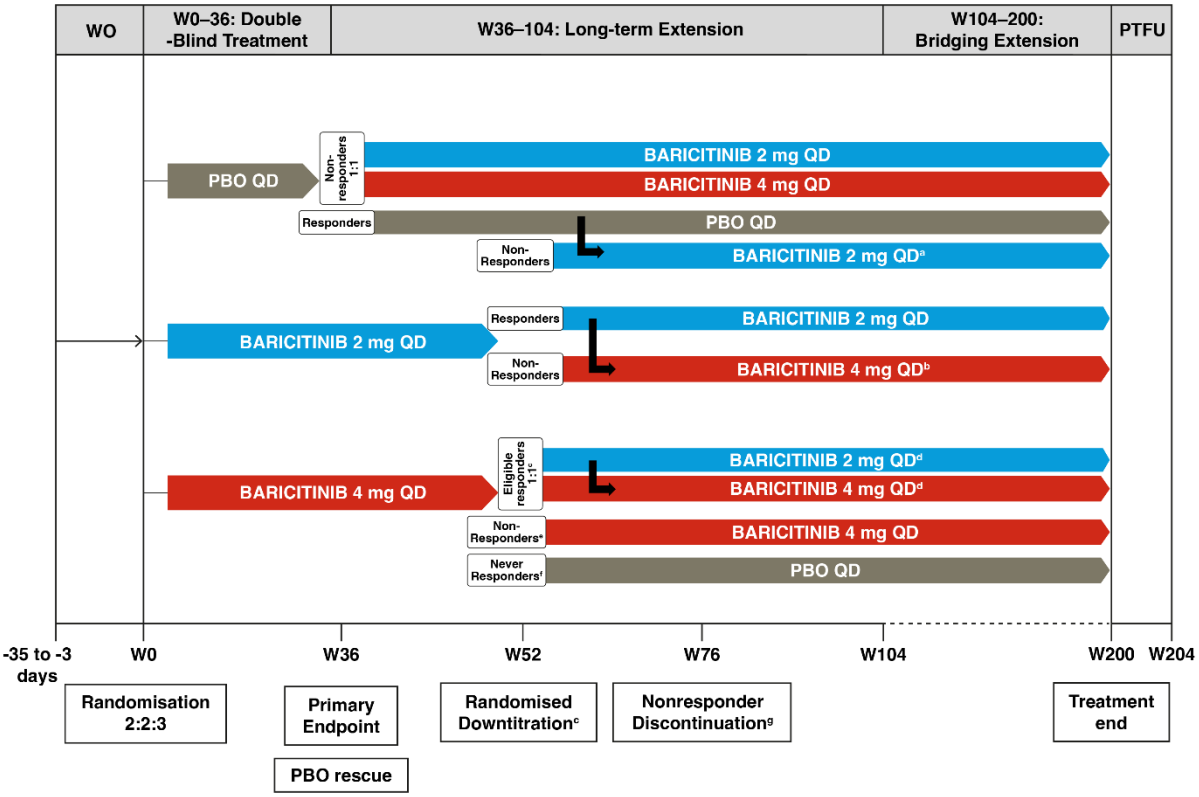


Footnotes: [a] Placebo responders stayed on placebo for remainder of the trial, even if relapse was observed later. [b] Patients with SALT ≤20 who stayed on the same dose of baricitinib from week 0 were randomised to stay on current baricitinib dose, or transitioned to placebo.[c] Responders participating in randomised withdrawal who experienced >20-point absolute worsening in total SALT score after week 52 were retreated with baricitinib dose to which they were originally randomised if they were randomised to placebo at week 52, OR continued to receive same dose of baricitinib if they were randomised to remain on baricitinib at week 52. [d] Non-responders at week 52 were rescued to baricitinib 4 mg if receiving baricitinib 2 mg from baseline, OR remained on baricitinib 4 mg if they were in the 4-mg group and achieved SALT ≤20 before week 52. [e] Never responders (never achieved SALT ≤20 by week 52 despite being in the baricitinib 4-mg group from baseline and had not experienced a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss) were automatically transitioned to placebo.[f] Non-responders at week 52 AND week 76 were automatically discontinued at week 76 unless they had a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss.

Abbreviations: EB: eyebrow; EL: eyelash; PTFU: post-trial follow up; QD: once daily; W: week.

Source: BRAVE-AA1 Clinical Study Report.

Figure 2. Study design of BRAVE-AA2 (Reproduced from Figure 2, Clarification Response)



Footnotes: [a] Placebo-treated patients not eligible for rescue to baricitinib at week 36 (due to spontaneous remission) were rescued to baricitinib if they were non-responders at week 52, OR if they experienced loss of treatment benefit after week 52. [b] Patients randomised to baricitinib 2 mg at week 0 were rescued to the 4-mg dose if they were non-responders at week 52, OR were responders at week 52 but experienced a >20-point worsening in SALT score after week 52. [c] Responders in the baricitinib 4-mg group (SALT ≤20 who stayed on 4 mg from week 0) were randomised to either stay on 4 mg OR transition to 2 mg.[d] Responders participating in the randomised down-titration who experienced a loss of treatment benefit after week 52 were re-treated with baricitinib 4 mg if they were randomised to the 2-mg dose at week 52, OR continued to receive baricitinib 4 mg if they randomised to remain on the 4-mg dose at week 52. [e] At week 52, non-responders (SALT >20) in the baricitinib 4-mg group since baseline who achieved SALT ≤20 before week 52 remained on 4 mg. [f] Never responders (never achieved SALT ≤20 by week 52 despite being in the baricitinib 4-mg group from baseline and had not experienced a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss) were automatically transitioned to placebo. [g] Non-responders at week 52 AND week 76 were automatically discontinued at week 76 unless they had a ≥2-point improvement from baseline in ClinRO measures for EB or EL hair loss.

Abbreviations: PTFU: post-trial follow up; QD: once daily; W: week.

Source: BRAVE-AA2 Clinical Study Report.

The EAG's clinical experts agreed with the company that the design of BRAVE-AA1 and BRAVE-AA2 was sufficiently similar to justify pooling the data, and the EAG considers this to be appropriate. The EAG's quality assessment of BRAVE-AA1 (Phase III portion) and BRAVE-AA2 is provided in Table 10. Overall, the EAG considers BRAVE-AA1 and BRAVE-AA2 to be high quality RCTs, in-line with the quality assessment provided by the company in Table 14 of the CS.

Table 10. The EAG's quality assessment of BRAVE-AA1 and BRAVE-AA2

Aspect of trial design or conduct	Section of CS or supporting information where details are reported	EAG's critique	
		BRAVE-AA1	BRAVE-AA2
Randomisation	CS B.2.3.1	**Appropriate** Participants were randomised 2:2:3 to placebo, 2 mg baricitinib or 4 mg baricitinib using an interactive web response system (IWRS).	
Concealment of treatment allocation	CS B.2.3.1 BRAVE-AA1 Protocol Section 5.1 BRAVE-AA2 CSR Table 4.1	**Appropriate** Treatment allocation was concealed by use of an IWRS	
Eligibility criteria	CS Table 6	**Some concerns** The eligibility criteria of the trials excluded certain adults with severe alopecia areata who may be less likely to respond to treatment, namely: • Those with current AA episodes at baseline >8 years with no episodes of regrowth and; • Males >60 years and females >70 years.	
Blinding	CS B.2.3.1	**Appropriate** Up to Week 36 the trials were double blind, with the patients, investigators and study team blinded to baricitinib or placebo assignment.	
Baseline characteristics	CS B.2.3.3	**Appropriate** Key baseline characteristics were balanced between the baricitinib 4 mg and placebo arms in both BRAVE-AA1 and BRAVE AA-2 (Table 11, Appendix Table 44). While UK centres were not included in either trial, the EAG's clinical experts did not expect any difference in demographic variables between the included centres and UK centres to be potential treatment effect modifiers.	
Statistical analysis			
Sample size and power	BRAVE-AA1 Protocol Section 10.1 BRAVE-AA2 CSR Table 4.1	**Appropriate** Sample sizes of approximately 625 patients (BRAVE-AA1) and 476 patients (BRAVE-AA2) were targeted to provide over (BRAVE-AA1) or approximately (BRAVE-AA2) 90% power to differences between baricitinib and placebo using the graphical testing procedure and the	

		following assumed response rates: 30% for baricitinib 4-mg, 20% for baricitinib 2-mg, and 5% for placebo. The actual sample sizes in the full analysis sets exceed the target sizes: 654 (BRAVE-AA1) and 546 (BRAVE-AA2). While the response rates in the power calculations were not clinically justified, these sample sizes are likely to detect most meaningful benefits of baricitinib 4 mg over placebo.
Handling of missing data	CS B.2.4.3	**Some concerns** Missing SALT data for [redacted] of patients at Week 36. Missing data were imputed using: i) non-responder imputation for categorical endpoints and, ii) modified last observation carried forward for continuous endpoints. In comparison to multiple imputation, these methods provide lower power to detect treatment effects. The EAG does not consider the analyses in the CS to bias results in favour of baricitinib because the analysis in the primary publication,[12] that used multiple imputation, estimated a larger treatment effect for baricitinib 4 mg than in the CS. As such, the company's handling of missing data was conservative for SALT analyses.
Outcome assessment	BRAVE-AA1 Protocol Sections 9.1.3.3. and 9.1.5 BRAVE-AA1 Protocol Sections 9.1.3.3. and 9.1.5	**Appropriate** Measurement of SALT score, i.e., the proportion of the scalp without hair coverage, was conducted by blinded investigators and is relatively objective. EQ-5D data were collected by self-report from blinded participants.
Analysis for estimate of effect	CS B.2.6.1	**Appropriate** The company's primary analyses and data used in the economic model are responder analyses (SALT ≤20, $SALT_{50}$ and $SALT_{75}$). While these may have lower power than analyses using continuous outcome variables, the EAG's clinical experts considered SALT ≤20 and $SALT_{75}$ to be clinically meaningful outcomes, and the results from the responder analyses were consistent with the continuous change from baseline analyses.

Abbreviations: AA; alopecia areata; CS: company submission; CSR: clinical study report; EAG: evidence assessment group; EQ-5D: EuroQol-5 Dimension; IWRS: interactive web response system; SALT: severity of alopecia tool

3.2.1 Randomisation

Participants in BRAVE-AA1 and BRAVE-AA2 were randomised 2:2:3 to placebo, 2 mg baricitinib or 4 mg baricitinib at Visit 2. Randomisation was stratified based on geographic region (North America and Japan), and duration of current AA episode at baseline (<4 years versus ≥4 years). Errors in data

entry led to [redacted] ([redacted]) patients in BRAVE-AA1 and [redacted] ([redacted]) patients in BRAVE-AA2 patients) having incorrect duration of current AA episode data at baseline. This led to errors in stratified randomisation. However, given the overall balance in baseline characteristics between the baricitinib 4 mg arm and the placebo arm (see Section 3.2.3 and Table 11), the EAG did not determine these errors likely to bias the results of the trials.

3.2.2 Eligibility criteria

The EAG has outlined in Section 2.3.1 how it considers the eligibility criteria of BRAVE-AA1 and BRAVEE-AA2 to most closely reflect the later-line population of the positioning of baricitinib. Due to the exclusion of older patients and those with current AA episodes >8 years and who had showed no sign of previous regrowth, the EAG suggests that the BRAVE-AA trials may slightly overestimate the treatment effectiveness of baricitinib in UK clinical practice for the later line population. In contrast, because trial participants were not excluded based on prior therapies, including contact immunotherapy, the EAG considers the BRAVE-AA trial data to likely underestimate treatment effectiveness in the first-line population.

The EAG also notes that participants in BRAVE-AA1 and BRAVE-AA2 had very few permitted concomitant medicines, and only [redacted] of patients had any that may target AA. While this is a reasonable for the patients receiving baricitinib, the EAG's clinical experts considered that many severe AA patients not receiving baricitinib would be treated after 6 months. While the basket of currently used therapies for severe AA have low likelihoods of success, some may still be more effective than placebo, which the EAG discusses this further in Section 3.4.

3.2.3 Participant characteristics

The baseline characteristics of patients in the placebo and baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2 are presented in Section B.2.3.3 in the CS. The EAG's clinical experts stated that the duration of current AA episode, presence of an atopic background and baseline SALT score might predict the likelihood of hair regrowth, and these data are presented in Table 11 alongside baseline health-related quality of life (HRQoL) data. These characteristics were largely balanced between the placebo and baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2. However, the BRAVE-AA2 placebo arm had a higher mean duration of current AA episode at baseline (4.68 years) than the baricitinib 4 mg arm (3.94 years), which might equate to a lower chance of response in the BRAVE-AA2 placebo arm than all other arms.

Table 11. Baseline characteristics of patients in the placebo and baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2 (adapted from Tables 9 and 10 of the CS)

Characteristic	BRAVE–AA1		BRAVE–AA2	
	Placebo (N=189)	baricitinib 4 mg (N=281)	Placebo (N=156)	baricitinib 4 mg (N=234)
Baseline characteristic highlighted by EAG clinical experts				
Mean (SD) duration of current AA episode, years	3.53 (3.65)	3.46 (3.37)	4.68 (5.490)	3.94 (3.353)
Atopic background, n (%)	73 (38.6)	97 (34.5)	67 (42.9)	87 (37.2)
Duration of current AA episode, n (%)				
<4 years	134 (70.9)	189 (67.3)	94 (60.3)	140 (59.8)
≥4 years	55 (29.1)	92 (32.7)	62 (39.7)	94 (40.2)
Mean (SD) SALT score	84.7 (17.82)	85.3 (18.18)	85.0 (17.79)	84.8 (18.08)
SALT category, n (%)				
Severe (SALT 50–94)	92 (48.7)	133 (47.3)	74 (47.7)	115 (49.1)
Very severe (SALT 95–100)	97 (51.3)	148 (52.7)	81 (52.3)	119 (50.9)
HRQoL baseline characteristics				
Mean (SD) Skindex–16 AA baseline domain scores				
Emotions	[redacted]	[redacted]	[redacted]	[redacted]
Functioning	[redacted]	[redacted]	[redacted]	[redacted]
Symptoms	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD) HADS total score				
HADS-Anxiety	6.7 (3.92)	6.1 (3.80)	5.9 (4.01)	6.4 (3.95)
HADS-Depression	4.0 (3.15)	4.0 (3.39)	3.7 (3.46)	3.8 (3.49)
EQ-5D-5L health state index	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D-5L VAS score	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 9, Table 10, Table 26, and Table 27

Abbreviations: AA: alopecia areata; CS: company submission; HADS: hospital anxiety and depression score; PRO: patient reported outcome; SALT: severity of alopecia tool; SD: standard deviation

The EAG's clinical experts noted that several of the baseline characteristics of patients in the BRAVE AA trials indicate that the trial participants had particularly severe and difficult to treat alopecia areata: a mean disease duration from the first onset of AA diagnosis of 12.2 years; a mean episode duration of 3.9 years; >50% had SALT 95-100 and around [redacted] had alopecia universalis. Such a

severity of patients may mean the BRAVE-AA trials slightly underestimate treatment effectiveness relative to UK practice, although the absence of high-quality demographic data around severe AA in the UK makes this uncertain.

BRAVE-AA1 and BRAVE-AA2 were international trials with no UK centres. BRAVE-AA1 recruited most patients from the USA (54.7% of patients) and South Korea (37.8% of patients), and BRAVE-AA2 recruited most patients from the USA (34.8% of patients) and from Asian sites (26.9% of patients). Despite certain demographics, such as race, differing systematically from UK practice, the EAG's clinical experts did not believe treatment efficacy would differ substantially between geographic region or across races, and agreed that other baseline characteristics were broadly similar to those that would be seen in UK practice, which are presented in Appendix Table 44.

As highlighted in Section 2.3, around [redacted] of participants in the BRAVE-AA trials had received prior therapies for AA, and over [redacted] had received systemic immunosuppressants/immunomodulators that the EAG's clinical experts noted would only be given for severe AA in UK practice. These data are presented in Table 12. The EAG's clinical experts stated that while some of these prior treatments would be common in NHS patients, such as topical corticosteroids from their GPs, others are not widely available or used in the UK, such as cryotherapy and phototherapy, and some, such as cyclosporin and topical immunotherapy are used in the UK but at a lower prevalence than was reported in the BRAVE-AA trials.

Table 12. Prior therapies received by patients in the placebo and baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2 (Adapted from CS Table 11 and CS Table 12)

	BRAVE-AA1		BRAVE-AA2	
	Placebo (N=189)	Baricitinib 4 mg (N= 281)	Placebo (N=156)	Baricitinib 4 mg (N= 234)
Prior therapy, n (%)	173 (91.5)	247 (87.9)	149 (95.5)	211 (90.2)
Topical therapy, n (%)	108 (57.1)	173 (61.6)	98 (62.8)	148 (63.2)
Topical IMT, n (%)	45 (23.8)	84 (29.9)	41 (26.3)	63 (26.9)
Intralesional therapy, n (%)	101 (53.4)	152 (54.1)	88 (56.4)	104 (44.4)
Systemic agents, n (%)				
Immunosuppressant/ immunomodulator	101 (53.4)	138 (49.1)	97 (62.2)	124 (53.0)
Corticosteroids	68 (36.0)	103 (36.7)	77 (49.4)	102 (43.6)
JAK inhibitor[a]	12 (6.3)	15 (5.3)	9 (5.8)	10 (4.3)
Others	57 (30.2)	88 (31.3)	54 (34.6)	52 (22.2)

Cyclosporin	46 (24.3)	69 (24.6)	27 (17.3)	27 (11.5)
Methotrexate	15 (7.9)	28 (10.0)	27 (17.3)	31 (13.2)
Other systemic (non-immunosuppressant), n (%)	17 (9.0)	28 (10.0)	15 (9.6)	18 (7.7)
Phototherapy, n (%)	23 (12.2)	54 (19.2)	28 (17.9)	37 (15.8)
Procedures, n (%)	30 (15.9)	65 (23.1)	35 (22.4)	47 (20.1)

Source: CS Table 11 and Table 12

Abbreviations: AA: alopecia areata; CI: confidence interval; IMT: immunotherapy; JAK: janus kinase

[a]Patients with prior inadequate response to JAK inhibitors were excluded from the trial, although no patients failed screening for this reason (inclusion/exclusion criteria #9, BRAVE-AA1 CSR and BRAVE-AA2 CSR)

3.2.4 *Outcome assessment*

The key clinical effectiveness outcome, SALT score, was assessed by investigators blinded to the treatment a patient was receiving. SALT score measurement is a relatively objective procedure in which the assessor compares each quarter of the scalp with a chart detailing how much hair is missing. The EAG's clinical experts noted how SALT measurement can be quite imprecise, and as such cautioned against using strict absolute thresholds, such as SALT ≤20 or $SALT_{75}$ to determine whether treatment should be continued, i.e., they might be unlikely to recommend a patient achieving a SALT score 21 to discontinue treatment.

The key HRQoL data were collected by self-report from blinded participants and the EAG does not have concerns about the validity of this data collection. Section 3.3.5 and Section 4.2.8.1 contains a critique of the company's claim of a lack of content validity for the EQ-5D-5L measure for severe AA.

3.3 Critique of the clinical effectiveness analysis

The SALT score was the primary focus of the company's clinical effectiveness analysis, and the absolute measure, SALT ≤20 was the primary outcome of the BRAVE-AA trials at Week 36. In the base case analysis, the company focuses on relative rather than absolute measures of hair regrowth. The definitions of the SALT outcomes used throughout the CS are presented in Table 13.

Table 13. Example definition of SALT outcomes used in the CS.

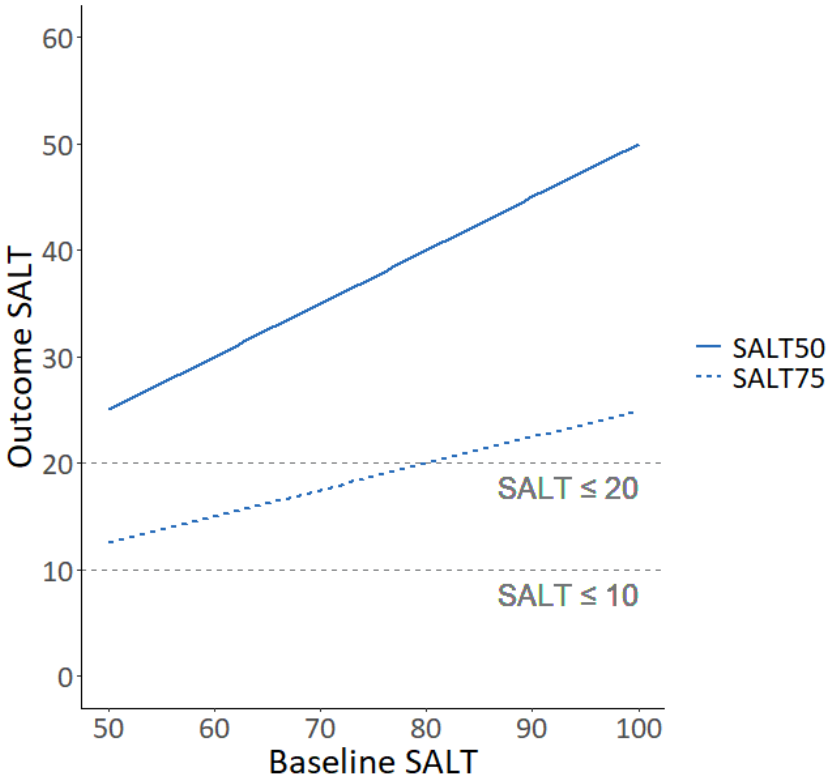
SALT Outcome	Definition
Absolute measures	
SALT ≤10, SALT ≤20	A SALT score of less than or equal to 10 (SALT ≤10) or 20 (SALT ≤20) at the timepoint.
Relative measures	

$SALT_{75}$, $SALT_{50}$	A 75% ($SALT_{75}$) or a 50% ($SALT_{50}$) reduction from baseline in an individual's SALT score at the timepoint.

Abbreviations: CS: company submission; SALT: severity of alopecia tool

Figure 3 displays the relationship between SALT ≤10, SALT ≤20, $SALT_{75}$ and $SALT_{50}$. SALT ≤10 is the most stringent criteria that the EAG's clinical experts agreed would be a strong clinically meaningful outcome for nearly all patients, alleviating the need for wig use. The EAG's clinical experts also agreed that SALT ≤20 and $SALT_{75}$ would be clinically meaningful outcomes and are near equivalent for severe AA patients. $SALT_{50}$, however, is a much less stringent criterion and the EAG's clinical experts doubted whether this would be a meaningful outcome for many patients, who would still have a large degree of hair loss and would likely still opt for wigs and/or head shaving.

Figure 3. The relationship between absolute and relative SALT thresholds used to define responders in the CS.



In the following section, the EAG critiques the clinical effectiveness analysis of the company in the CS. In general, the EAG considers the statistical comparisons between baricitinib 4 mg and placebo to be robust and to present a clear benefit of baricitinib 4 mg over placebo. The results of the additional analyses requested by the EAG in clarification question A11 were consistent with the

analyses presented in the CS, suggesting the company results are robust to different analytical approaches.

Throughout the CS, five different analysis sets were reported on, which are detailed in Table 13 of the CS. The clinical effectiveness used in the CS analyses were conducted in the full analysis sets (FAS) of BRAVE-AA1 and BRAVE-AA2, which comprised all patients randomised at baseline, or the pooled Week 36 efficacy population, which combines the BRAVE-AA1 FAS and the BRAVE-AA2 FAS. The EAG considers these the appropriate analysis sets to use for the clinical efficacy analysis. The safety analysis set comprised all patients randomised who receive at least one dose of study intervention and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first post-baseline visit.

3.3.1 SALT responder outcomes

The proportion of patients achieving a SALT ≤20, SALT ≤10, $SALT_{50}$ and $SALT_{75}$ response in BRAVE-AA1 and BRAVE-AA2 at Week 36 are presented in Table 14. Pooled across BRAVE-AA1 and BRAVE-AA2 baricitinib 4 mg arms, the $SALT_{50}$ response rate was [redacted] and $SALT_{75}$ response rate was [redacted]. The SALT ≤20 response rate was similar to $SALT_{75}$ at [redacted], and the SALT ≤10 response rate was [redacted]. For all statistical comparisons, baricitinib 4 mg had a significantly higher Week 36 response rate than placebo, with all p-values [redacted], and all odds ratios [redacted] (lowest 95% CI: [redacted]). Detailed tables of statistical results for each outcome measure are provided in Table 45, Table 46 and Table 47 of the Appendix.

Table 14. SALT response rates at Week 36 in BRAVE-AA1 and BRAVE-AA2 (FAS)

Week 36	BRAVE-AA1		BRAVE-AA2	
	Placebo (N=189)	Baricitinib 4 mg (N=281)	Placebo (N=189)	Baricitinib 4 mg (N=281)
SALT ≤20, % (95% CI)	5.3 (2.9 to 9.5)	35.2 (29.9 to 41.0)	2.6 (1.0 to 6.4)	32.5 (26.8 to 38.7)
SALT ≤10, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
$SALT_{50}$, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
$SALT_{75}$, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Tables 15, 18, 19 and 22

Abbreviations: AA: alopecia areata; CI: confidence interval; FAS: full analysis set; SALT: severity of alopecia tool

The EAG notes that while response rates were significantly higher for baricitinib 4 mg than placebo for all outcomes, only around [redacted] of patients achieved SALT ≤20 or $SALT_{75}$ by Week 36. However, the EAG also notes that over [redacted] of SALT ≤20 responders were also SALT ≤10 responders at Week 36, suggesting that many responders at the SALT ≤20 or $SALT_{75}$ thresholds had a large and clinically meaningful response. The EAG further notes that response rate results were replicated successfully between two large, multi-site, high-quality international trials. In addition, the results of the baricitinib 2 mg arms consistent with the baricitinib 4 mg arms, with the expected lower magnitude for the lower dose. Overall, the EAG considers the trials to have strong internal validity and that the SALT responder results are likely robust within the inclusion and exclusion criteria of the BRAVE-AA trials.

In addition to the dichotomous responder-based analysis, the company provided some data on the mean change from baseline in SALT score at Week 36 in Table 16 and Table 17 of the CS. The mean (SE) change from baseline in SALT score for baricitinib 4 mg was -45.79 (2.66) in BRAVE-AA1 and -47.45 (2.23) in BRAVE-AA2, compared to -8.13 (3.10) and -2.96 (2.72) in the BRAVE-AA1 and BRAVE-AA2 placebo arms, respectively. The EAG considers the result from the change from baseline analyses to be consistent with the responder-based analysis.

3.3.1.1 Week 52 and Week 76 data

The EAG's clinical experts agreed that around Week 36 is a reasonable time to assess the effect of a JAK inhibitor on patients. At Week 36, the double-blind treatment phase of BRAVE-AA1 and BRAVE-AA2 ended and non-responders in the placebo arm were randomised to baricitinib 2 mg or 4 mg rescue treatment. In contrast, patients who started on one of the baricitinib arms continued on this arm until at least Week 52. The pooled SALT ≤20 response data for these patients by visit until Week 52 are presented in Figure 4. At Week 52, [redacted] of patients in the baricitinib 4 mg had achieved SALT ≤20, an increase of [redacted] from Week 36 ([redacted] had achieved SALT ≤20 at Week 36).

Figure 4. Proportion of patients achieving SALT≤20 through Week 52 in the BRAVE-AA studies (pooled Week 52 efficacy population; primary censoring [NRI]). Reproduction of CS Figure 20.

Abbreviations: BARI: baricitinib; NRI: non-responder imputation; SALT: Severity of Alopecia Tool.

Some Week 76 data were also presented in Section B.2.8.2 of the CS. These data come from the randomised withdrawal sub-study of BRAVE-AA1 and the down titration sub-study of BRAVE-AA2. Of the [redacted] responders at Week 52 who were re-randomised to stay on baricitinib 4 mg in BRAVE-AA1, [redacted] maintained their SALT ≤20 response at Week 76. In BRAVE-AA2, of the [redacted] responders at Week 52 re-randomised to stay on baricitinib 4 mg in BRAVE-AA1, [redacted] maintained their SALT ≤20 response at Week 76.

Overall, the EAG considers that the Week 52 and Week 76 data presented in the CS may indicate that the Week 36 SALT ≤20 response rates underestimate the long-term efficacy of baricitinib 4 mg treatment. However, because of the uncertainty in these data and absence of placebo data from Week 36, the EAG consider the Week 36 data to be most appropriate to use in the economic analyses.

3.3.2 Subgroup analyses

The EAG's clinical experts outlined three variables that might be associated with the probability of hair regrowth: baseline SALT score or disease severity, length of current AA at baseline and, presence of an atopic background. In the company's prespecified subgroup analyses (CS Section

B.2.7), no significant subgroup-by-treatment interaction terms were observed in BRAVE-AA1 or BRAVE-AA2 and subgroup effects were not considered further. The EAG considers these analyses likely underpowered to detect meaningful subgroup-by-treatment interaction terms because:

- Dichotomised outcome and predictor variables were used when continuous data were available;
- There was a floor effect in the placebo response rate;
- The trial sample size was chosen only to provide appropriate power to detect a main effect of treatment in the primary efficacy analysis.

While the EAG recognises interaction modelling is usually the appropriate approach to detecting subgroup effects, the EAG considers assessing the magnitude of main effect of subgroup within the baricitinib 4 mg arm to be an appropriate measure of subgroup effects in the current appraisal. The company provide this analysis for baseline SALT score in Appendix E of the CS, and for length of current episode at baseline in the BRAVE-AA clinical study reports. These are presented below, and a subgroup analysis based on atopic background status requested in clarification question A14 is presented in Appendix Section 8.3.

3.3.2.1 Baseline severity

Patients in the baricitinib 4 mg who had severe AA at baseline, i.e., SALT 50–94 at baseline, had a higher probability of achieving SALT ≤20 at Week 36 than patients with very severe AA at baseline, i.e., SALT 95–100 at baseline (BRAVE-AA1: severe responders, [redacted]; very severe responders, [redacted]. BRAVE-AA2: severe responders, [redacted]; very severe responders, [redacted]), a relationship that was replicated at the lower 2 mg dose. The EAG notes that a relationship in this direction is expected, as severe patients with SALT 50–94 at baseline are already closer to the SALT ≤20 threshold than very severe patients. Nevertheless, the magnitude of the difference is notable, with severe patients having over twice the SALT ≤20 response rate as very severe patients. These data are presented in Figure 5 and Figure 6.

Figure 5. Proportion of patients with SALT≤20 at Week 36 by baseline AA severity in BRAVE–AA1 (FAS population; primary censoring rule). Reproduction of Figure 3 from CS Appendix E.

Footnotes: **p<0.05 vs placebo; **p<0.01 vs placebo; ***p<0.001 vs placebo.*

Abbreviations: AA, alopecia areata; BARI, baricitinib; FAS, full analysis set; NRI, non–responder imputation; PBO, placebo; SALT, Severity of Alopecia Tool.

Source: CS Appendix E

Figure 6. Proportion of patients with SALT≤20 at Week 36 by baseline AA severity in BRAVE–AA2 (FAS population; primary censoring rule). Reproduction of Figure 4 from CS Appendix E.

Footnotes: **p<0.05 vs placebo; ***p<0.001 vs placebo.*

Abbreviations: AA, alopecia areata; BARI, baricitinib; FAS, full analysis set; NRI, non–responder imputation; PBO, placebo; SALT, Severity of Alopecia Tool.

Source: CS Appendix E

3.3.2.2 *Length of current AA episode at baseline*

The length of current AA episode was highlighted by the EAG's clinical experts as a meaningful variable likely to predict hair regrowth. Table 15 provides SALT ≤20 response data for the baricitinib 4 mg and placebo arms of BRAVE-AA1 and BRAVE-AA2 by the duration of current AA episode at baseline (<4 years and ≥4 years categories). For the baricitinib 4 mg arm, a larger proportion of patients in the <4 years category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]) than those in the ≥4 years category (BRAVE-AA1, [redacted] BRAVE-AA2, [redacted]), a pattern also apparent in the placebo arm data (<4 years category: BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]; ≥4 years category; BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]). In response to clarification question A6, the company provided the correlation between baseline SALT score and baseline length of current AA episode at baseline. No meaningful correlation was observed (Pearson's r = 0.087 in the pooled placebo arms and Pearson's r = 0.039 in the pooled baricitinib 4 mg arms), suggesting that duration of current AA episode and baseline severity are relatively independent and meaningful subgroups.

Table 15. SALT ≤20 response of BRAVE-AA1 and BRAVE-AA2 by the duration of current AA episode at baseline (<4 years and ≥4 years categories)

	Duration of current AA episode at baseline	Week 36 SALT ≤20 response rate	
		BRAVE-AA1	BRAVE-AA2
Baricitinib 4 mg	<4 years	[redacted]	[redacted]
	≥4 years	[redacted]	[redacted]
Placebo	<4 years	[redacted]	[redacted]
	≥4 years	[redacted]	[redacted]

Source CS: Table 30, CS Table 31, BRAVE-AA1 CSR page 310, BRAVE-AA2 CSR page 324
Abbreviations: AA: alopecia areata; SALT: severity of alopecia tool

The company provided scenario analyses based on current duration of AA episode at baseline and results are given in Section 5.1.2.2.

3.3.3 *Withdrawal, down-titration, and relapse*

Data from the randomised withdrawal sub-study (BRAVE-AA1) and the down-titration sub-study (BRAVE-AA2) are presented in CS Section B.2.8.2. [redacted] SALT ≤20 responders in the BRAVE-AA1 4 mg baricitinib arm were re-randomised to placebo at Week 52. By Week 76, only [redacted] of these

patients ([redacted]) had maintained their SALT ≤20 response. [redacted] SALT ≤20 responders in the BRAVE-AA2 mg baricitinib arm were re-randomised to 2 mg baricitinib at Week 52. By Week 76, [redacted] of these patients ([redacted]) had maintained their SALT ≤20 response. The EAG notes that these data indicate that baricitinib is only therefore viable as a continued long-term treatment, and once patients have their treatment withdrawn, hair loss is common.

The company stated that data on trial-defined relapse were not yet available (clarification question A10). Relapse was measured in SALT ≤20 responders in the trial after Week 52 and was defined as a >20-point absolute worsening in total SALT score.

3.3.4 Non-SALT measures of hair regrowth

The company presents responder-based results of two non-SALT based hair loss-measures, the PRO Scalp Hair Assessment and the ClinRO measure for eyelash and eyebrow regrowth, in Table 16 and Table 17 of the CS. Similar to the SALT ≤20 responder analysis, approximately one third of patients in the baricitinib 4 mg arms achieved PRO responses and ClinRO responses at Week 36, compared to only around 5% of placebo patients for all measures. The EAG considers these results to be consistent with the SALT ≤20 responder analysis and assures that treatment with baricitinib 4mg leads to hair regrowth beyond the scalp.

3.3.5 Health-related quality of life

3.3.5.1 EQ-5D-5L

No meaningful differences were observed in EQ-5D-5L health state index or visual analogue score (VAS) between baseline and Week 36 for any arm in BRAVE-AA1 or BRAVE-AA2, with no more than a [redacted] mean increase in EQ-5D score across either the placebo or baricitinib 4 mg arm (Table 16).

Table 16. EQ-5D data from BRAVE-AA1 and BRAVE-AA2 at baseline and at Week 36.

EQ-5D	BRAVE-AA1		BRAVE-AA2	
	Placebo	Baricitinib 4 mg	Placebo	Baricitinib 4 mg
Health state index UK, mean (SD)				
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 36	[redacted]	[redacted]	[redacted]	[redacted]
VAS, mean (SD)				
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 36	[redacted]	[redacted]	[redacted]	[redacted]

Source: BRAVE-AA1 CSR pages 267, 272, 274 and 279; BRAVE-AA2 CSR pages 269, 274, 276 and 281

Abbreviations: EQ-5D: EuroQol-5 Dimension; SD: standard deviation; VAS; visual analogue scale

3.3.5.2 HADS

Mean change from baseline in the Hospital Anxiety Depression Scale (HADS) scores at Week 36 were presented in CS Table 25, and are reproduced for the placebo and baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2 in Table 17. Compared to baseline, HADS Anxiety decreased by a statistically significantly larger amount in the baricitinib 4 mg arms (BRAVE-AA1 mean change [SE]: [redacted]; BRAVE-AA2: [redacted]) than in the placebo arms (BRAVE-AA1 [SE]: [redacted]; BRAVE-AA2: [redacted]). However. these changes were lower than the most common definitions of the minimal clinically important difference (MCID) of around 1.7 to 2 for HADS scales,[22, 23] although this has not been validated in dermatology[24] or AA specifically. HADS Depression decreased in the baricitinib 4 mg arms of BRAVE-AA1 ([redacted]) and BRAVE-AA2 ([redacted]) but increased slightly from baseline in the placebo arms (BRAVE-AA1 [SE]: [redacted]; BRAVE-AA2: [redacted]). Only the difference between baricitinib 4 mg and placebo in BRAVE-AA2 for HADS Depression was statistically significant at $p < 0.05$.

Table 17. Mean change from baseline in HADS-Anxiety and HADS-Depression scores at Week 36 for the placebo and baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2

Week 36	BRAVE-AA1		BRAVE-AA2	
	Placebo (N = 189)	Baricitinib 4 mg (N = 281)	Placebo (N = 156)	Baricitinib 4 mg (N = 234)
HADS Anxiety				
Mean (SD) baseline score	[redacted]	[redacted]	[redacted]	[redacted]
LSM (SE)	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs placebo	[redacted]	[redacted]	[redacted]	[redacted]
HADS Depression				
Mean (SD) baseline score	[redacted]	[redacted]	[redacted]	[redacted]
LSM (SE)	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs placebo	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 25

Abbreviations: HADS: Hospital anxiety depression scale; LSM: least squares mean; SD: standard deviation; SE: standard error

3.3.5.3 SF-36 and Skindex-16

In addition to EQ-5D and HADS data, the company presented HRQoL data using the SF-36 and Skindex-16 measures at Week 36. In the SF-36 measure:

- There were no significant differences in change from baseline SF-36 physical component score between placebo and baricitinib 4 mg arms in either BRAVE-AA1 or BRAVE-AA2 (CS Table 28);
- There was a statistically significantly greater increase for the baricitinib 4 mg arm versus placebo in BRAVE-AA2 in the SF-36 mental component score, but not, however, in BRAVE-AA1 (CS Table 29). Moreover, the increase in baseline for the baricitinib 4 mg arm, [redacted] (95% CI difference from placebo: [redacted]), is below the most common definitions of the MCID for SF-36, around 3 to 5 points,[25, 26] although data in dermatology is scarce.[27-29] In the baricitinib 4 mg arm of BRAVE-AA1, the SF-36 mental component score numerically worsened from baseline (mean change from baseline: baricitinib 4 mg [SE], [redacted] [redacted]; placebo: [redacted].

In contrast to the SF-36 measure, a large benefit of baricitinib 4 mg over placebo was observed in the Skindex-16 measure, adapted for AA. Skindex-16 is a patient reported questionnaire designed to measure the effects of skin disease on quality of life, which comprises three domains: symptoms, emotions and functioning. Patients in the baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2 had a statistically significant improvement in mean change from baseline to Week 36 in the emotions and functioning domain, that was also significantly greater than the change from baseline in placebo (CS Table 23). While there was also a greater reduction in the symptom component of the Skindex-16 in the baricitinib 4 mg arms of BRAVE-AA1 and BRAVE-AA2, this reduction was only statistically significantly greater than placebo in BRAVE-AA2 (CS Table 23). Moreover, no specific MCID has been validated to date for the Skindex-16 or the AA-adapted version which make the clinical meaning of these results difficult to interpret.[30]

3.3.5.4 EAG critique of BRAVE-AA1 and BRAVE-AA2 trial HRQoL data

In the BRAVE-AA trials, no improvement in HRQoL was observed in EQ-5D for baricitinib 4 mg over placebo by Week 36 and only modest improvements were observed in the HADS and SF-36 measures. A larger improvement was observed in the Skindex-16 scale; a skin disease specific scale designed to be sensitive to quality-of-life changes caused by skin disease.

The company provided three arguments for why the EQ-5D data observed in BRAVE-AA1 and BRAVE-AA2 may be unsuitable for use in the economic models. The EAG does not find these arguments convincing and provides comments on them in Table 18, although the EAG's clinical experts did note reservations about the suitability of each of the EQ-5D scales for measuring QoL in AA, highlighting

how the majority of benefits will be most visible in psychosocial functioning, which the EQ-5D only captures partially.

Table 18. The EAG's critique of the company's discussion of the limitations and representative of the EQ-5D data collected din BRAVE-AA1 and BRAVE-AA2

Company Argument	EAG comment
Patients' baseline HRQoL data were near ceiling in BRAVE-AA1 and BRAVE-AA2, limiting the scope for patients' HRQoL to be improved by treatment. At baseline, patients median EQ-5D was [redacted] or higher in all BRAVE-AA arms, i.e., around the UK population norm for EQ-5D of 0.91 for males and females aged 35-44.[31]	BRAVE-AA1 and BRAVE-AA2 were judged to be high quality international clinical trials with large sample sizes, and the EQ-5D data were replicated between the trials. The EAG considers it likely the trials have appropriately measured a high baseline EQ-5D at the population level.
The AA patients who could have gained most utility benefit through hair regrowth may have been excluded in the BRAVE-AA studies. Specifically, patients with the presence of significant uncontrolled neuropsychiatric disorder, or who were clinically judged by the investigator to be at risk for suicide were excluded from the trials.	In response to clarification question A8 the company confirmed [redacted] of significant uncontrolled neuropsychiatric disorders, but noted that these patients may not have chosen or been invited to screening for the trial—something the EAG's clinical experts agreed with, and pointed to a systematic review and meta-analysis showing a negative relationship between HRQoL and anxiety and depression in AA.[32] However, the EAG does not consider the company to have provided sufficient evidence that either: i) the trial populations were missing a large cohort of patients who would have had low baseline EQ-5D scores nor, ii) provided evidence of the size of this cohort in UK clinical practice.
Scales such as the EQ-5D may lack content validity for indications like AA, which may not lead to issues with patients' mobility, cause pain or impede usual activities (three of the five domains of the EQ-5D).	The company have not provided relevant psychometric data demonstrating a lack of content validity of the EQ-5D in severe AA, and nevertheless continued to use the EQ-5D measure from the Adelphi DSP study in their base case analysis. While it is plausible that not all domains of the EQ-5D may be directly affected by AA—something noted by the EAG's clinical experts—it is possible that AA may have indirect effects on these domains.

Abbreviations: AA: alopecia areata; EAG: evidence assessment group; EQ-5D: EuroQol-5 Dimension; HADS: Hospital anxiety depression scale; HRQoL: health related quality of life.

Hence, the EAG finds it plausible that the BRAVE-AA trials have adequately estimated only a small gain in utility following baricitinib 4 mg treatment at the population level and considers the EQ-5D data collected in the BRAVE-AA1 and BRAVE-AA2 trials to be suitable to inform decision making. Nevertheless, the EAG recognises that severe AA can and does have large negative impacts on quality of life for some patients, something highlighted by the EAG's clinical experts. The EAG

believes that this may not equate to large changes in EQ-5D score at the population level, however, because:

- In the large sample, high-quality BRAVE-AA trials, many severe AA patients had a genuinely near-ceiling EQ-5D;
- Only [redacted] of baricitinib 4 mg patients achieved the clinically meaningful SALT ≤20 response rate at Week 36, meaning that any treatment effect at the population level on HRQoL would be diluted by the [redacted] non-responders.

Hence, because only a minority of patients may have EQ-5D deficits at baseline, and only a minority of these patients will likely respond to treatment, any EQ-5D improvements at the population level are likely to be small. The EAG's clinical experts also noted that:

- HRQoL benefits may lag behind a treatment response in severe AA, as a patient adjusts to the changes in their appearance;
- Baricitinib is not a curative treatment, and patients may continue to suffer anxiety because baricitinib needs to be taken continuously over a long period to maintain hair regrowth, with missed doses potentially resulting in hair loss.

3.3.6 Safety data

The company present the adverse events (AEs) observed in BRAVE-AA1 and BRAVE-AA2 in Section B.2.10 of the CS. A slightly higher proportion of patients treated with baricitinib 4 mg in BRAVE-AA1 ([redacted]) and BRAVE-AA2 ([redacted]) had at least one treatment emergent adverse event (TEAE) than in the placebo arms (BRAVE-AA1 placebo: [redacted]; BRAVE-AA2 placebo: [redacted]). Similarly, a slightly higher proportion of patients treated with baricitinib 4 mg in BRAVE-AA1 ([redacted]) and BRAVE-AA2 ([redacted]) had at least one serious adverse event (SAE) than in the placebo arms (BRAVE-AA1 placebo: [redacted]; BRAVE-AA2 placebo: [redacted]). These data are presented in Table 19, alongside pooled data from the baricitinib 4 mg arm extension phases up to August 2021 (providing approximately 6 months additional data following the first data cuts in February 2021). These data are in-line with the BRAVE-AA1 and BRAVE-AA2 data from Week 36

Table 19. Overview of adverse events in the BRAVE-AA studies up to Week 36, and from the extension phase up to August 2021.

	BRAVE-AA1	BRAVE-AA2	Pooled Extension Phase

	Placebo (N=189)	Baricitinib 4 mg (N=280)	Placebo (N=154)	Baricitinib 4 mg (N=233)	Baricitinib 4 mg (N=540)
Patients with ≥1 TEAE, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Deaths	0	0	0	0	[redacted]
SAEs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs leading to permanent discontinuation from study intervention, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AEs leading to discontinuation from study, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; SAE: serious adverse event; TEAE: treatment emergent adverse event.

Source: CS Table 37 and Table 42

While the rates of AEs were slightly higher for baricitinib 4 mg over placebo, the EAG considers that baricitinib 4 mg had a relatively safe safety profile over the study period. Notably, adverse events of special interest due to the mechanism of baricitinib were not greatly elevated over placebo (Table 20).

Table 20. Adverse events of special interest across all treatment groups in the BRAVE-AA trials. Adapted from CS Table 41.

Adverse event	BRAVE-AA1		BRAVE-AA2	
	Placebo n (%)	Baricitinib 4 mg n (%)	Placebo n (%)	Baricitinib 4 mg n (%)
	(N = 189)	(N = 280)	(N = 154)	(N = 233)
Patients with ≥1 TE infection	[redacted]	[redacted]	[redacted]	[redacted]
TE herpes zoster	[redacted]	[redacted]	[redacted]	[redacted]
TE herpes simplex	[redacted]	[redacted]	[redacted]	[redacted]
Positively adjudicated MACE	[redacted]	[redacted]	[redacted]	[redacted]
Positively adjudicated VTE	[redacted]	[redacted]	[redacted]	[redacted]
Positively adjudicated ATE	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal perforation	[redacted]	[redacted]	[redacted]	[redacted]
Nonmelanoma skin cancer	[redacted]	[redacted]	[redacted]	[redacted]
Malignancies other than NMSC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; MACE: major adverse cardiovascular event; NMSC: nonmelanoma skin cancer; TE: treatment emergent.

Source: CS Table 41, CS page 104

The EAG notes that long-term safety data are not yet available for baricitinib in adults with severe AA. The EAG's clinical experts noted the importance of post marketing pharmacovigilance and highlighted the uncertainty they might feel in giving a patient such an immunomodulator for a long period of time, with serious infections, thromboembolic disease and malignancy being highlighted as long-term safety concerns. The EAG's clinical experts noted these are similar concerns that they have for atopic dermatitis patients who are already being prescribed baricitinib in the NHS, and that longer-term data with up to 4.6 years of follow-up for baricitinib in the treatment of rheumatoid arthritis are still somewhat uncertain, although the results were consistent with the short-term data from the primary publication of the clinical trials.[33]

3.3.6.1 Baricitinib 2mg dose

The EAG notes that the rate of AEs observed in the BRAVE-AA trials was lower for the baricitinib 2 mg. This dose may be used for patients:

- Aged ≥75 years;
- With a history of chronic or recurrent infections;
- Who have dose tapered.

The EAG notes, however, that no patients aged ≥75 years were included in the BRAVE-AA trials. As such the safety data from the trial are unlikely to be representative of this population.

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

The company conducted a feasibility assessment for a network meta-analysis (NMA) comparing the drugs comprising best supportive care for severe AA with baricitinib (Section B.2.9.2 of the CS). The company concluded that no NMA or indirect treatment comparisons (ITCs) were feasible because only two placebo-controlled randomised controlled trials (RCTs) formed a connected evidence loop with BRAVE-AA1 and BRAVE-AA2. Neither of these RCTs reported similar outcomes to BRAVE-AA1 or BRAVE-AA2 at similar timepoints, nor reported sufficient or similar treatment effect modifying baseline characteristics to the BRAVE-AA trials.

Twenty-one RCTs or observational studies did not form a connected network with BRAVE-AA1 and BRAVE-AA2. For these studies:

- Key baseline characteristics were poorly reported and often differed to the BRAVE-AA studies;
- Most studies (17 out of 21) included patients with mild or moderate AA;
- Most studies did not report similar outcomes to BRAVE-AA1 and BRAVE-AA2, and not at similar timepoints.

The EAG considers the company's feasibility assessment to be thorough, and the EAG agrees with the company's decision not to perform an NMA or ITCs. Any unanchored comparisons would be at very high risk of bias and the EAG does not believe it would be possible to appropriately adjust for treatment effect modifiers that differ between the studies.

In the absence of viable NMA or ITC data, the EAG considers there to be some unresolved uncertainty concerning the comparative effectiveness of baricitinib against some, but not all, of the current best supportive care therapies for severe AA. The EAG agrees with the company that most supportive care therapies have very limited effectiveness in treating severe AA. For these therapies, the placebo arm of BRAVE-AA1 and BRAVE-AA2 is an acceptable approximation for the treatment response. However, for DPCP there is evidence of a treatment effect above that of no treatment in some severe AA patients who can tolerate the treatment. Nevertheless, the EAG does not consider comparing baricitinib directly with DPCP to be relevant to this submission as:

- Only a minority of patients receive DPCP, and it causes strong allergic reactions in many of these patients (e.g., serve eczema was reported as treatment-emergent adverse event in 31% of patients in a large meta-analysis[10]);
- The magnitude of the effectiveness of DPCP may have been overestimated by various biases in the efficacy analyses of DPCP trials (see company's response to clarification question A9);
- Over 25% of patients in the BRAVE-AA1 and BRAVE-AA2 trials had been previously treated with topical immunotherapy (and hence are likely DPCP failures), and as such are likely a more severe population than those in the DPCP trials.

Hence, while the EAG maintains that there is some unresolved uncertainty around how much more effective baricitinib might be over DPCP, the EAG notes this is likely a sizeable benefit both in efficacy and safety. As outlined in Section 2.3.3, the EAG considers no active treatment, informed by the placebo arm of the BRAVE-AA trials, to be the most appropriate comparator for this appraisal.

3.5 Conclusions of the clinical effectiveness section

In the CS, the company has presented clinical effectiveness and safety evidence in support of baricitinib for treating adults with severe AA, an indication for which patients have a clear unmet need. The company's evidence comes primarily from the international, placebo-controlled BRAVE-AA1 and BRAVE-AA2 randomised controlled trials. The EAG assessed BRAVE-AA1 and BRAVE-AA2 to be high quality trials, and to provide strong evidence of a clinically meaningful benefit of baricitinib 4 mg over placebo. Pooled across the BRAVE-AA trials, [redacted] of participants in the baricitinib 4 mg achieved the primary SALT ≤20 outcome at Week 36, compared to only [redacted] in the placebo arms, although [redacted] is still a minority of patients.

The EAG considers the trial populations of BRAVE-AA1 and BRAVE-AA2 to be suitably similar to patients in the UK to inform decision making, despite neither of the trials including UK centres. The EAG noted that the exclusion of males >60 years and females >70 years, and the exclusion of patients with AA episode durations >8 years at baseline and no previous sign of regrowth may have biased the efficacy estimates in BRAVE-AA1 and BRAVE-AA2 in favour of baricitinib over placebo. However, the EAG considered any overestimation of the efficacy of baricitinib likely to be balanced by several features of the trial that might cause the efficacy to be underestimated, namely a baseline population with a high rate of treatment-experience, a relatively severe population at baseline and a conservative use of non-responder imputation. The EAG therefore concludes that the trials provide a reasonably unbiased estimate of the efficacy of baricitinib 4 mg in adults with severe AA.

The EAG considers there to be two clinically meaningful variables that predict a patient's response to baricitinib 4 mg treatment: baseline SALT score and baseline duration of current AA episode. Patients with a shorter duration of AA episode and patients with a lower SALT score at baseline are more likely to achieve a clinically meaningful response than patients with a longer duration of current AA episode and patients with higher SALT scores at baseline. The EAG notes, however, that AA episode duration and SALT score are continuous variables for which categorical subgroups cannot be clinically defined.

The EAG considers the BRAVE-AA trials to provide high-quality health-related quality of life data on adults with severe AA. There were no significant increases in baseline for EQ-5D or SF-36 in either of the baricitinib 4 mg arms of BRAVE AA-1 or BRAVE AA-2, and the EAG considers this to be likely due to: i) a genuine high baseline quality of life in many adults with severe AA and, ii) a low absolute response rate to baricitinib 4 mg treatment, i.e., [redacted]. Hence, while many patients may have a

greatly reduced quality of life because of their severe AA, treatment with baricitinib 4 mg may only lead to modest HRQoL benefits at the population level.

The EAG considers baricitinib 4 mg to have displayed a relatively strong short-term safety profile in the BRAVE-AA trials. Over the 36-week double-blind treatment phases of BRAVE-AA1 and BRAVE-AA2, there was only a slightly higher rate of treatment emergent adverse events (TEAEs) and serious adverse events (SAEs) in the baricitinib 4 mg arm versus the placebo arm. The absolute rate of SAEs was low, and the safety profile of baricitinib 4 mg in the BRAVE-AA trials was consistent with its safety profile from trials informing TA681[15] and TA466.[16] Nevertheless, the EAG considers there to be uncertainty concerning the long-term safety of baricitinib 4 mg in treating adults with severe AA, with the EAG's clinical experts noting reservations about providing long-term immunomodulators to otherwise healthy young adult patients with severe AA.

Similarly, the EAG considers there to be uncertainty concerning the long-term effectiveness of baricitinib 4 mg in treating adults with severe AA. The small amount of data available at Week 52 and Week 72 suggests that the [redacted] response rate at Week 36 may be an underestimate of the long-term effectiveness of baricitinib. However, in the absence of comparative data at these timepoints, the EAG considers the Week 36 response rates to be the most robust estimate of the relative efficacy of baricitinib at a clinically relevant timepoint. However, data from later timepoints in the withdrawal sub study of BRAVE-AA1 and down-titration study in BRAVE-AA2 suggest that hair loss is common as soon as treatment is stopped, and treatment efficacy reduced upon down-titration.

In general, the EAG considers the submitted evidence to suitably match the decision problem defined in the final scope issued by NICE.[11] However, the EAG notes that the comparator proposed by the company, "Watch and wait" with active monitoring, is not a common management strategy used in clinical practice for adults with severe AA. The EAG agrees with the company that the no active treatment component of "Watch and wait" is a common management strategy used by patients with severe AA, but notes that this is not usually associated with intensive monitoring. The EAG considered three plausible comparators for baricitinib 4 mg:

- A comparison with DPCP, the most effective treatment used to treat severe AA in clinical practice and only active treatment recommended by the 2012 BAD Guidelines.[9] The EAG notes, however, that there exist no data to perform a valid indirect comparison between

DPCP and baricitinib 4 mg in adult severe AA patients, that DPCP is only available to a minority of AA patients, leads to severe adverse reactions in many patients and has a high rate of relapse;

- A comparison with the "basket" of non-DPCP therapies currently used to treat severe AA, primarily systemic immunosuppressants and systemic corticosteroids. Again, the EAG notes the lack of suitable evidence to perform such indirect comparisons. Moreover, where evidence exists, it suggests these therapies are ineffective;
- A comparison with no active treatment and discharge from care.

The EAG, in agreement with the company, considers a comparison with no active treatment to be the most relevant for the current submission as it reflects a commonly used treatment option by adults with severe AA, and is the only comparison for which high-quality data comparative effectiveness data are available. However, the EAG considers no active treatment and discharge from care, rather than "Watch and wait" with active monitoring, to be the appropriate comparison. The EAG nevertheless notes a large degree of heterogeneity and a lack of data around the treatment pathway of AA and severe AA in UK clinical practice, and potential differences between the prevalent and incident populations of adults with severe AA in clinical practice. For treatment experienced severe AA patients, i.e., the prevalent population in UK practice that would receive baricitinib at the point of approval, the EAG considers no active treatment and discharged from care to be the most appropriate comparator. For newly diagnosed cases of severe AA, the EAG considers it likely that contact immunotherapy, systemic immunosuppressants or systemic steroids may be trialled, but their use is heterogenous and access inequitable, and only contact immunotherapy and wig use are recommended by the 2012 BAD Guidelines. The EAG does not consider there to be a widely accepted standard of care for these patients, and in lieu of robust treatment pattern or comparative effectiveness data, the EAG considers no active treatment and discharge from care to be an acceptable comparator for this population.

4 Cost effectiveness

Table 21 below presents the incremental cost-effectiveness results of the company's updated (i.e., post clarification) base case results.

Table 21. Company's base case results post clarification

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
'Watch and wait'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	18,072
Probabilistic results							
'Watch and wait'	[redacted]	-	[redacted]	-	-	-	-
Baricitinib	[redacted]	-	[redacted]	[redacted]	-	[redacted]	17,942

Abbreviations: ICER, incremental cost effectiveness ratio, LYG, life year gained; QALY, quality adjusted life year.

4.1 EAG comment on the company's review of cost effectiveness evidence

The company carried out a systematic literature review (SLR), using a single search strategy, to identify existing:

- Cost-effectiveness studies for the treatment of adult patients with severe alopecia areata (AA);
- Health-state utility values (HSUVs) for patients with severe AA; and,
- Cost and resource use evidence for the treatment of adult patients with severe AA.

Searches were initially run in August 2021 and were last updated in January 2022. A summary of the Evidence Assessment Group's (EAG's) critique of the methods implemented by the company to identify relevant evidence is presented in Table 22. Due to time constraints, the EAG was unable to replicate the company's searches and appraisal of identified abstracts.

Table 22. EAG's critique of company's systematic literature review

Systematic review step	Section of CS in which methods are reported			EAG assessment of robustness of methods
	Cost effectiveness evidence	HRQoL evidence	Resource use and costs evidence	
Search strategy	Appendix G	Appendix G	Appendix G	Appropriate

Inclusion/ exclusion criteria	Appendix G	Appendix G	Appendix G	Appropriate
Screening	Appendix G	Appendix G	Appendix G	Appropriate
Data extraction	Appendix G	Appendix H	Appendix I	Appropriate
Quality assessment of included studies	Appendix G	Appendix G	Appendix G	Appropriate

Abbreviations: CS, company submission; EAG, evidence assessment group; HRQoL, health related quality of life.

The SLR identified a total of 597 records. The SLR did not identify any cost-effectiveness studies for any treatment for AA. A total of 30 publications related to health-related quality of life (HRQoL) and four costs studies were identified by the SLR.

Of the 30 extracted and HRQoL studies, one reported AA quality of life (AAQOL) index values, 17 reported Dermatology Life Quality Index (DLQI) scores, four reported Hospital Anxiety and Depression Scale (HADS) scores, eight reported Skindex values, three reported EQ-5D values directly and SF-36 values were reported in six studies. However, none of the extracted utility data were deemed suitable by the company and instead utility values used in the model are from a company sponsored study (the Adelphi disease-specific programme (DSP) study)[13, 14] with scenarios informed by utility data from the BRAVE-AA1 and BRAVE-AA2 trials (See Section 4.2.8).

The company considered none of the four cost papers to be useful to inform the economic analysis as the cost data were not UK specific. As with utilities informing the model, the company obtained UK specific resource use from the Adelphi DSP study.[13, 14] Please refer to Section 4.2.9 for further details on the cost and resource use data applied in the model.

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

4.2.1 NICE reference case checklist

Table 23 summarises the EAG's appraisal of the company's economic evaluation against the requirements set out in the NICE reference case checklist for the base-case analysis, with reference to the NICE final scope outlined in Section 2.

Table 23. NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	All relevant health effects for adult patients with severe AA have been included

Perspective on costs	NHS and PSS	All relevant costs have been included and are based on the NHS and PSS perspective
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Cost-utility analysis has been provided by the company. Fully incremental analysis not required as there is only one relevant comparator in the analysis.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Lifetime horizon (100 years of age)
Synthesis of evidence on health effects	Based on systematic review	The company performed an appropriate systematic review
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs based on EQ-5D from a company sponsored Adelphi DSP study.[13, 14]
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	EQ-5D obtained from a company sponsored Adelphi DSP study[13, 14] which included AA patients with mild, moderate, severe and very severe disease. EQ-5D data were available directly from patients in the BRAVE-AA1 and BRAVE-AA2 trials, but the company only explored this in a scenario analysis.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The EQ-5D data from the company sponsored Adelphi DSP study[13, 14] were based only on responses from UK patients.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The economic evaluation matches the reference case.
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 included in the analysis have been sourced using NHS reference costs, PSSRU and the NHS Drug tariff.[34-36]
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Discount rate of 3.5% has been used for both costs and health effects.

Abbreviations: AA, alopecia areata; DSP, disease specific programme; EAG, evidence review group; NHS, national health service; PSS, personal social services; QALY, quality adjusted life year

4.2.2 Population

The modelled population considered by the company for this Single Technology Appraisal (STA) are adults with severe AA, aligned with the [redacted] marketing authorisation for baricitinib. Severe AA was defined as patients with a Severity of Alopecia Tool (SALT) score higher or equal to 50 at baseline and is reflective of the inclusion criteria for the key trials, BRAVE-AA1 and BRAVE-AA2. Additionally, the company explored subgroups based on disease severity (severe defined as SALT score between 50-94 and very severe, defined as SALT score between 95-100).

Baseline characteristics of the modelled population are based on pooled data from BRAVE-AA1 and BRAVE-AA2. Baseline age and sex of the population included in the model are [redacted] years of age and [redacted] male, which the EAG's clinical experts considered were reflective of the patient population in the UK. Baseline characteristics for the severe subgroup analysis are presented in Table 47 of the company submission (CS).

Generally, the modelled population and subgroups are in line with NICE final scope. However, as mentioned in Section 2.3.1, the BRAVE-AA trial population is most similar to the prevalent population in clinical practice who would be eligible to receive baricitinib at the point of approval, i.e. a later-line treatment experienced population. Furthermore, the BRAVE-AA trials provide data on a narrower population than those who could receive baricitinib in clinical practice, as patients with baseline AA episodes >8 years and males >60 years and females >70 years were excluded. Such patients would be eligible to receive baricitinib per the [redacted] marketing authorisation but may be less likely to achieve hair regrowth. Although, the EAG considers that impact on treatment response based on excluded patients is balanced out by the fact that around [redacted] of participants in the BRAVE-AA trials had received, and likely failed on, prior therapies, including over [redacted] receiving therapies used specifically for severe AA in UK practice. Please see Section 2.3.3 for further details.

4.2.3 Interventions and comparators

The intervention considered in the economic analysis is baricitinib 4 mg, once daily. Additionally, the SmPC states that, "*a dose of 2 mg once daily may be appropriate for patients such as those aged ≥ 75 years and for patients with a history of chronic or recurrent infections. A dose of 2 mg once daily may also be considered for patients who have achieved sustained control of disease activity with 4 mg once daily and are eligible for dose tapering*".[37] However, the company has not considered dose

tapering or subgroup analysis by age to account for the 2 mg dose in the CS. However, it should be noted that in the model, less than [redacted] of patients aged over 75 years remain on treatment.

The comparator in the analysis is 'Watch and wait', which the company assumes is akin to the placebo arm in the BRAVE-AA1 and BRAVE-AA2 trials. As such, the 'Watch and wait' arm in the economic model is associated with no drug acquisition costs but includes costs associated with regular monitoring. The NICE final scope lists the main comparator as established clinical management without baricitinib.[11]

4.2.3.1 EAG critique

As mentioned in Sections 2.2.1 and Section 2.3.3, the EAG agrees with the company that the no active treatment component of "Watch and wait" is a common management strategy used for adults with severe AA. However, the EAG's clinical experts considered that patients would not be regularly monitored if they are not on treatment. Additionally, the EAG's clinical experts advised that there is no standard treatment pathway. Patients with severe disease are most likely to have had systemic corticosteroids or systemic immunosuppressants but response to treatment is limited. In particular, the EAG's clinical experts advised that for the prevalent population, patients are likely to have explored all available treatment options and that a significant proportion of patients may not take up treatment. As such, for the prevalent population, patients are likely to manage their condition using wigs or complete hair removal. For the incident (or newly diagnosed) population, the EAG considers (based on advice from its clinical experts) that patients are likely to be less treatment-experienced than the prevalent population but does not consider any specific systemic immunosuppressants or systemic corticosteroids to be an established standard of care for these patients, and considers no active treatment to be a realistic endpoint for these patients.

As such, the EAG considers that the relevant comparator for the patient population is 'discharged from care' and thus the assumption of active monitoring of patients not on any treatment is not reflective of UK clinical practice. As such, for the EAG preferred assumptions, the comparator is defined as 'discharged from care' and the removal of associated costs of monitoring (discussed further in Section 4.2.9) are excluded.

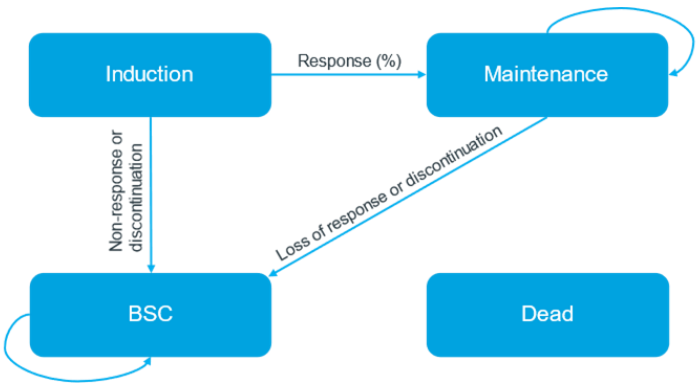
During the clarification stage, the EAG requested the company to explore dose tapering scenarios in line with the guidance in the Summary of Product Characteristics (SmPC). In their clarification response, the company stated that the cost for 4 mg and 2 mg is the same and thus does not affect

costs. Furthermore, the company explained that patients with a sustained response who are down titrated to 2 mg and do not maintain their response will resume the 4 mg dose to restore their previous response. As such the company did not consider a dose tapering scenario to be informative as remedial measures will be employed and thus in the long-term, the cost-effectiveness of baricitinib is unlikely to be affected. However, there is no direct evidence to suggest a "loss and regain" effect when remedial measures are used for patients who have lost response based on dose tapering. Nonetheless, the EAG considers the company's justification for not exploring dose tapering to be reasonable.

4.2.4 *Modelling approach and model structure*

A single *de novo* Markov model was developed in Microsoft Excel© to assess the cost-effectiveness of baricitinib 4 mg compared with a 'Watch and wait' approach for the treatment of adults with severe AA. The company structured the model using previous economic models for other dermatological disorders, such as psoriasis and atopic dermatitis due to a lack of AA models identified in the literature. The aim of the model developed by the company was to estimate the treatment pathway for patients beginning treatment for severe AA (first-line treatment). To capture all costs and benefits associated with treatment until death, the health states within the model include induction, maintenance, best supportive care (BSC) and death. Figure 7 presents the schematic of the Markov model.

Figure 7. Model structure (Figure 26 in company submission)



Abbreviations: BSC, best supportive care.

All patients enter the model via the induction state and either start treatment on baricitinib 4 mg or are regularly monitored ('Watch and wait'). The duration of the induction phase is 36 weeks, and patients transition through nine tunnel states, each lasting four weeks in duration. At any point

during the induction phase, patients can transition to the BSC health state due to all-cause treatment discontinuation (excluding discontinuation due to lack of efficacy).

Following the end of the 36-week induction phase, patients in the baricitinib 4 mg and 'Watch and wait' treatment groups are assessed on their response to treatment. Responders to treatment at Week 36 (defined as achieving $SALT_{50}$) transition to the Maintenance health state where they remain until loss of response or treatment discontinuation due to other causes (all cause discontinuation). Baricitinib patients that transition to the Maintenance health state at Week 36 continue to remain on a 4 mg dose as treated in the induction phase. Patients on 'Watch and wait' who enter the maintenance phase continue with regular monitoring. Patients transitioning to the Maintenance health state are stratified into $SALT_{50}$ and $SALT_{75}$ subgroups depending on relative hair regrowth to allow for differences in utility to be captured. In the model, after the 36-week treatment response assessment, patients remained either $SALT_{50}$ or $SALT_{75}$ unless they discontinued treatment for any reason and thus transition to the BSC health state. Please see Section 4.2.5 for further details on the definition of treatment response and treatment discontinuation applied in the model.

Non-responders were classified as those who fail to achieve $SALT_{50}$ at the end of the induction phase and transition to the BSC state alongside those who discontinued treatment during the induction phase. Patients in the BSC state remain there until the end of the model time horizon or death. At any point in the model time horizon, patients can transition to death from all health states and no patients can experience remission after the 36-week treatment response assessment (that is, transition from being a non-responder to a responder). Transition probabilities to death reflect the UK general population mortality rates (see Section 4.2.7 for further details).

The model was designed to capture responses to treatment over a lifetime horizon (until a patient reaches 100 years of age) and model cycle length was 4 weeks. No half-cycle correction was included in the model due to the short cycle length. The perspective of the analysis was based on the UK National Health Service (NHS) with an annual discount rate of 3.5% being applied for both costs and quality-adjusted life-years (QALYs) captured by the model as per the NICE reference case.

4.2.4.1 EAG critique

The EAG considers the company's model structure to be appropriate and allows important differences in costs and QALYs to be captured. Additionally, the model structure is similar to previous analyses of similar dermatological diseases, such as atopic dermatitis.[15, 38] However, the

EAG has key issues with the underlying assumptions included for each the health states, which are explored throughout the rest of Section 4. In particular, the EAG considers the company's definition of treatment response and the distinction between a patient achieving $SALT_{50}$ and $SALT_{75}$ to allow for addition utility gain may not be reasonable and this is further explored in Sections 4.2.5 and 4.2.8.

4.2.5 Treatment response

In the model, the primary treatment response measure was the achievement of $SALT_{50}$ at Week 36 based on pooled data from the BRAVE-AA1 and BRAVE-AA2 trials. The treatment response of $SALT_{50}$ is a relative measure of response and is defined by the company as at least a 50% improvement from baseline SALT score. In addition to the outcome of $SALT_{50}$, the company also included the outcome of $SALT_{75}$ (defined as at least a 75% improvement from baseline SALT score), as a way of capturing additional quality of life benefit associated with achieving an increased relative improvement in hair growth. The company also explored treatment response by severity as additional scenarios.

Treatment response data included in the company's base case is presented in Table 24.

Table 24. Pooled treatment response at Week 36 from BRAVE-AA1 and BRAVE-AA2

Intervention	$SALT_{50}$ (SE)	$SALT_{75}$ (SE)
Baseline SALT 50-100 patients (FAS population)		
Baricitinib 4 mg	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]
Baseline SALT 50-94 patients (severe population)		
Baricitinib 4 mg	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]
Baseline SALT 95-100 patients (very severe population)		
Baricitinib 4 mg	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]

Abbreviations: FAS, full analysis set; SALT, Severity of Alopecia Tool; SE, standard error.

At any point during the induction phase (prior to the Week 36 treatment response assessment point), patients can transition to the BSC health state due to all cause discontinuations excluding lack of efficacy. Table 25 presents the treatment discontinuation data used during the induction phase of the model. The 36-week data were adjusted to reflect the 4-week cycles included in the model.

Table 25. Pooled 36-week treatment discontinuation (excluding lack of efficacy) from BRAVE-AA1 and BRAVE-AA2

Population (Baseline SALT scores)	Baricitinib 4 mg	'Watch and wait'
SALT 50-100 patients (FAS population)	[redacted]	[redacted]
SALT 50-94 patients (severe population)	[redacted]	[redacted]
SALT 95-100 patients (very severe population)	[redacted]	[redacted]

Abbreviations: FAS, full analysis set; SALT, Severity of Alopecia Tool; SE, standard error.

4.2.5.1 Long-term treatment discontinuation

In the model, at Week 36, patients will either move into the maintenance health state if they have achieved a treatment response at Week 36 or move to BSC if they do not achieve a response. Patients in the maintenance health remain on treatment (baricitinib 4 mg or 'Watch and wait') and only transition to the BSC health state due to all-cause discontinuation. All-cause discontinuation was defined as discontinuation from treatment for all causes including lack of efficacy.

For the 'Watch and wait' arm of the model, the pooled Week 0-36 all-cause discontinuation rate from the placebo arms of the BRAVE-AA1 and BRAVE-AA2 trials was used, due to a lack of data beyond 36 weeks. The 36-week all-cause discontinuation data for 'Watch and wait' was then converted into an annual rate to be used for the model.

For baricitinib 4 mg, pooled Week 0-52 all cause discontinuation data from the BRAVE-AA1 and BRAVE-AA2 was used. Table 26 presented the long-term discontinuation data used in the model. The annual all-cause discontinuation data were adjusted to reflect the 4-week cycles included in the model.

Table 26. Annual pooled all-cause discontinuation from BRAVE-AA1 and BRAVE-AA2

Population (Baseline SALT scores)	Baricitinib 4 mg	'Watch and wait'
SALT 50-100 patients (FAS population)	[redacted]	[redacted]
SALT 50-94 patients (severe population)	[redacted]	[redacted]
SALT 95-100 patients (very severe population)	[redacted]	[redacted]

Abbreviations: FAS, full analysis set; SALT, Severity of Alopecia Tool; SE, standard error.

4.2.5.2 EAG critique

The EAG primary concern with the company's approach to treatment response included in the model was the definition of response employed at Week 36. In the BRAVE-AA1 and BRAVE-AA2 trials, the primary endpoint was the proportion of patients achieving SALT≤20 at Week 36. A response of SALT≤20 indicated scalp hair loss of less than 20% (or ≥80% scalp coverage with hair). The EAG considers that SALT≤20 represents an absolute measure of response, which its clinical experts considered was a more clinically meaningful outcome for patients as, definitively, they will have at least 80% hair regrowth on the scalp and thus may stop wearing wigs or shaving their head. The company's base case approach of using $SALT_{50}$ is a relative improvement from baseline in hair regrowth on the scalp and thus may still be patchy and require the use of wigs or hair removal.

During the clarification stage, the EAG requested the company to provide a scenario exploring the outcome of SALT≤20 in the model. Additionally, based on the advice from the EAG's clinical experts, a scenario exploring SALT≤10 (defined as scalp hair loss of less than 10% or ≥90% scalp coverage with hair) was also requested. The company provided pooled Week 36 treatment response data using SALT≤20 and SALT≤10 (Table 27) and ran these data in a scenario (presented in Section 5.1.2.2).

Table 27. Pooled treatment response at Week 36 from BRAVE-AA1 and BRAVE-AA2

Intervention	SALT≤20 (95% CI)	SALT≤10 (95% CI)
Baseline SALT 50-100 patients (FAS population)		
Baricitinib 4 mg	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]
Baseline SALT 50-94 patients (severe population)		
Baricitinib 4 mg	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]
Baseline SALT 95-100 patients (very severe population)		
Baricitinib 4 mg	[redacted]	[redacted]
'Watch and wait'	[redacted]	[redacted]

Abbreviations: CI, confidence interval; FAS, full analysis set; SALT, Severity of Alopecia Tool.

The EAG considers that SALT≤20 is the most appropriate definition of response at Week 36 for use in the model as it is the primary endpoint in the key BRAVE trials and based on the EAG's clinical experts, is a more clinically meaningful outcome for patients. As such, the EAG has included SALT≤20 at Week 36 in its preferred assumptions, presented in Section 6.4.

The EAG notes that in a randomised double-blind trial, where placebo is the comparator, it can be argued that any observed placebo response is due to: increased medical attention as a result of being in an RCT, an unconscious expectation by the patient and the investigator that the patient will improve, as well as a patient's profound desire to get better. This is particularly the case when considering outcomes that have a subjective component like an assessment of SALT score. In clinical practice, where patients will know what treatment they are receiving or, as in this particular case, know when they have been discharged from clinical care, the EAG considers it unlikely that an observed placebo response from a clinical trial would occur.

However, it could also be argued that the factors influencing a perceived placebo response are also present in the active treatment group, in this case baricitinib. To negate the potential additional benefit present in the outcomes for baricitinib, the EAG considers it reasonable to use the placebo group from the trial as a surrogate for 'discharged from care'. With the rationale being that any "placebo effect" in both arms will "cancel out" and the incremental results from the cost-effectiveness analysis (and so the resulting ICER) will be based solely on the "true" treatment effect of baricitinib. With regards to treatment discontinuation for the comparator arm in the model, the EAG considers that this should be viewed as the placebo effect waning, as patients are not on active treatment.

The EAG considered that for long-term all-cause discontinuation, using Week 0-52 data for baricitinib 4 mg for the maintenance phase may not be representative of discontinuation of patients with a sustained response and during the clarification stage, requested the company to explore all cause discontinuation based on data for Week 36-52. The company supplied the requested data (adjusted to an annual rate), which estimated all-cause discontinuation for baricitinib 4 mg to be [redacted], based on Week 36-52 data, which is lower than the company's base case estimate of [redacted]. Results of the scenario are presented in Section 5.1.2.2 and this has been included in the EAG preferred assumptions.

4.2.6 Adverse events

The company did not include the impact of adverse events (AEs) in the model as they considered observed AEs from BRAVE-AA1 and BRAVE-AA2 to be mild and would not have a significant impact on HRQoL or costs.

4.2.6.1 EAG critique

The company state that their approach of not including AEs in the economic model is aligned with TA681 and TA534, but the EAG considers this is not accurate.[15, 38] In both TA681 and TA534, the impact of AEs was included in the economic models.[15, 38] As such, during the clarification stage, the EAG requested the company to include a scenario which considers the impact of AEs in terms of costs. The EAG focussed only on costs as it considered that the impact of AEs would be captured in the utility estimates derived from BRAVE-AA1 and BRAVE-AA2, which is the preferred source of utility data (see Section 4.2.8). The company provided this scenario using pooled AE rates from BRAVE-AA1 and BRAVE-AA2 (presented in Table 28). However, the company provided no justification for the inclusion of specific AEs nor the definition used (such as treatment-emergent or serious AEs). Additionally, no sources were provided for the costs used for the scenario.

The impact on the incremental cost-effectiveness ratio (ICER) from the inclusion of costs for AEs was minimal (see Section 5.1.2.2), but the EAG considers it is good practice to include costs associated with AEs in the model, especially when treatment is long-term if response is achieved. As the EAG was unable to verify the inputs used in the company's AE scenario and was unable to produce an alternative scenario due to a paucity of time, costs of AEs have not been included in the EAG base case. Nonetheless, the EAG requests that during technical engagement, the company provides a more thorough description and justification of their approach to the inclusion of AEs and assumed unit costs to treat each AE and update the scenario if necessary.

Table 28. AEs and costs included in the company scenario analysis

Adverse event	Unit cost*	Baricitinib 4 mg		'Watch and wait'	
		Induction (36-weeks)	Maintenance (annual)	Induction (36-weeks)	Maintenance (annual)
Upper respiratory tract infection	£39.00	[redacted]	[redacted]	[redacted]	[redacted]
Nasopharyngitis	£39.00	[redacted]	[redacted]	[redacted]	[redacted]
Headache	£206.34	[redacted]	[redacted]	[redacted]	[redacted]
Acne	£171.53	[redacted]	[redacted]	[redacted]	[redacted]
Total cost	-	**£19.48**	**£27.79**	**£15.86**	**£22.70**

Abbreviations: AE, adverse event

*Source or justification of cost not provided by the company

4.2.7 Mortality

Treatment with baricitinib is assumed not to impact mortality. As such, the company included background mortality such that the transition probability to the death state per model cycle was equal for both arms. All-cause mortality was based on Office for National Statistics (ONS) UK lifetables.[39]

4.2.8 Health-related quality of life

In the base case analysis, QALYs accrued by the patient cohort in each model cycle are dependent on the utility attributable to each model health state and an age-related reduction in quality of life.

The BRAVE-AA1 and BRAVE-AA2 trials collected EQ-5D data, as well as data from the SKINDEX-16 AA, SF-36 and HADS questionnaires, up to Week 36 directly from patients but the company stated that the values obtained from the trials were insensitive to changes in the severity of AA and lacked content validity as baseline values were almost the same as UK age- and sex-adjusted general population values. As such, the utility values informing the economic model were derived from a company sponsored Adelphi DSP study.[13, 14] Utilities based on EQ-5D and HADS data from the BRAVE-AA1 and BRAVE-AA2 trials were explored in a scenario analysis.

The Adelphi DSP study collected EQ-5D-5L data from patients with AA in Europe (including the UK). The study was initiated in October 2021. Details of the Adelphi DSP study were limited in the CS and as such, the EAG requested further information during the clarification stage. In their response to clarification, the company provided the questionnaire used for patient reported disease burden, a data file of the responses from the study as well as the overall objective (used as proxy for the study protocol) for the utility aspect of the Adelphi DSP study (the other aspect was to obtain resource use from treating physicians of patients with AA).

The objective of the utility aspect of the Adelphi DSP study was to characterise the patient reported disease burden of AA based on physician-rated current severity by:

- Describing SKINDEX-16 AA, HADS, EQ-5D-5L and Work Productivity and Activity Impairment (WPAI);
- Reporting on concordance in patient and physician ratings of severity; and
- Assessing predictors of patient reported burden measures (Skindex-16 AA, HADS, EQ-5D and WPAI).

However, there were sections in the questionnaire (after the patient filled out their responses to the utility instruments) that were focused on the effects of AA on work and daily life as well as the patients' feelings about their condition.

Overall, there were [redacted] responses to the EQ-5D questionnaire. Responses were stratified by physician-reported current severity of a patient's AA episode and only responses for severe and very severe were considered for the analysis ([redacted] responses). The crosswalk algorithm by Hernandez et al.[40] was used to convert the EQ-5D-5L values to the EQ-5D-3L. The health state utility values (HSUVs) from the final analysis of the Adelphi DSP study (provided during the clarification stage) informing the model are presented in Table 29. It should be noted that the same utility values were used for the severity subgroup analysis.

Table 29. Utility values informing the model from the Adelphi DSP study

Health state	Utility value (SE)	Comments
Induction (up to Week 36)	[redacted]	Baseline score for the severe and very severe subgroup.
Maintenance - $SALT_{50}$	[redacted]	Utility value for the moderate severity subgroup. In the model, the company implemented the utility value as a change from baseline utility gain, calculated as the difference between baseline scores for moderate and severe/very severe subgroup ([redacted])
Maintenance - $SALT_{75}$	[redacted]	Utility value for the mild severity subgroup. In the model, the company implemented the utility value as a change from baseline utility gain, calculated as the difference between baseline scores for mild and severe/very severe subgroup ([redacted])
BSC	[redacted]	Baseline score for the severe and very severe subgroup.

Abbreviations: BSC, best supportive care; DSP, disease-specific programme; SALT, Severity of Alopecia Tool; SE, standard error.

Utilities in the model were adjusted for age, as per the NICE methods guide.[41] The multiplicative approach was used as recommended by the Decision Support Unit (DSU) Technical Support Document (TSD) 12.[42] General population utility values adjusted for age and sex were obtained from the HSE 2014 dataset, as recommended by the DSU.[40]

4.2.8.1 EAG critique

The EAG considers the company's justification for not using pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials (lack of sensitivity and content validity) is a criticism of the EQ-5D tool and not the methods to obtain the data used in the trial. As such, the EAG considers the company's criticism

of the trial EQ-5D data extends to the EQ-5D data obtained from the company sponsored Adelphi DSP study. The EAG asked the company to explain why EQ-5D data from the Adelphi DSP study were more appropriate than the EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials. The company stated that there was a substantial ceiling effect present in the trial data. The company estimated that around [redacted] of participants in the BRAVE-AA1 and BRAVE-AA2 trials reported a score of perfect health at baseline (score of 11111) and as such an improvement in HRQoL would not be obtained at Week 36 for these patients. The company explained that in the Adelphi study, the ceiling effect was observed, but not to the same extent. However, the company did not provide the overall proportion of patients reporting a score of perfect health from the Adelphi DSP study, rather they presented data by each domain of the EQ-5D (Table 26 of the company clarification response).

Additionally, the company only reported perfect health score data for the severe and very severe subgroup but did not supply any information for the mild and moderate subgroup which inform the utility gain in the model for patients who achieve $SALT_{50}$ and $SALT_{75}$ outcomes used in the company base case analysis. The EAG considers that mild and moderate severity patients in the Adelphi DSP study are more likely to report scores of perfect health as their disease is, by definition, less severe, and unlikely to have a more significant impact on HRQoL compared with severe and very severe patients. As such, the utility gain in the economic model may be biased if there was a high proportion of mild and moderate severity patients reporting a score of perfect health in the Adelphi DSP study.

Additionally, based on the physician survey used for the Adelphi DSP study,[13, 14] definition of severity was not based on SALT score, but categories of severity (mild, moderate, severe, and very severe). The EAG considers the severity categories from the Adelphi DSP study represent absolute scalp hair coverage and the difference between the severe/very severe and moderate/mild severity reflects an assumed change in absolute hair regrowth rather than the relative change from baseline assumed in the economic model.

Conversely, the pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials is based only on patients with severe and very severe disease at baseline (defined as SALT 50-100) and the change from baseline at Week 36 estimated by the company is an observed change from baseline score for patients achieving $SALT_{50}$ and $SALT_{75}$ outcomes used in the company base case analysis, rather than baseline scores for patients with mild and moderate severity AA.

In BRAVE-AA1 and BRAVE-AA2, EQ-5D-5L was measured at Weeks 0, 12, 24 and 36. In the long-term extension phase of BRAVE-AA1 and BRAVE-AA2, EQ-5D-5L was measured at Weeks 52, 64 and 76. Patients were followed up for 200 weeks. Overall, there were 341 responses in the placebo arm and 514 responses in the baricitinib 4 mg arm but pooled data were used to inform the HSUVs. The following linear model was used to analyse the EQ-5D data:

$$cEQ5D = a0 + a1\ EQ5Dbl + a2\ SALTcat + a3\ AGE$$

Abbreviations: cEQ5D, Change in EQ5D; EQ5Dbl, Baseline EQ5D; SALTcat, SALT improvement categories at Week 36 (<50%, ≥50% to <75%, ≥75%); AGE, Age in years.

As with the utility values from the Adelphi DSP study, the company used the crosswalk algorithm by Hernandez et al.[40] to convert EQ-5D-5L values to the EQ-5D-3L. Table 30 presents the pooled EQ-5D health state data from BRAVE-AA1 and BRAVE-AA2 trials. As mentioned previously, the company ran a scenario using HSUVs from the BRAVE-AA1 and BRAVE-AA2 trials and results of the scenario are presented in Section 5.1.2.2. It should be noted that age-matched general population utility value is 0.91. The utility values for patients that achieve $SALT_{50}$ or $SALT_{75}$ using data from BRAVE-AA1 and BRAVE-AA2 are only just below the general population value. Based on feedback from the EAG's clinical experts, this may not be unreasonable as for the majority of patients with severe AA, there is not a significant impact on HRQoL.

Table 30. Health state utility data - pooled EQ-5D data from BRAVE-AA1 and BRAVE-AA2

Health state	Utility value (SE)	Comment
Induction (up to Week 36)	[redacted]	Baseline score for SALT 50-100 FAS population
Maintenance - $SALT_{50}$	[redacted]	Change from baseline for patients achieving $SALT_{50}$
Maintenance - $SALT_{75}$	[redacted]	Change from baseline for patients achieving $SALT_{75}$
BSC	[redacted]	Baseline score for SALT 50-100 FAS population

Abbreviations: BSC, best supportive care; FAS, full analysis set; SALT, Severity of Alopecia Tool; SE, standard error.

The source of utility values in the model is a primary driver of cost-effectiveness. As recommended in the NICE methods guide,[41] the reference case should report the measurement of changes in health-related quality of life directly from patients. As such, the EAG considers the pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials represents a more robust source of utility data that matches the NICE reference case and should be used in the cost-effectiveness analysis for the base case.

As mentioned in Section 4.2.5.2, the EAG considers that the primary trial outcome of SALT≤20 is a more appropriate measure of treatment effectiveness. As such, the EAG requested, and the company provided, change from baseline at Week 36 for patients achieving SALT≤20 based on pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials ([redacted]). Upon request, the company also supplied the change from baseline at Week 36 for patients achieving SALT≤10 based on pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials ([redacted]). The EAG ran a scenario using the SALT≤20 change from baseline utility gain in combination with the outcome of SALT≤20 for treatment effectiveness in the model, as well as the same scenario using data for SALT≤10 and results are presented in Section 6.3. The EAG considers the SALT≤20 combined scenario is a more appropriate approach to the cost-effectiveness analysis and has included it in the EAG preferred assumptions, presented in Section 6.4.

The EAG notes that the experience of severe AA can vary between patients. The EAG's clinical experts advised that for most patients, HRQoL may only be mildly affected and thus may not be that different to the general population but equally HRQoL is severely affected for a few patients (primarily driven by adverse mental health). Additionally, the EAG's clinical experts advised that overtime, patients may come to terms with their hair loss, while a few may remain distressed about their condition. Thus, the EAG acknowledges that there is a small, but heterogenous, patient population that is more adversely affected in terms of HRQoL but that the demographics of this population are difficult to identify clinically and consistently, and it is beyond the scope of assessment to identify that group. Nonetheless, the EAG ran two scenarios around the EAG base case to identify the QALY gain needed for the ICER to reach the £20,000 and £30,000 cost-effectiveness threshold and these are presented in Section 6.4.

4.2.9 Resource use and costs

The costs included in the economic model consist of drug acquisition costs, monitoring resource use and costs, costs associated with BSC, and costs associated with the management of the psychological burden of AA. The details of each are given in the following subsections. Unit costs used in the model were based on 2020/21 price years.

Many of the company's resource use assumptions were informed by the company sponsored Adelphi DSP study.[13, 14] One objective of the Adelphi DSP study was to describe treatment patterns associated with AA based on physician rated severity and this feedback was used to inform the

resource use in the model (described in the following subsections). Only data obtained from physicians treating severe and very severe patients in the UK ([redacted]) were used for the model.

4.2.9.1 Drug acquisition costs

Baricitinib is given as a fixed-dose 4 mg tablet taken once daily and is also available as a 2 mg dose. The list price of a 28-tablet pack of 2 mg or 4 mg is £805.56. There is currently a patient access scheme (PAS) in place for baricitinib such that the fixed price pack is [redacted]. The daily cost of baricitinib 4 mg is [redacted].

No drug acquisition costs were applied to the 'Watch and wait' arm of the model. Instead, patients who are allocated to 'Watch and wait' are actively monitored and these costs are described in Section 4.2.9.2.

4.2.9.2 Monitoring resource use and costs

During the induction and maintenance phases of the economic model, patients on baricitinib and 'Watch and wait' are actively monitored. The company used feedback from clinical experts to inform the assumptions around monitoring during the induction and maintenance phases of the economic model, presented in Table 31. Unit costs for monitoring resource are presented in Table 48 of the company's clarification response and were sourced from NHS reference costs 2020/21 and PSSRU.[34, 35]

The total cost of monitoring in the induction phase (36 weeks) baricitinib 4 mg and 'Watch and wait' was £1,022.75 and £1,011.86, respectively. The annual cost of monitoring in the maintenance health state for baricitinib and 'Watch and wait' was £371.71 and £357.19, respectively. The main difference in costs in the induction and maintenance health state between baricitinib 4 mg and 'Watch and wait' was the inclusion of blood monitoring for baricitinib patients.

Table 31. Monitoring resource use and costs

Resource use	Proportion of patients	Baricitinib 4 mg		'Watch and wait'	
		Induction (36 weeks)	Maintenance (annual)	Induction (36 weeks)	Maintenance (annual)
Dermatologist outpatient consultation	100%	4.0	2.0	4.0	2.0
Dermatologist nurse visit	100%	1.0	0.5	1.0	0.5
Full blood count	100%	3.0*	4.0	0.0	0.0
Wig use (modacrylic wig)	80%	2.0	0.0	2.0	0.0

Orthotics	80%	1.0	0.0	1.0	0.0
*Updated in the company's clarification response.					

4.2.9.3 BSC health state costs

In the BSC health state, costs included drug acquisition and monitoring costs as well as disease management costs. Costs associated with the management of the psychological burden of AA were also included in the BSC health state and are described further in Section 4.2.9.4.

Treatments and the proportion of patients on each treatment included in the BSC health state are presented in Table 32 and were informed by the company's Adelphi DSP study. Treatment dosage was based on each treatment's SmPC. Unit costs were obtained from the NHS drug tariff[36] and are presented in Table 56 of the CS. Confidential medicines unit (CMU) and Drugs and pharmaceutical electronic market information tool (eMIT) prices are available for medicines included in BSC, as such the EAG has produced a confidential appendix to the EAG report. Please refer to Appendix 8.4. for the source of the confidential prices used in the confidential appendix.

The company also included costs of monitoring patients while on treatment in the BSC (Table 33) and assumptions were based on clinical expert opinion obtained by the company. Unit costs for monitoring in the BSC health state were based on NHS reference costs 2020/21.[35] The total annual cost of drug acquisition and monitoring in the BSC health state was estimated to be £3,683.10.

Table 32. Drug acquisition costs in the BSC health state

Treatment	Dose and frequency	Number of doses per year	Proportion of patients*	Annual cost
Ciclosporin	4 mg/kg QD	108,114 (4mg * 74kg * 365.25 days)	13.72%	£355.70
Methotrexate	20 mg per week	1,040 (20mg * 52 weeks)	12.86%	£3.25
Azathioprine	2 mg/kg body weight QD, for 1 year	54,057 (2mg * 74kg * 365.25 days)	2.57%	£3.10
Intralesional steroids (triamcinolone acetonide)	5 mg repeated every other week	130 (5mg * 26 weeks)	9.43%	£0.46
DPCP (contact immunotherapy) treatment	Weekly treatment for 9 months	36 (4 times per month for 9 months)	21.63%	£890.79
Prednisolone*	0.4 mg/kg QD	10,811 (0.4mg * 74kg * 365.25 days)	17.15%	£10.46

TCS: Mometasone scalp lotion*	2ml QD	730.5 (2 ml * 365.25 days)	24.77%	£13.15
Minoxidil 5% foam (topical)*	1 g BID (men) or 1 g QD (women) - discontinue if no improvement after 16 weeks (men) or 24 weeks (women)	202 [39.3%§ *(2g * 7 days * 16 weeks) + 60.7%§*(1g * 7 days * 24 weeks)]	5.72%	£3.02
Minoxidil tablets	20 mg QD	7,305 (20mg * 365.25 days)	0.00%	£0.00
Mycophenolate Mofetil	1 g BID, for 1 year	730,500 (2,000mg * 365.25 days)	2.86%	£4.59
Anthralin 0.1% cream	1.5 g QD	242 (1.5g * 7 days * 23 weeks)	5.72%	£1.04
Patients not currently on treatment	4 mg/kg QD	108,114 (4mg * 74kg * 365.25 days)	12.00%	£0.00
Total cost		-	-	**£1,285,56**

Abbreviations: BID, twice per day; DPCP, diphenylcyclopropenone; QD, once per day; TCS, topical corticosteroids

§Sex distribution based on SALT 50-100 pooled FAS population from BRAVE-AA1 and BRAVE-AA2.

*Updated as part of the company's clarification response

Table 33. Drug monitoring in the BSC health state

Treatment	Description of monitoring[35]	Unit cost[35]	Frequency of visits per year	Proportion of patients*	Annual cost
Ciclosporin	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	9	13.72%	£211.81
Methotrexate	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	9	12.86%	£198.57
Azathioprine	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	9	2.57%	£39.71
Intralesional steroids (triamcinolone acetonide)	JC43C – OPROC – Minor Skin Procedures, 19 years and over – Dermatology	£250.70	18	9.43%	£425.65
DPCP (contact immunotherapy) treatment	JC43C – OPROC – Minor Skin Procedures, 19 years and over – Dermatology	£250.70	36	21.63%	£1,952.11
Prednisolone	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	13	17.15%	£382.43
TCS: Mometasone ointment	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	4	24.77%	£169.97

Minoxidil 5% foam (topical)	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	4	5.72%	£39.22
Minoxidil tablets	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	4	0.00%	£0.00
Mycophenolate Mofetil	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	9	2.86%	£44.13
Anthralin 0.1% cream	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	4	5.72%	£39.22
Patients not currently on treatment	-	-	-	12.00%	-
Total cost	-	-	-	-	**£3,502.82**

Abbreviations: DPCP, diphenylcyclopropenone; TCS, topical corticosteroids

*Based on company clinical expert opinion and updated as part of the company's clarification response.

In addition to drug acquisition and monitoring costs, the company included additional disease monitoring costs, based on feedback obtained from clinical experts (i.e. tests, wig use and orthotics) and the company's Adelphi DSP study (dermatologist visits). Table 34 presents the disease management costs applied in the BSC health state. Unit costs were sourced from NHS reference costs 2020/21 and PSSRU.[34, 35] The total annual cost of disease management in the BSC health state was estimated to be £468.47.

Table 34. Disease monitoring costs in the BSC health state

Resource use	Description of monitoring[34, 35]	Unit cost[34, 35]	Frequency per year	Proportion of patients*	Annual cost
Dermatologist outpatient consultation	Weighted average of WF01A-D and WF02A-D - Dermatology	£171.53	2.00	13%	£41.17
Dermatologist nurse visit	PSSRU, 15 minutes of hospital nurse Band 6 patient related time	£28.25	0.50	13%	£1.70
Thyroid function	DAPS05 - Haematology	£3.63	4.00	100%	£14.52
Vitamin D	DAPS05 - Haematology	£3.63	4.00	100%	£14.52
Ferritin	DAPS05 - Haematology	£3.63	4.00	100%	£14.52
Full blood count	DAPS05 - Haematology	£3.63	4.00	100%	£14.52
Liver function	DAPS04 - Clinical biochemistry	£1.85	4.00	100%	£7.40
Renal function	DAPS04 - Clinical biochemistry	£1.85	4.00	100%	£7.40

Tuberculosis	DAPS07 - Microbiology	£10.18	4.00	100%	£40.72
Lipids	DAPS05 - Haematology	£3.63	4.00	100%	£14.52
Wig use (modacrylic wig)	Wigs and fabric supports on the NHS	£75.70	2.00	80%	£121.12
Orthotics	Service Code 658 - Total Outpatient Attendances	£220.46	1.00	80%	£176.37
Total	-	-	-	-	**£468.47**

Abbreviations: BSC, best supportive care.

4.2.9.4 Costs of Psychological management of AA

The company assumed that patients with severe AA will incur costs to manage the psychological burden of AA. The costs of psychological management of severe AA were split between pharmacological (Table 36) and non-pharmacological costs (Table 35) and were assumed to occur in the induction phase and the BSC health state of the economic model. Costs were sourced from PSSRU and CG90.[34, 43] It should be noted that CG90 was replaced by NG222 after the company produced the CS, but costs were not dissimilar between the two sources.[43, 44]

Table 35. Non-pharmacological support costs included in the model

Resource use & description	Unit cost (PSSRU)[34]	Proportion of patients*	Resource use in induction*	Resource use in BSC*
Psychiatrist visit - NICE NG222** - band 7 HI therapist (with MBCT qualification).	£112.00	5.00%	3.00	4.00
Psychologist visit - NICE NG222** - One-hour direct contact (band 5 PWP).	£50.00	10.00%	3.00	4.00
Self-help with support - 1 GP session.	£39.23	12.38%	0.75	1.00
Group exercise & one GP referral visit - 30 sessions x 1 hour each; 1 therapist (band 5 PWP) and 8 participants per group = 30 therapist.	£186 + £39.23	0.75%	0.75	1.00
Interpersonal psychotherapy & one GP referral visit - 8 sessions x 1 hour each = 8 therapist hours per service user (band 7 HI therapist).	£873 + £39.23	0.75%	0.75	1.00
Counselling & one GP referral visit - 12 sessions x 1 hour each = 12 therapist hours per service user (band 7 HI therapist).	£873 + £39.23	1.13%	0.75	1.00
Total costs	-	-	**£49.54**	**£66.05**

Abbreviations: BSC, best supportive care; HI, high intensity; MBCT, mindfulness-based cognitive therapy; PWP, psychological well-being practitioner.

*Based on company clinical expert opinion

**CG90 was replaced by NG222 after the company submission and as such, the relevant tables in NG222 are Table 85-86 of Evidence Review B.[43, 44]

Table 36. Cost of pharmacological treatment for the psychological treatment of severe AA

Treatment	Proportion of patients*	Total cost (incl. 4 GP visits)**	Induction		BSC	
			Resource use	Cost	Resource use	Cost
Sertraline	16.50%	£161.42	0.75	£19.98	1.00	£26.63
Escitalopram	16.50%	£160.26	0.75	£19.83	1.00	£26.44
Duloxetine	5.00%	£164.59	0.75	£6.16	1.00	£8.23
Total costs	-	-	-	**£45.98**	-	**£61.31**

Abbreviations: AA, alopecia areata; BSC, best supportive care; GP, general practitioner; HI, high intensity; MBCT, mindfulness-based cognitive therapy; PWP, psychological well-being practitioner.

*Based on company clinical expert opinion

**Obtained from Table 86, Evidence Review B of CG90. CG90 was replaced by NG222 after the company submission and as such, the relevant table in NG222 is Table 83 of Evidence Review B.[43, 44] However, the costs between the original guidance and the update are not dissimilar.

† Cost of GP visit = £39.23[34]

4.2.9.5 EAG critique

The EAG identified several issues with the company's assumptions around resource use and costs that were deemed by the EAG's clinical experts not to align with UK clinical practice. Primarily, the EAG considers the costs in the model for 'Watch and wait' and the BSC health state to be overestimated based on feedback from the EAG's clinical experts. Overestimation of costs in the BSC health state is a key issue as patients in both arms of the model spend a substantial amount of time in the BSC health state accruing costs with no associated benefit (utility for this health state is set to baseline, see Section 4.2.8).

As mentioned in Section 4.2.3, the EAG's clinical experts considered that 'Watch and wait' for patients with severe AA does not happen in the NHS and patients would not be regularly monitored if they were not receiving treatment. Additionally, the EAG's clinical experts considered that a range of treatments may be given to patients but that these are not very effective. In the company's own Adelphi DSP study, it was estimated that the majority of severe/very severe patients were treatment experienced ([redacted]).[13, 14] Thus, it is likely that if response to treatment is not achieved, patients will not engage with further treatment or will not be followed up (effectively patients are discharged from care). Furthermore, a significant proportion of patients may not take up treatment and instead opt for using wigs to manage their hair loss. The EAG considers that lack of engagement with

treatment or being discharged from care has implications for the costs assumed in the BSC health state for both arms of the model, as patients transition to this health state upon loss of treatment response or treatment discontinuation for any other reason. Thus, the EAG considers that monitoring costs for patients on 'Watch and wait' (what the EAG refers to as 'discharged from care') and disease management costs in the BSC health state should be excluded.

During the clarification stage, the EAG requested, and the company supplied, a scenario where monitoring costs for 'Watch and wait' in both the induction phase and Maintenance health state were removed from the model (see Section 5.1.2.2). Additionally, the EAG requested a scenario where disease monitoring costs in the BSC health state are removed from both arms of the model. However, the company only provided a scenario where disease monitoring costs in the BSC were excluded for the baricitinib 4 mg arm only. As such, the EAG ran a scenario where disease monitoring costs are excluded from the BSC health state for both arms of the model and results are presented in Section 6.3.

A key assumption made by the company which affects both arms of the model in the induction phase and the BSC health state, is the inclusion of costs associated with psychological support for severe AA patients (non-pharmacological interventions). The EAG's clinical experts advised that the company's assumptions of psychological care support were optimistic, and that provision of support is extremely limited given the current pressures faced by the NHS. The EAG's clinical experts did consider the company's assumptions around pharmacological treatment for the management of the psychological burden of severe AA to be reasonable. During the clarification stage, the EAG requested the company to provide a scenario where psychological support costs were removed from the model. The company supplied the requested scenario, but upon further investigation, the EAG found that the company's scenario excluded pharmacological treatment costs in addition to the costs of psychological support. As such, the EAG ran a scenario excluding only the non-pharmacological psychological support costs and the results are presented in Section 6.3.

For patients in the induction phase of the model and in the BSC health state, provision of wigs and orthotics has been assumed to occur for 80% of patients, with 2 wigs and 1 orthotic supplied for both induction and annually in the BSC health state. The EAG's clinical experts advised that wigs and orthotics are predominantly used by female patients, of which at baseline in the model, 60.7% are female. Furthermore, the induction phase of the model is only 36 weeks, yet the assumptions made for wigs and orthotics resource use are the same as the BSC health state, which represents annual

usage. The EAG requested scenarios exploring wigs and orthotics for only female patients and only one wig for the induction phase. The company supplied the requested scenarios, and this demonstrated that changes to the assumptions around wigs and orthotics had minimal impact on the ICER (see Section 5.1.2.2).

Overall, the EAG considers the following assumptions to be a more accurate reflection of costs incurred by severe AA patients in the model and has included these in its preferred assumptions, presented in Section 6.4:

- Exclusion of monitoring costs for 'Watch and wait' in the induction phase of the model (comparator is assumed to be 'discharged from care').
- Exclusion of disease management costs in the BSC health state for both arms of the model.
- Exclusions of psychological support costs (non-pharmacological intervention) in the induction phase and BSC health state for both arms of the model.
- Only one wig assumed in the induction phase for both arms of the model. The EAG decided not to include the assumption of wigs and orthotics use only for female patients only as this may be a strong assumption and the impact on the ICER was minimal.

5 Cost effectiveness results

5.1.1 *Company's cost effectiveness results*

Table 37 presents the cost-effectiveness results of the company's updated (i.e., post clarification) base case deterministic and probabilistic analyses. The company performed probabilistic sensitivity analysis (PSA) to assess the joint parameter uncertainty around base case results. Incremental results from the company's PSA, arising from 1,000 simulations.

In the base case probabilistic analysis, an incremental quality-adjusted life-year (QALY) gain of [redacted] over 'Watch and wait' along with additional costs of [redacted] for the baricitinib 4 mg, generates an incremental cost-effectiveness ratio (ICER) of £17,942 per QALY. The net monetary benefit (NMB) using the £30,000 threshold is [redacted] and the net health benefit (NHB) is [redacted]. A positive NHB implies that overall population health would be increased as a result of the new intervention

A proposed confidential patient access scheme (PAS) discount for baricitinib is applied in the company's base case and is therefore reflected in the results presented in this report. Confidential medicines unit (CMU) and Drugs and pharmaceutical electronic market information tool (eMIT) prices are available for medicines included in best supportive care (BSC) and as such the Evidence Assessment Group (EAG) has produced a confidential appendix to the EAG report. Analyses included in the confidential appendix include the company base case results, and sensitivity and scenario analyses.

Table 37. Company's base case results post clarification

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
'Watch and wait'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	18,072
Probabilistic results							
'Watch and wait'	[redacted]	-	[redacted]	-	-	-	-
Baricitinib	[redacted]	-	[redacted]	[redacted]	-	[redacted]	17,942

Abbreviations: ICER, incremental cost effectiveness ratio, LYG, life year gained; QALY, quality adjusted life year.

A PSA scatterplot is presented in Figure 8 and a cost-effectiveness acceptability curve (CEAC) is presented in Figure 9. Based on these analyses, the probability that baricitinib is cost effective versus

'Watch and wait' is [redacted] at a willingness to pay (WTP) threshold of £20,000 and [redacted] at a WTP threshold of £30,000.

The EAG considers the parameters and respective distributions chosen for PSA to be generally sound. The EAG also considers the probabilistic results to be comparable to the deterministic results.

Figure 8. Cost-effectiveness plane - PSA scatterplot: baricitinib 4 mg vs 'Watch and wait' (company's clarification response appendix, Figure 18)

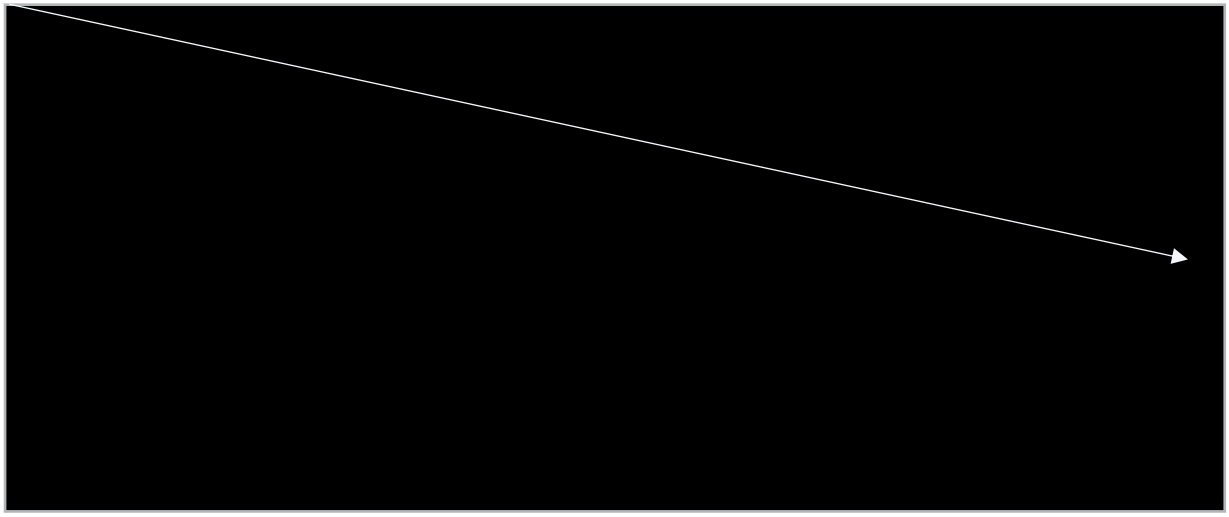


Figure 9. Cost-effectiveness acceptability curve: baricitinib 4 mg vs 'Watch and wait' (company's clarification response appendix, Figure 19)

5.1.2 Company's sensitivity analyses

5.1.2.1 One way sensitivity analysis

The company conducted one-way sensitivity analyses (OWSAs) to assess the impact, on the ICER, of varying specific parameters in isolation and to identify the main model drivers. The results are illustrated using the tornado diagram in Figure 10. The ICER was most sensitive to the frequency and monitoring resource use for diphenylcyclopropenone (DPCP) treatment included in the BSC health state, followed by the Severity of Alopecia Tool 50 ($SALT_{50}$) 36-week response rate for baricitinib 4 mg.

Figure 10. Tornado plot (company's clarification response appendix, Figure 20)

Abbreviations: BSC, best supportive care; DPCP, diphenylcyclopropenone; HSUV, health state utility value; SALT, severity of alopecia tool.

5.1.2.2 Scenario analysis

The company undertook a series of scenario analyses to assess the impact of applying alternative assumptions to key model parameters. In addition, the company conducted several additional scenario analyses requested by the EAG. Results of all the scenario analyses conducted by the company are presented in Table 38. Several requested scenarios were not provided by the company, as such the EAG have conducted these additional scenario analyses and provided the results in Section 6.3.

Table 38. Company scenario analyses - deterministic

Results per patient	Baricitinib 4 mg (1)	'Watch and wait' (2)	Incremental value (1-2)
Company updated base case - post clarification			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			18,072
Starting population with SALT 50-94 (severe subgroup)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			25,154
Starting population with SALT 95-100 (very severe subgroup)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			12,685
Response based on $SALT_{75}$			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			16,490
Utilities based on pooled EQ-5D data from BRAVE-AA1 and BRAVE-AA2			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			174,446
Utilities based on pooled HADS data from BRAVE-AA1 and BRAVE-AA2			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			55,483
Proportion of patients on BSC drugs based on clinical expert opinion			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			Dominant
EAG requested scenarios			
Response based on SALT≤20			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			17,071
Response based on SALT≤20 - Starting population with SALT 50-94 (severe subgroup)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]

ICER (£/QALY)			18,773
Response based on SALT ≤20 - Starting population with SALT 95-100 (very severe subgroup)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			16,929
Response based on SALT≤10			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			20,782
Response based on SALT≤20 and duration of AA episode <4 years			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			16,154
Response based on SALT≤20 and duration of AA episode >4 years			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			18,982
Long-term all-cause discontinuation based on Week 36-52 data for baricitinib 4 mg ([redacted])*			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			16,293
Inclusion of costs for AEs			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			18,348
Removal of monitoring costs for 'Watch and wait' in the induction phase and Maintenance health state			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			20,887
Wig costs weighted by proportion of females (60.67%)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			18,732
Inclusion of only one wig on the induction phase			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			18,068

Abbreviations: AA, alopecia areata; BSC, best supportive care; EAG, Evidence Review Group; HADS, Hospital Anxiety and Depression Scale; ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool

*The EAG has presented results for this scenario from the model, as it could not verify the company's ICER presented in the clarification response (B4bi)

5.1.3 Model validation and face validity check

For the model validation, the company stated that quality control checks were performed by an analyst not involved in the development of the economic model. Additionally, the company provided the model quality assurance checklist used for the validation, which the EAG considers provided a thorough and appropriate check of the model.[45] Consequently, the EAG did not identify any model errors.

6 Additional economic analysis undertaken by the EAG

6.1 Model corrections

The Evidence Assessment Group (EAG) did not identify any model corrections.

6.2 Exploratory and sensitivity analyses undertaken by the EAG

In Section 4 of this report, the EAG has described several scenarios that warrant further exploration in addition to the company's own sensitivity and scenario analyses to ascertain the impact of these changes on the incremental cost effectiveness ratio (ICER). The deterministic scenarios that the EAG has performed are as follows and results are presented in Section 6.3:

- Treatment response at Week 36 defined as achieving Severity of Alopecia Tool (SALT) score of less than or equal to 20 (SALT≤20) in combination with utility values for baseline and change from baseline associated with achieving SALT≤20 from the BRAVE trials (Section 4.2.5.2 and 4.2.8.1):
    - Full analysis set (FAS) - baseline SALT values of 50-100.
    - Severe subgroup - baseline SALT values of 50-94.
    - Very severe subgroup - baseline SALT values of 95-100.
- SALT≤10 and in combination with utility values for baseline and change from baseline associated with achieving SALT≤10 from the BRAVE trials - FAS only (Section 4.2.5.2 and 4.2.8.1).
- No placebo response - all patients in the comparator arm move to the best supportive care (BSC) health state at Week 36 (Section 4.2.5.2)
- Removal of disease monitoring costs in the BSC health state for both arms of the model (Section 4.2.9.5).
- Removal of non-pharmacological support costs (Section 4.2.9.5).

6.3 EAG scenario analysis

Table 39 presents the deterministic results of the EAG exploratory analyses described in Section 6.2. Results reported include the company's proposed patient access scheme (PAS); a fixed pack price of [redacted].

Table 39. Results of the EAG's scenario analyses

	Results per patient	Baricitinib 4 mg	'Watch and wait'	Incremental value
0	**Company base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			18,072
1	**SALT≤20 at Week 36 + baseline and CFB utility from BRAVE trials - FAS population**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			118,494
2	**SALT≤20 at Week 36 + baseline and CFB utility from BRAVE trials - severe population**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			130,303
3	**SALT≤20 at Week 36 + baseline and CFB utility from BRAVE trials - very severe population**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			117,510
4	**SALT≤10 at Week 36 + baseline and CFB utility from BRAVE trials - FAS population**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			129,068
5	**No placebo response (SALT≤20 at Week 36 + CFB utility from BRAVE trials - FAS population)**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			86,343
6	**Removal of disease monitoring costs in the BSC health state for both arms of the model**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			63,941
7	**Removal of non-pharmacological support costs**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			18,679

Abbreviations: BSC, best supportive care; CFB, change from baseline; EAG, Evidence Assessment Group; FAS, full analysis set; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

6.4 EAG preferred assumptions

In this section, the EAG presents its base case ICER for baricitinib 4 mg for treating severe alopecia areata (AA). As discussed in Section 4, the EAG considers that the definition of the comparator arm should be 'discharged from care' based on advice obtained from its clinical experts. As a reminder, the EAG's clinical experts did not recognise the company's comparator of 'Watch and wait' as reflective of what happens in the NHS for patients with severe AA. According to the EAG's clinical experts, patients would not be regularly monitored if they are not on treatment. Additionally, the EAG's clinical experts considered that a significant proportion of patients may not take up treatment (effectively discharged from care) and instead opt for using wigs to manage their hair loss. As such, the comparator for the EAG's base case is 'discharged from care'.

The following assumptions were incorporated into the EAG's base case:

- Treatment response at Week 36 defined as achieving SALT≤20.
    - utility values for baseline and change from baseline associated with achieving SALT≤20Long-term all-cause discontinuation based on Week 36-52 data for baricitinib 4 mg.
- No monitoring costs in the induction phase and Maintenance health state for 'Watch and wait' (comparator defined as 'discharged from care').
- Removal of disease monitoring costs in the BSC health state for both arms of the model.
- Removal of non-pharmacological psychological support costs.
- One wig assumed in the induction phase for both arms of the model.

The EAG considers that costs of adverse events (AEs) should be included in the EAG's preferred base case. However, as mentioned in Section 4.2.6.1, the EAG was unable to verify the inputs used in the company's AE scenario and was unable to produce an alternative scenario due to a paucity of time. Nonetheless, the EAG requests that during technical engagement, the company provides a more thorough description and justification of their approach to the inclusion of AEs and assumed unit costs to treat each AE and update the scenario if necessary.

Table 40 presents the impact of each assumption on the ICER and Table 41 presents the EAG's deterministic and probabilistic base case results. Table 42 presents the severity subgroup analysis around the EAG base case but it should be noted that probabilistic subgroup results could not be obtained due to a problem with the probabilistic sensitivity analysis (PSA) function in the model.

In the EAG base case probabilistic analysis, an incremental quality-adjusted life-year (QALY) gain of [redacted] over 'discharged from care' along with additional costs of [redacted] for the baricitinib 4 mg, generates an ICER of £423,775 per QALY. The net monetary benefit (NMB) using the £30,000 threshold is [redacted] and the net health benefit (NHB) is [redacted]. The EAG considers that the ICERs are highly sensitive due to the small incremental costs and quality-adjusted life-year (QALY) gain, such that small changes cause a substantial impact.

Additionally, as mentioned in Section 4.2.8.1, the EAG acknowledges that there is a small, but heterogenous, patient population that is more adversely affected in terms of Health-related quality of life (HRQoL) but that the demographics of this population are difficult to identify clinically and consistently, and it is beyond the scope of assessment to identify that group. Nonetheless, the EAG ran two scenarios around the EAG base case and severity subgroup analysis to identify the QALY gain needed for the ICER to reach the £20,000 and £30,000 cost-effectiveness threshold and these are presented in Table 43. The results of the threshold analysis demonstrate that for the overall population, a QALY gain of [redacted] to [redacted] is needed for the ICER to be within the £20,000 to £30,000 threshold. Thus, the EAG advises the committee to consider if the estimated QALY gain needed for baricitinib 4 mg to be cost-effective is plausible for the condition under consideration.

Table 40. EAG's preferred model assumptions - FAS population

Preferred assumption	Section in EAG report	Deterministic ICER (£/QALY)	Cumulative ICER (£/QALY)
Company base case post clarification	**-**	**18,072**	**18,072**
SALT≤20 at Week 36	4.2.5.2	17,071	17,071
SALT≤20 at Week 36 + baseline and CFB utility from BRAVE trials	4.2.5.2 and 4.2.8.1	118,494	118,494
Long-term all-cause discontinuation based on Week 36-52 data for baricitinib 4 mg ([redacted])	4.2.5.2	16,293	107,217
No monitoring costs in the induction phase and Maintenance health state for 'Watch and wait' (comparator defined as 'discharged from care')	4.2.9.5	20,887	126,309
Removal of disease monitoring costs in the BSC health state for both arms of the model	4.2.9.5	63,941	419,926
Removal of non-pharmacological psychological support costs	4.2.9.5	18,679	423,809
One wig assumed in the induction phase for both arms of the model	4.2.9.5	18,068	423,775
EAG preferred base case	**-**	-	**423,775**

Abbreviations: BSC, best supportive care; CFB, change from baseline; EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

Table 41. EAG's base case

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
'Discharged from care'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	423,775
Probabilistic results							
'Discharged from care'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	-	[redacted]	379,030

Abbreviations: ICER, incremental cost effectiveness ratio, LYG, life year gained; QALY, quality adjusted life year.

Table 42. Deterministic scenarios around the EAG base case

	Results per patient	Baricitinib 4 mg	'Discharged from care'	Incremental value
0	**EAG base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			423,775
1	**Severe subgroup - baseline SALT 50-95**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			407,212
2	**Very severe subgroup - baseline SALT 95-100**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			456,573

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

Note: the same baseline utility ([redacted]), change from baseline ([redacted]) and treatment discontinuation rate ([redacted]) have been used for the subgroups as for the base case as the relevant data were not available by severity.

Table 43. Threshold analysis on QALY gain needed for £20,000 to £30,000 cost-effectiveness threshold

Population	QALY gain - £20,000 threshold	QALY gain - £30,000 threshold
Full analysis set - baseline SALT 50-100	[redacted]	[redacted]
Severe subgroup - baseline SALT 50-94	[redacted]	[redacted]
Very severe subgroup - baseline SALT 95-100	[redacted]	[redacted]

Abbreviations: EAG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

6.5 Conclusions of the cost effectiveness sections

Generally, the EAG considers the company's submitted cost-effectiveness analysis adheres to the decision problem defined in the NICE final scope. However, for the comparator (which is listed in the final scope as established clinical management without baricitinib 4 mg), current treatment of severe alopecia areata is variable across the NHS and clinician dependent as many treatments are ineffective and there is use of off label medicines. Additionally, the EAG's clinical experts considered that a significant proportion of patients may not take up treatment (effectively discharged from care) and instead opt for using wigs to manage their hair loss. The EAG agrees with the company that the no active treatment component of "Watch and wait" is a common management strategy used for adults with severe AA. However, the EAG's clinical experts considered that patients would not be regularly monitored if they are not on treatment. As such, the EAG considers that 'discharged from care' is the relevant comparator for the analysis. The EAG's preferred definition for the comparator has implications for costs included in the model as monitoring costs in the induction phase and Maintenance health states are no longer relevant. Additionally, the EAG's clinical experts considered that if patients do achieve a sufficient treatment response on any treatment (including Janus Kinase [JAK] inhibitors), they are unlikely to engage with further care. As such, much of the costs included in the BSC health state is likely to not be incurred by patients.

As such, the EAG considers that a true reflection of the cost-effectiveness of baricitinib 4 mg is a comparison where, in the absence of baricitinib 4 mg, patients manage their hair loss with the use of wigs and orthotics and for a small proportion of patients, antidepressants are required to manage the psychological burden of severe AA.

Additionally, the EAG considers (based on advice from its clinical experts) that patients value an absolute change in scalp hair regrowth and a relative change from their baseline hair loss (as measure by SALT scores) may still require hair removal or use of wigs due to patchy regrowth. As such, the use of $SALT_{50}$ as the measure of treatment response at Week 36 is not considered clinically meaningful and instead SALT≤20 is the EAG's preferred definition of treatment response and aligns with the primary endpoint of the key baricitinib trials, BRAVE-AA1 and BRAVE-AA2.

The utilities used are a key driver in the model as based on the EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials, patients have a relatively high baseline utility and there is not a substantial increase in health-related quality of life (HRQoL) from achieving a response to treatment (whether that is the company's base case definition of $SALT_{50}$ or the primary endpoint of SALT≤20 in the trials).

The company argue that the EQ-5D is insensitive to changes in the severity of AA and lacked content validity as baseline values were almost the same as UK age- and sex-adjusted general population values. However, the company hasn't supplied sufficient evidence to validate the lack of content validity with the EQ-5D nor has it demonstrated why patients should have a substantial change in their QoL.

The EAG's clinical experts advised that for most patients' HRQoL may only be mildly affected and thus may not be that different to the general population but equally HRQoL is severely affected for a few patients (primarily driven by adverse mental health). Additionally, the EAG's clinical experts advised that overtime, patients may come to terms with their hair loss, while a few may remain distressed about their condition. Thus, the EAG acknowledges that there is a small, but heterogenous, patient population that is more adversely affected in terms of HRQoL but that the demographics of this population are difficult to identify clinically and consistently, and it is beyond the scope of assessment to identify that group. However, the EAG has estimated the QALY gain needed to reach the £20,000 and £30,000 cost-effectiveness thresholds and advises the committee to consider if the estimated QALY gain needed for baricitinib 4 mg to be cost-effective is plausible for the condition under consideration.

7 References

1. Harries M, Macbeth AE, Holmes S, Chiu WS, Gallardo WR, Nijher M, et al. The epidemiology of alopecia areata: a population-based cohort study in UK primary care*. *British Journal of Dermatology* 2022; 186: 257-65.
2. Petukhova L, Cabral RM, Mackay-Wiggan J, Clynes R, Christiano AM. The genetics of alopecia areata: What's new and how will it help our patients? *Dermatol Ther* 2011; 24: 326-36.
3. Darwin E, Hirt PA, Fertig R, Doliner B, Delcanto G, Jimenez JJ. Alopecia Areata: Review of Epidemiology, Clinical Features, Pathogenesis, and New Treatment Options. *Int J Trichology* 2018; 10: 51-60.
4. Pratt CH, King LE, Jr., Messenger AG, Christiano AM, Sundberg JP. Alopecia areata. *Nat Rev Dis Primers* 2017; 3: 17011.
5. Trueb RM, Dias M. Alopecia Areata: a Comprehensive Review of Pathogenesis and Management. *Clin Rev Allergy Immunol* 2018; 54: 68-87.
6. Xing L, Dai Z, Jabbari A, Cerise JE, Higgins CA, Gong W, et al. Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. *Nature medicine* 2014; 20: 1043-9.
7. Fukuyama M, Ito T, Ohyama M. Alopecia areata: Current understanding of the pathophysiology and update on therapeutic approaches, featuring the Japanese Dermatological Association guidelines. *J Dermatol* 2022; 49: 19-36.
8. Triyangkulsri K, Suchonwanit P. Role of janus kinase inhibitors in the treatment of alopecia areata. *Drug design, development and therapy* 2018; 12: 2323-35.
9. Messenger AG, McKillop J, Farrant P, McDonagh AJ, Sladden M. British Association of Dermatologists' guidelines for the management of alopecia areata 2012. *British Journal of Dermatology* 2012; 166: 916-26.
10. Lee S, Kim BJ, Lee YB, Lee W-S. Hair Regrowth Outcomes of Contact Immunotherapy for Patients With Alopecia Areata: A Systematic Review and Meta-analysis. *JAMA Dermatology* 2018; 154: 1145-51.
11. National Institute for Health and Care Excellence. Baricitinib for treating severe alopecia areata - Final scope [10941], 2022. Available from: https://www.nice.org.uk/guidance/gid-ta10941/documents/final-scope. Date accessed: September 2022.
12. King B, Ohyama M, Kwon O, Zlotogorski A, Ko J, Mesinkovska NA, et al. Two Phase 3 Trials of Baricitinib for Alopecia Areata. *The New England journal of medicine* 2022.
13. Adelphi Real World. Alopecia Areata Disease Specific Programme Physician Survey - Survey Materials. 2020.
14. Eli Lilly and Company. Disease Burden, Disease Management, Treatment Patters and Resouce Use in Patients with Alopecia Areata - Protocol. 2022.
15. National Institute for Health and Care Excellence (NICE). Baricitinib for moderate to severe atopic dermatitis: Technology appraisal guidance [TA681]. 2021. Available from: https://www.nice.org.uk/guidance/ta681
16. National Institute for Health and Care Excellence (NICE). Baricitinib for moderate to severe rheumatoid arthritis: Technology appraisal guidance [TA466]. 2017. Available from: https://www.nice.org.uk/guidance/ta466
17. National Institute for Health and Care Excellence. Equality impact assessment – Scoping: Baricitinib for treating severe alopecia areata [ID3979] 2022. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10941/documents. Date accessed.
18. King B, Ko J, Forman S, Ohyama M, Mesinkovska N, Yu G, et al. Efficacy and safety of the oral Janus kinase inhibitor baricitinib in the treatment of adults with alopecia areata: Phase 2 results from a randomized controlled study. *Journal of the American Academy of Dermatology* 2021; 85: 847-53.

19. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; 343: d5928.
20. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. York: Centre for Reviews and Dissemination, University of York; 2008.
21. Gupta AK, Carviel JL, Foley KA, Shear NH, Piraccini BM, Piguet V, et al. Monotherapy for Alopecia Areata: A Systematic Review and Network Meta-Analysis. *Skin Appendage Disord* 2019; 5: 331-7.
22. Lemay KR, Tulloch HE, Pipe AL, Reed JL. Establishing the Minimal Clinically Important Difference for the Hospital Anxiety and Depression Scale in Patients With Cardiovascular Disease. *J Cardiopulm Rehabil Prev* 2019; 39: E6-E11.
23. Wynne SC, Patel S, Barker RE, Jones SE, Walsh JA, Kon SS, et al. Anxiety and depression in bronchiectasis: Response to pulmonary rehabilitation and minimal clinically important difference of the Hospital Anxiety and Depression Scale. *Chron Respir Dis* 2020; 17: 1479973120933292.
24. Canadian Agency for Drugs and Technologies in Health. Clinical Review Report. Dupilmab (Dupixent), 2020. Available from: https://www.cadth.ca/dupilumab-0.
25. Frendl DM, Ware JE, Jr. Patient-reported functional health and well-being outcomes with drug therapy: a systematic review of randomized trials using the SF-36 health survey. *Med Care* 2014; 52: 439-45.
26. Ogura K, Bartelstein MK, Yakoub MA, Nikolic Z, Boland PJ, Healey JH. Minimal clinically important differences in SF-36 global score: Current value in orthopedic oncology. *J Orthop Res* 2021; 39: 2116-23.
27. Both H, Essink-Bot ML, Busschbach J, Nijsten T. Critical review of generic and dermatology-specific health-related quality of life instruments. *J Invest Dermatol* 2007; 127: 2726-39.
28. Shikiar R, Willian MK, Okun MM, Thompson CS, Revicki DA. The validity and responsiveness of three quality of life measures in the assessment of psoriasis patients: results of a phase II study. *Health Qual Life Outcomes* 2006; 4: 71.
29. Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life-a systematic review. *Journal of clinical epidemiology* 2017; 89: 188-98.
30. Secrest AM, Hopkins ZH, Frost ZE, Taliercio VL, Edwards LD, Biber JE, et al. Quality of Life Assessed Using Skindex-16 Scores Among Patients With Acne Receiving Isotretinoin Treatment. *JAMA Dermatol* 2020; 156: 1098-106.
31. Kind P, Hardman G, Macan S. UK population norms for EQ-5D. Discussion paper 172. *University of York* 1999.
32. Rencz F, Gulacsi L, Pentek M, Wikonkal N, Baji P, Brodszky V. Alopecia areata and health-related quality of life: a systematic review and meta-analysis. *Br J Dermatol* 2016; 175: 561-71.
33. Taylor PC, Takeuchi T, Burmester GR, Durez P, Smolen JS, Deberdt W, et al. Safety of baricitinib for the treatment of rheumatoid arthritis over a median of 4.6 and up to 9.3 years of treatment: final results from long-term extension study and integrated database. *Ann Rheum Dis* 2022; 81: 335-43.
34. Jones K, Burns A. PSSRU: Unit Costs of Health and Social Care (2021). 2021. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-of-health-and-social-care-2021/.
35. National Health Service (NHS). 2020/21 National Cost Collection data. 2020. Available from: https://www.england.nhs.uk/national-cost-collection/. Date accessed.
36. Service NH. National Health Service (NHS) Business Services Authority. Drug Tariff (April 2022). 2022.
37. European Medicinces Agency (EMA). Olumiant (baricitinib) SmPC. Available from: https://www.ema.europa.eu/en/documents/product-information/olumiant-epar-product-information_en.pdf. . Date accessed: September 2022.

38. National Institute for Health and Care Excellence. Dupilumab for treating moderate to severe atopic dermatitis [TA534]. 2018. Available from: https://www.nice.org.uk/guidance/ta534.
39. Statistics OfN. Population Estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2020. June 2021. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2020. Date accessed: September 2022.
40. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by age and sex for the UK: Report by the Decision Support Unit., 2022. Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02/DSU%20Age%20based%20utility%20-%20Final%20for%20website.pdf. Date accessed: October 2022.
41. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual, 2022. Available from: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741. Date accessed: September 2022.
42. Ara R, Wailoo AJ. NICE DSU Technical Support Document 12: The use of health state utility values in decision models., 2011. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD12-Utilities-in-modelling-FINAL.pdf. Date accessed.
43. National Institute for Health and Care Excellence. Depression in adults: recognition and management [CG90], 2009. Available from: https://www.nice.org.uk/guidance/cg90.
44. National Institue for Health and Care Excellence. Depression in adults: treatment and management [NG222], 2022. Available from: https://www.nice.org.uk/guidance/ng222. Date accessed: September 2022.
45. Lilly E. NICE Cost-Effectiveness Model for Baricitinib in AA: Model Sanity Checklist. 2022.

8 Appendices

8.1 Additional baseline characteristics

Baseline characteristics reported in the CS but not presented in the main body of the EAG report are reproduced in Table 44.

Table 44. Baseline characteristics reported for BRAVE-AA1 and BRAVE-AA2 not included in Table 11.

Characteristic	BRAVE–AA1		BRAVE–AA2	
	Placebo (N=189)	baricitinib 4 mg (N=281)	Placebo (N=156)	baricitinib 4 mg (N=234)
Mean (SD) age, years	37 (12.91)	36 (13.27)	37 (12.35)	38 (12.65)
Female, n (%)	109 (57.7)	165 (58.7)	98 (62.8)	144 (61.5)
Geographic region, n (%)				
North America	103 (54.5)	153 (54.4)	54 (34.6)	82 (35.0)
Asia	70 (37.0)	107 (38.1)	42 (26.9)	63 (26.9)
Rest of the world	16 (8.5)	21 (7.5)	60 (38.5)	89 (38.0)
Race, n (%)				
White	83 (44.1)	123 (43.9)	85 (54.5)	144 (61.5)
Asian	78 (41.5)	114 (40.7)	51 (32.7)	67 (28.6)
Black or African American	17 (9.0)	28 (10.0)	16 (10.3)	18 (7.7)
American Indian or Alaska Native	8 (4.3)	8 (2.9)	0 (0.0)	0 (0.0)
Native Hawaiian or Other Pacific Islander	1 (0.5)	1 (0.4)	0 (0.0)	0 (0.0)
Mean (SD) BMI, kg/m2	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD) duration since onset of AA, years	12.6 (11.2)	11.8 (11.1)	11.79 (10.190)	11.89 (11.122)
Age of AA onset, n (%)				
<18 years	[redacted]	[redacted]	57 (36.5)	74 (31.6)
≥18 years	[redacted]	[redacted]	99 (63.5)	160 (68.4)
Patients with AU, n (%)	74 (39.2)	127 (45.2)	66 (42.3)	111 (47.4)
ClinRO for eyebrow hair loss, n (%)				
2	53 (28.3)	73 (26.3)	46 (30.1)	49 (21.0)
3	71 (38.0)	115 (41.4)	66 (43.1)	112 (48.1)
ClinRO for eyelash hair loss, n (%)				
2	38 (20.3)	74 (26.6)	31 (20.3)	43 (18.5)
3	58 (31.0)	93 (33.5)	59 (38.6)	97 (41.6)

PRO for Scalp Hair Assessment				
3 (50–94% hair loss)	72 (38.1)	102 (36.4)	60 (38.5)	78 (33.3)
4 (95–100% hair loss)	109 (57.7)	173 (61.8)	91 (58.3)	137 (58.5)

Abbreviations: AA: alopecia areata; AU: alopecia universalis; ClinRO; Clinician reported outcome; PRO: patient reported outcome SD: standard deviation

8.2 SALT responder statistical analyses

Table 45 provides the detailed statistical analysis, including number of responders and differences, odds ratios and p-values versus placebo for baricitinib 4 mg in BRAVE-AA1 and BRAVE-AA2 at Week 36 for SALT ≤20. Table 46 provides these data for SALT ≤10, and Table 47 for the relative treatment response outcomes, $SALT_{50}$ and $SALT_{75}$.

Table 45. SALT ≤20 response at Week 36 for BRAVE-AA1 and BRAVE-AA2 (FAS)

SALT ≤20, Week 36	BRAVE-AA1		BRAVE-AA2	
	Placebo (N=189)	4 mg baricitinib (N=281)	Placebo (N=189)	4 mg baricitinib (N=281)
Response, n (%) (95% CI)	10 (5.3) (2.9, 9.5)	99 (35.2) (29.9, 41.0)	4 (2.6) (1.0, 6.4)	76 (32.5) (26.8, 38.7)
Difference (95% CI) vs placebo	N/A	29.9 (23.2, 36.2)	NA	29.9 (23.1, 36.3)
Odds ratio (95% CI) vs placebo	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs placebo	N/A	<0.001	NA	<0.001

Source: CS Table 15

Abbreviations: AA: alopecia areata; CI: confidence interval; FAS: full analysis set; SALT: severity of alopecia tool

Table 46. SALT ≤10 response at Week 36 for BRAVE-AA1 and BRAVE-AA2 (FAS)

SALT ≤10, Week 36	BRAVE-AA1		BRAVE-AA2	
	Placebo (N=189)	4 mg baricitinib (N=281)	Placebo (N=189)	4 mg baricitinib (N=281)
Response, n (%) (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs placebo	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio (95% CI) vs placebo	[redacted]	[redacted]	[redacted]	[redacted]
p-value vs placebo	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 22

Abbreviations: AA: alopecia areata; CI: confidence interval; FAS: full analysis set; SALT: severity of alopecia tool

Table 47. $SALT_{50}$ and $SALT_{75}$ response at Week 36 for BRAVE-AA1 and BRAVE-AA2 (FAS)

Week 36	BRAVE-AA1		BRAVE-AA2	
	Placebo (N=189)	4 mg baricitinib (N=281)	Placebo (N=156)	4 mg baricitinib (N=234)
$SALT_{50}$				
Response, n (%) (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO		[redacted]		[redacted]
Odds ratio (95% CI) vs PBO		[redacted]		[redacted]
p-value vs placebo		[redacted]		[redacted]
$SALT_{75}$				
Response, n (%) (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Difference (95% CI) vs PBO		[redacted]		[redacted]
Odds ratio (95% CI) vs PBO		[redacted]		[redacted]
p-value vs placebo		[redacted]		[redacted]

Abbreviations: AA: alopecia areata; CI: confidence interval; FAS: full analysis set; SALT: severity of alopecia tool

8.3 Atopic background subgroup analysis

Table 48 provides SALT ≤20 response data for the baricitinib 4 mg and placebo arms of BRAVE-AA1 and BRAVE-AA2 by atopic background status (no atopic background and atopic background categories). For the baricitinib 4 mg arm, a larger proportion of patients in the atopic background category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]) than those in the no atopic background category (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]), a pattern also apparent in the placebo arm data (no atopic background category: BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]; atopic background category; BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]).

Table 48. SALT ≤20 response of BRAVE-AA1 and BRAVE-AA2 by atopic background status at baseline

	Atopic background	Week 36 SALT ≤20 response rate	
		BRAVE-AA1	BRAVE-AA2
Baricitinib 4 mg	No atopic background	[redacted]	[redacted]
	Atopic background	[redacted]	[redacted]

Placebo	No atopic background	[redacted]	[redacted]
	Atopic background	[redacted]	[redacted]

Source: Clarification questions Table 6 and Table 7

Abbreviations: AA: alopecia areata; SALT: severity of alopecia tool

8.4 Source of the confidential prices used in the confidential appendix

Table 49. Source of the confidential prices used in the confidential appendix

Treatment	Source of price/type of commercial arrangement
Ciclosporin	CMU
Methotrexate	List price
Azathioprine	CMU
Intralesional steroids (triamcinolone acetonide)	List price
DPCP treatment	NHSE
Prednisolone	eMIT
Prednisolone	eMIT
TCS: Mometasone ointment	eMIT
Minoxidil 5% foam (topical)	List price
Minoxidil tablets	List price
Mycophenolate Mofetil	List price
Anthralin / dithranol 0.1% cream	List price
Sertraline	List price
Escitalopram	eMIT
Duloxetine	eMIT

Abbreviations: CMU, confidential medicines unit; DPCP, diphenylcyclopropenone; contact immunotherapy; eMIT, Drugs and pharmaceutical electronic market information tool; NHSE, National Health Service England.

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on Thursday 20 October 2022** 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.

Please underline all confidential information, and separately highlight information that is submitted as 'commercial in confidence' in turquoise, all information submitted as 'academic in confidence' in yellow, and all information submitted as 'depersonalised data' in pink.

Section 1: Major Issues

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 31 states: "…*and a less effective 2 mg dose may be used in patients >75 years, leading to a lower probability of treatment response*."	Please amend this wording to "…and a ~~less effective~~ 2 mg dose may be used in patients >75 years, **which may lead** ~~leading~~ to a lower probability of treatment response"	The BRAVE-AA trials did not enrol patients >75 years, so it is inaccurate to draw this conclusion about the efficacy of 2 mg baricitinib in these patients, especially given that the lower efficacy observed in the 2 mg arm of the clinical trials (in younger patients) may not be applicable due to the reduced renal function likely to be present in these older patients and the subsequent increased exposure to baricitinib due to reduced renal elimination.	This is not a factual error. No change required.
Page 73 states: "*The treatment response of $SALT_{50}$ is a relative measure of response and is defined by the company as at least a 50% improvement from baseline SALT score."*	Please provide some additional context as follows: "The treatment response of $SALT_{50}$ is a relative measure of response and is defined by the company as at least a 50% improvement from baseline SALT score. **A relative improvement from baseline score utilised in the model**	While the company accepts that the EAG's preference for defining treatment response may differ to that of the company, the company considers it important to acknowledge that the approach used in the	This is not a factual error. No change required.

	as the definition of response is aligned with previous appraisals in other dermatological indications, including atopic dermatitis ([TA681] [TA466]) and psoriasis [TA350]."	company submission (CS) is aligned with the modelling precedent in other dermatology indications, including atopic dermatitis and psoriasis.	

Section 2: Minor comments

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 18 and Page 80 state: "However, *the company did not provide the overall proportion of patients reporting a score of perfect health from the Adelphi DSP study."*	Please amend to "However, the company **was not able to** provide the overall proportion of patients reporting a score of perfect health from the Adelphi DSP study **as these data were not available to the company at the time of submission**."	While the company acknowledges that they did not provide the overall proportion of patients reporting a score of perfect health from the Adelphi DSP study, the company considers it important to provide additional context that these data were not provided as they were not available to the company at the time of submission, rather than being omitted by choice. These data will be	This is not a factual error. No change required.

		provided during Technical Engagement.	
Page 20 states: *"Current AA episode duration and baseline SALT score are clinically meaningful variables that predict treatment response and vary substantially in the trial."*	Please clarify whether the EAG are claiming that current AA episode duration and baseline SALT score are clinically meaningful variables within the BRAVE-AA trials specifically or are established in the literature as treatment effect modifiers generally in patients with AA.	It is currently unclear whether the EAG are stating that current AA episode duration and baseline SALT score are clinically meaningful variables within the BRAVE-AA trials specifically or more broadly in patients with AA.	This is not a factual error. No change required.
Page 45 states: "…; *a mean disease duration of around 12 years; a mean episode duration of around 3.5 years; >50% had SALT 95-100 and around 40% had alopecia universalis."*	Please amend to "….; a mean **duration from the first onset of AA diagnosis of 12.2** years; a mean episode duration of **3.9** years; >50% had SALT 95-100 and [redacted] had alopecia universalis."	Minor amendments to improve the accuracy of the report.	Thank you for providing the exact figures. The EAG report has been updated.
Page 54 states: *"For the baricitinib 4 mg arm, a larger proportion of patients in the <4 years category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted])*	Please update to "For the baricitinib 4 mg arm, a larger proportion of patients in the <4 years category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]) than those in the ≥4 years category (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]), a pattern **is** also apparent in the placebo arm data (<4 years category: BRAVE-AA1, [redacted];	Typographical errors.	Thank you for highlighting this. The EAG report has been updated.

<table>
<tr>
<td>than those in the ≥4 years category (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]), a pattern also apparent in the placebo arm data (<4 years category: BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]; ≥4 years category; BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]).”</td>
<td>BRAVE-AA2, [redacted]; ≥4 years category; BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]).”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 54; Table 15. (SALT ≤20 response of BRAVE-AA1 and BRAVE-AA2 by the duration of current AA episode at baseline [<4 years and ≥4 years categories])</td>
<td>Please amend the table as follows:
<table>
<tr><th></th><th>Duration of current AA episode at baseline</th><th colspan="2">Week 36 SALT ≤20 response rate</th></tr>
<tr><th></th><th></th><th>BRAVE-AA1</th><th>BRAVE-AA2</th></tr>
<tr><td rowspan="2">Baricitinib 4 mg</td><td><4 years</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>≥4 years</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td rowspan="2">Placebo</td><td><4 years</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>≥4 years</td><td>[redacted]</td><td>[redacted]</td></tr>
</table></td>
<td>Typographical errors.</td>
<td>Thank you for highlighting this. The EAG report has been updated.</td>
</tr>
<tr>
<td>Page 55 states: “…re-randomised to placebo at Week 56.” and “…baricitinib arm were</td>
<td>Page 55: Please amend to “…re-randomised to placebo at Week 52.” and “…baricitinib arm were re-randomised to 2 mg baricitinib at Week 52.”</td>
<td>Typographical errors.</td>
<td>Thank you for highlighting this. The EAG report has been updated.</td>
</tr>
</table>

<table>
<tr>
<td>re-randomised to 2 mg baricitinib at Week 56.”

Page 64 states: “The small amount of data available at Week 56 and Week 72 suggests that..”</td>
<td>Page 64: Please amend to “The small amount of data available at Week 52 and Week 72 suggests that..”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 57 states: “A larger improvement was observed in the Skindex-36 scale…”</td>
<td>Please update to “A larger improvement was observed in the Skindex-16 scale…”</td>
<td>Typographical error.</td>
<td>Thank you for highlighting this. The EAG report has been updated.</td>
</tr>
<tr>
<td>Page 77–78; Table 28 (AEs and costs included in the company scenario analysis)</td>
<td>Please amend the Total Cost row in the table using the values presented below:
<table>
<tr><th></th><th>Induction</th><th>Maintenance ($SALT_{50}$)</th><th>Maintenance ($SALT_{75}$)</th></tr>
<tr><td>Baricitinib 4mg</td><td>£18.94</td><td>£24.90</td><td>£60.49</td></tr>
<tr><td>Watch and Wait</td><td>£15.08</td><td>£5.87</td><td>£3.07</td></tr>
</table>
</td>
<td>The Company were unable to replicate the values in the Total Costs row within Table 28 in the EAG report. Furthermore, it should be noted that for Maintenance, the AE costs for baricitinib and ‘watch and wait’ are disaggregated by response level ($SALT_{50}$ and $SALT_{75}$) in the company base case. As requested, a more thorough description of the approach to the inclusion of AEs in the model will be provided during Technical Engagement.</td>
<td>Not a factual error. Please refer to the company’s instructions on how to run the AE scenario in Table 45 of the company clarification response. Once the company’s instructions are followed, the total costs presented in Table 28 can be found in cells F255:I255 in the ‘Treatment costs’ tab of the model.</td>
</tr>
</table>

			Additionally, the scenario the company provided, no instructions were given in terms of disaggregated costs for the different SALT categories, as no disaggregated AE data were provided. Thus the EAG considers the company's approach presented in this response may be an update to what was provided during the clarification stage.
Page 83 states: "*There is currently a patient access scheme (PAS) in place for baricitinib such that the fixed price pack is [redacted] (discount of [redacted]).*"	Page 83: Please amend to "There is currently a patient access scheme (PAS) in place for baricitinib such that the fixed price pack is [redacted] ~~(discount of [redacted])~~."	The legal construction of the PAS price of baricitinib is as a fixed pack price that is not dependent on a percentage discount. As such, only the fixed unit price should be referred to in this appraisal.	Thank you for highlighting this distinction. The EAG report has been updated.

Page 86 states: *"Table 34 presents the disease management costs applied in the BSC health state. Unit costs were sourced from NHS reference costs 2020/21 and PSSRU.[34, 35] The total annual cost of disease management in the BSC health state was estimated to be £354.20."*	Please amend to "Table 34 presents the disease management costs applied in the BSC health state **post clarification**. Unit costs were sourced from NHS reference costs 2020/21 and PSSRU.[34, 35] The total annual cost of disease management in the BSC health state was estimated to be **£468.47 post clarification**."	The annual costs of disease management in the BSC health state were updated post clarification, so the company considers it important to clarify this in the text and to provide the updated values.	Thank you for highlighting this error. The EAG report has been updated.
Page 89 states: "*Additionally, the EAG requested a scenario where disease monitoring costs in the BSC health state are removed from both arms of the model. However, the company only provided a scenario where disease monitoring costs in the BSC were excluded for the baricitinib 4 mg arm only*."	Please amend the wording to: "Additionally, the EAG requested a scenario where disease management costs are removed ~~from both arms of the model~~. ~~However~~ **Therefore**, the company ~~only~~ provided a scenario where disease monitoring costs in the BSC were excluded for the baricitinib 4 mg arm ~~only~~ **as the EAG's clinical experts advised that if a patient's condition does not adequately respond to a Janus Kinase (JAK) inhibitor, they would not be given any further treatment**.	The clarification question referred to within this section was phrased "Please provide a scenario where disease management (not including wig use and orthotics) and drug acquisition and monitoring costs in the BSC health state are removed" and did not specify whether this was from the treatment arm or comparator arm of the model. The wording of the question suggests that the EAG's clinical experts stated	This is not a factual error. No change required.

<table>
<tr><td></td><td></td><td>that BSC would not be used following non-response to treatment. This indicates that they may be referring to the anticipated treatment pathway following the introduction of JAK-inhibitors. The company therefore considers it important to update this text to reflect the ambiguity of the clarification question rather than suggesting the scenario provided by the company was incorrect.</td><td></td></tr>
<tr><td>Page 95; Table 38. (Company scenario analyses – deterministic.)</td><td>Please amend the table as follows:<table><tr><th colspan="4">Response based on SALT≤10</th></tr><tr><td>Total costs (£)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td>QALYs</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td colspan="3">ICER (£/QALY)</td><td>£16,490</td></tr></table></td><td>Typographical errors.</td><td>This is not a factual error. Please refer to the company instructions to run the SALT10 scenario in the model (cell H16 in the ‘Main’ tab). After selecting SALT10 in cell D16 of the ‘Main’ tab, SALT 75 in cell D14 of the ‘Main’ tab needs to be reselected.</td></tr>
</table>

Page 98 states: "*Results reported include the company's proposed patient access scheme (PAS) of [redacted]."*	Please amend to "Results reported include the company's proposed patient access scheme (PAS); **a fixed pack price of [redacted]**".	The legal construction of the PAS price of baricitinib is as a fixed pack price that is not dependent on a percentage discount. As such, only the fixed unit price should be referred to in this appraisal.	Thank you for highlighting this distinction. The EAG report has been updated.
Page 110 states: "*For the baricitinib 4 mg arm, a larger proportion of patients in the <4 years category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]) than those in the ≥4 years category (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]), a pattern also apparent in the placebo arm data (<4 years category: BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]; ≥4 years category; BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]).*"	Please amend to "For the baricitinib 4 mg arm, a larger proportion of patients in the <4 years category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]) than those in the ≥4 years category (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]), a pattern also apparent in the placebo arm data (<4 years category: BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]; ≥4 years category; BRAVE-AA1, [redacted]**;** BRAVE-AA2, [redacted])."	Typographical errors.	Thank you for highlighting this. The EAG report incorrectly refers to the baseline duration of current AA episode data here rather than the atopic background subgroup analysis. The wording has been updated to, "Table 48 provides SALT ≤20 response data for the baricitinib 4 mg and placebo arms of BRAVE-AA1 and BRAVE-AA2 by atopic background status (no atopic background and atopic background categories). For the

			baricitinib 4 mg arm, a larger proportion of patients in the atopic background category achieved SALT ≤20 at Week 36 (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]) than those in the no atopic background category (BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]), a pattern also apparent in the placebo arm data (no atopic background category: BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]; atopic background category; BRAVE-AA1, [redacted]; BRAVE-AA2, [redacted]).” The column header in Table 48 has also been updated to, “Atopic background” instead of “Duration of current AA episode at

			baseline", and the reporting of confidence intervals in the table has been changed to the format (XX to XX) from (XX, XX) to be in-line with the rest of the EAG report.

Section 3: Confidentiality Highlighting Amendments

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG response
Page 21–22; Table 7. (Deterministic scenarios around the EAG base case.)	The total costs (£) and QALY values for the "baricitinib 4 mg", "'Discharged from care'" and "incremental value" columns are commercially confidential.	Please amend the highlighting in the table as follows:	Thank you for highlighting this error. The EAG report has been updated.

	Results per patient	Baricitinib 4 mg	'Discharged from care'	Incremental value
0	**EAG base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			423,775

<table>
<tr><td></td><td></td><td>
<table>
<tr><td>1</td><td colspan="4">Severe subgroup - baseline SALT 50-95</td></tr>
<tr><td></td><td>Total costs (£)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td></td><td>QALYs</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td></td><td colspan="3">ICER (£/QALY)</td><td>407,212</td></tr>
<tr><td>2</td><td colspan="4">Very severe subgroup - baseline SALT 95-100</td></tr>
<tr><td></td><td>Total costs (£)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td></td><td>QALYs</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td></td><td colspan="3">ICER (£/QALY)</td><td>456,573</td></tr>
</table>
</td><td></td></tr>
<tr><td>Page 50</td><td>Pooled data for BRAVE-AA1 and BRAVE-AA2 are unpublished, so are academic in confidence.</td><td>Please amend the highlighting to: “At Week 52, [redacted] of patients in the baricitinib 4 mg had achieved SALT ≤20, an increase of [redacted] from Week 36 ([redacted] had achieved SALT ≤20 at Week 36).</td><td>Thank you for highlighting this error. The EAG report has been updated.</td></tr>
<tr><td>Page 51</td><td>Week 52 data for BRAVE-AA1 and BRAVE-AA2 are unpublished, so are academic in confidence.</td><td>Please amend the highlighting to: “Of the [redacted] responders at Week 52 who were re-randomised to stay on baricitinib 4 mg in BRAVE-AA1, [redacted] maintained their SALT ≤20 response at Week 76. In BRAVE-AA2, of the [redacted] responders at Week 52…”</td><td>Thank you for highlighting this error. The EAG report has been updated.</td></tr>
</table>

Page 55	Week 52 data for BRAVE-AA1 and BRAVE-AA2 are unpublished, so are academic in confidence.	Please amend the highlighting to: "[redacted] SALT ≤20 responders in the BRAVE-AA2 mg baricitinib arm were re-randomised…"	Thank you for highlighting this error. The EAG report has been updated.
Page 82	Pooled data for BRAVE-AA1 and BRAVE-AA2 are unpublished, so are academic in confidence.	Please amend highlighting to: "Upon request, the company also supplied the change from baseline at Week 36 for patients achieving SALT≤10 based on pooled EQ-5D data from the BRAVE-AA1 and BRAVE-AA2 trials ([redacted])".	Thank you for highlighting this error. The EAG report has been updated.
Page 83	Data from the Adelphi DSP are unpublished, so are academic in confidence.	Please update the highlighting to: "Only data obtained from physicians treating severe and very severe patients in the UK ([redacted]) were used for the model."	Thank you for highlighting this error. The EAG report has been updated.

Section 4: Minor Typographical and Grammatical Errors

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**EAG response**
Page 20 states: "*This may cause the trials to underestimate treatment effectives in newly diagnosed severe AA patients…"*	Please amend to "This may cause the trials to underestimate treatment **effectiveness** in newly diagnosed severe AA patients…"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.

Page 32 states: *"the BRAVE-AA trial participants having varying, and often long, lengths of current AA episode at baseline, with mean durations >3.5 years for all arms."*	Please amend this wording to "the BRAVE-AA trial participants having varying, and often long, lengths of current AA episode at baseline, with mean durations **≥**3.5 years for all arms."	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.
Page 24 states: "*In the CS, Baricitinib is positioned as: i) a first-line treatment for severe AA, and ii) a later-line treatment to treat patients who do not respond…"*	Please amend to "In the CS, Baricitinib is positioned as: i) a first-line treatment for severe AA, and ii) a later-line treatment to treat patients **with severe AA** who do not respond…"	Minor typographical error.	Thank you for highlighting this. The EAG report has been updated.
Page 44 states: "*However, the BRAVE-AA2 placebo arm having a higher mean duration of current AA episode at baseline (4.68 years) than the baricitinib 4 mg arm…"*	Please update to "However, the BRAVE-AA2 placebo arm **has** a higher mean duration of current AA episode at baseline (4.68 years) than the baricitinib 4 mg arm..."	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated to, "However, the BRAVE-AA2 placebo arm **had** a higher mean duration of current AA episode at baseline (4.68 years) than the baricitinib 4 mg arm..."

Page 55 states: *"the PRO Scalp Hair Assessment and the ClinRO measure for eyelash and eyebrow, in Table 16 and Table 17 of the CS."*	Please update to "the PRO Scalp Hair Assessment and the ClinRO measure for eyelash and eyebrow **regrowth**, in Table 16 and Table 17 of the CS."	Minor typographical error.	Thank you for highlighting this. The EAG report has been updated.
Page 55; Table 16. (EQ-5D data from BRAVE-AA1 and BRAVE-AA2 at baseline and at Week 36.)	Please amend the wording in the second row to "Health state index UK, mean (**SD**)"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated. The EAG also updated the wording in the fifth row to VAS, mean (**SD**) in line with pages 267 and 272 of the BRAVE-AA1 CSR. The abbreviations footer has been updated to include "SD: standard deviation" instead of "SE: standard error".
Page 58 states: *"The EAG does not find these arguments convincing and provides comments on them in Table 18, although the EAG's clinical experts*	Please amend wording to: "The EAG does not find these arguments convincing and provides comments on them in Table 18, although the EAG's clinical experts ~~*The EAG's clinical experts*~~ did note…."	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.

The EAG's clinical experts did note…"			
Page 58 states "*highlighting how the majority of benefits will be most visible in psychosocial functioning with the EQ-5D only captures partially."*	Please update to "highlighting how the majority of benefits will be most visible in psychosocial functioning **domain which** the EQ-5D only captures partially."	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated to, "highlighting how the majority of benefits will be most visible in psychosocial functioning**, which** the EQ-5D only captures partially."
Page 61 states: "*This does may be used for patients:…"*	Please update to "This **dose** may be used for patients:…"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.
Page 71 states: "*The aim of the model developed by the company was to estimate the treatment pathways for patients"*	Please update to "The aim of the model developed by the company was to estimate the treatment **pathway** for patients"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.
Page 79 states: "*Utilities in the model in the model were adjusted for age"*	Please amend wording to "Utilities ~~in the model~~ in the model were adjusted for age"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.

Page 84 states: "*Confidential medicines unit (CMU) and Drugs and pharmaceutical electronic market information tool (eMIT) prices are available for medicines included in BSC and such the EAG has produced…"*	Please amend to "Confidential medicines unit (CMU) and Drugs and pharmaceutical electronic market information tool (eMIT) prices are available for medicines included in BSC**, as such**, the EAG has produced…"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.
Page 84 states: "*The company also included costs of monitoring patients while on treatment in the BSC (Table 33)…*"	Please amended wording to "The company also included costs of monitoring patients while on treatment in ~~the~~ BSC (Table 33)…"	Minor typographical error.	Thank you for highlighting this error. The EAG report has been updated.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

B.1 Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

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.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

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 redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Thursday 1 December 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

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.

Technical engagement response form

About you

Table 1 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
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	N/A

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2. Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Issue 1: Definition of the comparator Which of the following is considered to be standard of care in the treatment of severe alopecia areata? • 'Watch and wait' comprising no active treatment and frequent monitoring? • Treatment with	No	Based on clinical expert opinion, the EAG proposes that 'discharge from care' is the most appropriate comparator for baricitinib in severe alopecia areata (AA). Consequently, the EAG propose that the monitoring costs included in the Induction and Maintenance health states should be removed from the base case cost-effectiveness analysis. In the Company Submission (CS), established clinical management is defined as 'watch and wait', in which patients receive no treatment, followed by best supportive care (BSC) after non-response. The Company stated that 'watch and wait' reflects the current treatment pathway and experts consider this approach to be a legitimate option for many patients with severe AA, the indication of relevance to this submission.[1] Based on clinical expert opinion sought by the Company, patients undergoing 'watch and wait' were modelled to incur disease monitoring costs such as dermatologist outpatient consultations. However, it is plausible that patients not currently receiving active treatment may not be monitored and, given that removal of monitoring costs from the comparator arm has minimal impact on the incremental cost-effectiveness ratio (ICER), the Company is content to accept the removal of monitoring costs included in the Induction and Maintenance health states for the comparator arm.

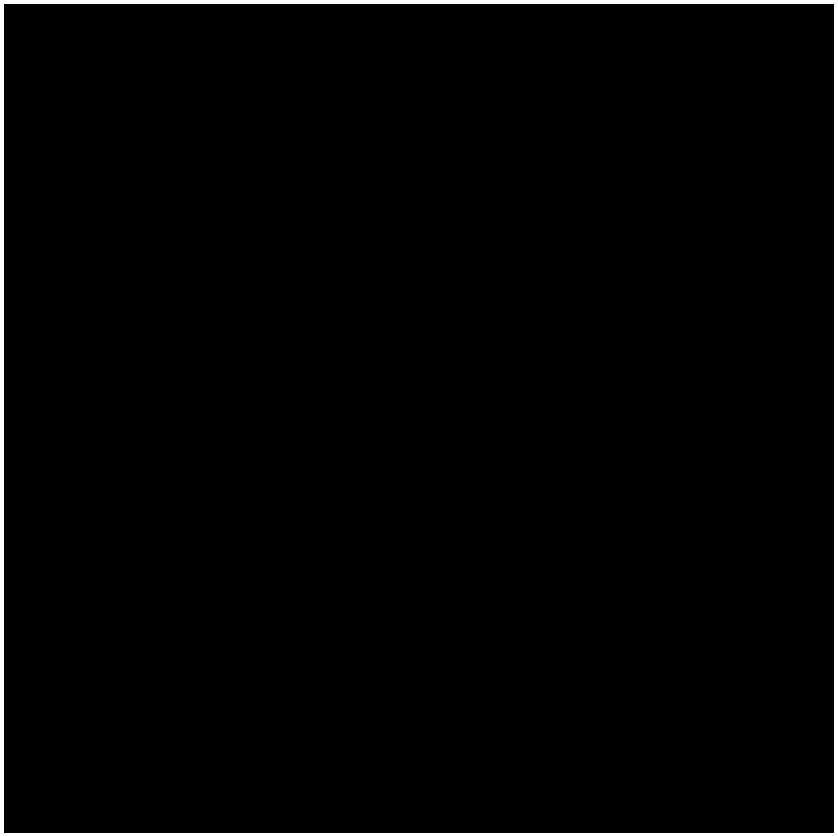
diphenylcyclopropenone (DPCP)? • Treatment with a basket of low-effectiveness non-DPCP therapies including systemic corticosteroids and immunosuppressants? • No active treatment and discharge from care?		However, the Company do not accept the EAG's proposed terminology for the comparator, since the use of 'discharged from care' implies that patients with severe AA do not engage further with their dermatologist and will not continue to seek treatment for their condition, even while their severe AA persists. While this may be accurate for a proportion of patients who have exhausted all available BSC treatment options, the Company do not consider this to be an accurate reflection of the treatment pathway for a proportion of patients with severe AA, as discussed further in response to Issue 4. Given this, the Company would like to propose that the comparator arm is referred to as 'No Active Treatment' to reflect the fact that patients in the comparator arm do not incur monitoring costs during induction phase or when they are not receiving treatment, but may engage subsequently with BSC therapies if they do not experience spontaneous remission. Subject to the alternative terminology proposed by the Company for the comparator arm, the updated base case results following the removal of monitoring costs are presented in Table 5. Regarding the other comparators listed in the questions for Technical Engagement (DPCP and a basket of non-DPCP therapies), the Company agrees with the overall conclusions drawn by the EAG and therefore considers these to be unsuitable comparators for baricitinib in this indication: • There is a lack of robust evidence to support the efficacy of DPCP in this indication, the quality of published efficacy evidence for DPCP is generally low with some studies not accounting for non-completers and/or discontinuations in efficacy results, and available evidence demonstrates a high rate of relapse both during treatment and following discontinuation.[2] In addition, this treatment is generally not well-tolerated and is not widely available in the UK, resulting in inequitable access.[1] • A basket of off-label non-DPCP treatments may be an appropriate comparator for some newly diagnosed patients with severe AA. However, the role of these treatments in the current pathway of care is very unclear with no apparent treatment pattern and comparisons with these treatments would also be associated with significant uncertainty due to the paucity of data.
Issue 2: Definition of treatment response at Week 36	No	The EAG has proposed that SALT≤20 is aligned with the primary endpoint from the BRAVE-AA trials and represents a more clinically meaningful benefit for patients than $SALT_{50}$ and therefore, SALT≤20 should be used as the definition of treatment response in the Induction period in the cost-effectiveness model.

What is considered a clinically meaningful change in the Severity of Alopecia Tool (SALT)? • SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair) • $SALT_{50}$: at least a 50% improvement from baseline SALT score • $SALT_{75}$: at least a 75% improvement from baseline SALT score		In the CS, $SALT_{50}$ was used as the definition of response in the base case, though the $SALT_{75}$ response was also used in the model when $SALT_{50}$ was selected in order to obtain a more granular calculation of the total QALYs. To the Company's knowledge, there is no universally accepted definition of what constitutes a 'clinically meaningful change in SALT' but some AA experts have considered a 50% improvement from baseline in SALT score as a reasonable efficacy target for an AA treatment.[3] Therefore, to align with clinical expert opinion, $SALT_{50}$ was deemed to be the most appropriate definition of treatment response.[3] In addition, the Company maintains that the use of SALT≤20 response is at risk of being overly restrictive, given that concomitant pattern baldness may lead to a ceiling on the maximum possible response in some patients, and that continued hair regrowth was observed in the BRAVE-AA trials in some patients who failed to meet SALT≤20 at Week 36, as demonstrated by the Week 52 data presented in Section B.2.8.2 of the CS. However, since the EAG's proposed amendments minimally impact the ICER, the Company is content to accept the EAG's proposed amendment to update the definition of response to SALT≤20 in order to resolve this issue, subject to the caveats of its failure to account for the issues associated with gradual hair regrowth and concomitant pattern baldness. The base case results from this amendment are presented in Table 5.
Issue 3: Source of utilities in the model • For the majority of people with	Yes	The EAG raised concerns that the utility values derived from the Adelphi Disease Specific Programme (DSP) were subject to the same limitations of those derived in the BRAVE-AA trials, i.e. the observed ceiling effect, and that the Company had not provided sufficient evidence to validate the lack of content validity of the European Quality of Life-5 Dimensions (EQ-5D) instrument for measuring health-related quality of life (HRQoL) in patients with severe alopecia areata (AA). Therefore, the EAG proposes that

severe alopecia areata, would their health-related quality of life be similar to the general population? • Which utility value best represents health-related quality of life for people with severe alopecia areata during the induction phase of treatment? Value from Adelphi DSP study (EAR, Table 29) or value using pooled data from BRAVE-AA1 and BRAVE-AA2		utilities derived from the pooled BRAVE-AA trial EQ-5D data should be used in the base case cost-effectiveness analysis. Despite being subject to the content validity issues of the EQ-5D instrument and still exhibiting some degree of associated ceiling effect, the Company maintains their position that the utility values derived from the Adelphi DSP better represent the HRQoL for people with severe AA than the BRAVE-AA trial-derived values, while also still aligning with the NICE reference case.[4] Considering the proportion of patients in 'perfect health' at baseline in the BRAVE-AA trials versus the Adelphi DSP, it is clear that the ceiling effect is observed to a greater extent in the BRAVE-AA trial data due to the fact that the HRQoL of patients in the BRAVE-AA trial is not representative of the population with severe AA in UK clinical practice. Therefore, the Company considers that the Adelphi DSP EQ-5D values are a more appropriate source of utilities for use in the base case. The reasons for this are outlined in detail below. **Evidence supporting the HRQoL impairment associated with AA** Studies have clearly established that severe AA has a profound psychological and psychosocial impact on patients, and, in the majority of patients, this translates into a significant impairment in HRQoL compared to the general population. The systematic review and meta-analysis conducted by Rencz *et al.* (2016) identified 479 studies, of which 21 met the inclusion criteria.[5] The study designs included cross-sectional designs (n=12), case control studies (n=5), a prospective cohort study (n=1), a retrospective cohort study (n=1), a randomised controlled trial (RCT [n=1]) and a non-RCT (n=1). The analysis demonstrated that AA was generally associated with a substantial negative effect on HRQoL, as measured by various general, dermatology-, hair-, scalp- and AA-specific HRQoL measures.[5] This meta-analysis also demonstrated that SF-36 outcomes were significantly poorer among AA patients than the general population, particularly across the "role emotional", "mental health" and "vitality" domains.[5] A multicentre, observational, cross-sectional study conducted by Balieva *et al.* across 13 European countries, measured the HRQoL of 5,369 participants (4,010 patients and 1,359 controls) and, similarly, reported the greatest impact on the depression/anxiety domain of the EQ-5D instrument among patients with AA, leading to an overall reduction in EQ-5D VAS scores compared to healthy controls (69.7±18.1 versus 82.2±15.5). Accordingly, it was identified that patients with AA had a fourfold risk of anxiety/depression versus healthy controls,

(EAR, Table 30)?		compared to a threefold and twofold risk for urticaria and atopic dermatitis versus healthy controls, respectively.[6] Similarly, an international multicentre, observational and case-control study conducted by Titeca *et al*. (2020) found patients with AA to experience greater HRQoL impairments due to increased levels of anxiety and depression, as measured by the EQ-5D and Hospital Anxiety Depression Scale (HADS) instruments.[7] Compared to controls (n=1,359), patients with AA (n=37) scored significantly higher on the HADS-A (7.9 versus 5.6) and HADS-D (5.4 versus 3.6) and, according to the EQ-5D instrument, lower scores were observed for patients with AA across the activity, pain, and anxiety and depression dimensions.[7] Qualitative investigations among patients with AA further highlight the substantial proportion of patients that experience emotional and functional impairments as a result of the symptoms of AA.[8] In interviews among participants with severe scalp loss (defined as ≥50% scalp hair loss according to the Severity of Alopecia Tool), the majority (56%) of patients explicitly discussed feelings of insecurity/inadequacy/self-consciousness, while 47% of patients discussed sadness/depression and 42% discussed anxiety/worry/fear/stress.[8] A qualitative survey assessing the experiences of patients living with AA similarly identified the hair loss to be “emotionally devastating for many” and not the few as the EAG suggests, with participants reporting that AA is “emotionally damaging”, “extremely distressing” and is a disease capable of bringing on “feelings of huge depression and suicidal thoughts”.[9] As such, while a small proportion of patients may eventually “come to terms” with their hair loss, this does not appear to be the dominant experience based on available evidence, nor does it negate the extended period of distress and reduced HRQoL which may precede this acceptance.[9] **Evidence supporting the anticipated HRQoL gain associated with hair regrowth** As described in Section B.1.3.2 of the Company Submission (CS), available evidence also suggests that patients with more severe disease and greater scalp involvement have a greater HRQoL impairment compared to those with milder disease (less extensive hair loss).[5, 10-12] For instance, in the study by Abedini *et al.* (2018), patients with severe AA had significantly greater HRQoL impairment compared with mild AA patients, as measured by the Dermatology Life Quality Index (DLQI) (10.7±7.5 and 5.4± .8, respectively).[11] Similarly, in the study by Edson-Heredia *et al.* (2022) which explores the severity and burden of AA in

		Japan, patients with severe AA reported higher scores on all domains of the Skindex-16 questionnaire (emotions, symptoms and functioning [p < 0.005]) and scored higher on the HADS-A and HADS-D, indicating greater levels of anxiety and depression amongst patients with more severe disease.[13] This may reflect the increased visibility of the scalp hair loss in severe AA, which patients often report as being the most bothersome aspect of AA.[14] As such, while the HRQoL outcomes associated with hair regrowth are poorly studied in patients with AA, it would be expected that an improvement in the Severity of Alopecia Tool (SALT) score would lead to concomitant improvements in HRQoL, given that hair regrowth (an improvement in SALT) is effectively a proxy for the difference in HRQoL reported between patients with severe versus mild AA. **Limitations of the EQ-5D instrument in AA** Contrary to the extensive evidence from the peer-reviewed literature highlighted above, the utility values derived from the pooled BRAVE-AA EQ-5D data appear to suggest that patients with severe AA are not experiencing HRQoL impairment, given that the mean baseline utility value is comparable to the age and sex-adjusted utility value for the general population and many patients report perfect health (resulting in a significant ceiling effect to any response to treatment).[15, 16] The BRAVE-AA EQ-5D data also appear to suggest that patients experience very limited, if any, improvement in HRQoL after experiencing a significant improvement in SALT score, as demonstrated by the lack of correlation between trial participants' SALT and EQ-5D-3L scores at both baseline and Week 36, highlighted in Question B9(b) of the clarification questions. As noted by a clinical expert for whom additional input was sought post-clarification questions (24th October 2022), AA therefore does not appear to affect HRQoL using these generic measures, but this this is not because AA does not affect HRQoL, rather it is because these questionnaires are not fit for purpose to assess HRQoL in patients with severe AA.[17] The Company therefore wishes to re-iterate that the BRAVE-AA EQ-5D data are unsuitable and lack any face validity for capturing the HRQoL detriment of severe AA and the subsequent HRQoL gain experienced by patients responding to baricitinib, and invites the Committee to consider whether it truly plausible, as

		posited by the EAG, that the significant hair regrowth observed during the BRAVE-AA trials, as shown in Figure 1, is not associated with significant improvements in HRQoL. **Figure 1. Photographs of changes in absolute SALT scores over time through Week 36 in patients treated with baricitinib in BRAVE-AA1**  **Abbreviations**: SALT: Severity of Alopecia Tool.

		Reliability of the Adelphi DSP-derived utilities for capturing HRQoL While the Company acknowledges that the Adelphi DSP data are also affected by the limitations of the EQ-5D instrument in severe AA highlighted above, this does not negate the fact that the Adelphi values are more suitable than the values generated in the BRAVE-AA trials due to the inflated ceiling effect associated with the BRAVE-AA EQ-5D data, which does not occur to the same extent in the Adelphi DSP data. As outlined in the response to clarification question B7, the proportion of patients reporting perfect health in the BRAVE-AA trials was ([redacted] [Figure 2]), which is likely to be higher than the proportion of patients with severe AA in UK clinical practice. The proportion reporting perfect health in the severe/very severe group in the Adelphi DSP ([redacted] [Figure 3]), which was used as a proxy of the baseline utility in the base case, appears to be more appropriate. As such, the Company acknowledge that though the Adelphi DSP data are still subject to the content validity issue, the proportion of patients reporting perfect health is likely to be more reflective of the UK patient population with severe AA and the Adelphi DSP is therefore likely to have appropriately estimated the HRQoL detriment associated with severe AA. However, it is clear that the BRAVE-AA trial data were subject to a far greater ceiling effect than the Adelphi DSP, whereby [redacted] of the participants in the BRAVE-AA trials were unable to report an improvement in HRQoL, even if they experienced significant hair regrowth. The difference between the EQ-5D values for the mild and severe/very severe groups from the Adelphi DSP (used as a proxy for the HRQoL change from baseline following achievement of SALT≤20) is therefore likely to be more representative of the utility gain associated with hair regrowth than the trial change from baseline data. Consequently, while there may be certain methodological benefits associated with the BRAVE-AA utilities, including that the efficacy and HRQoL estimates are generated from the same patients, these benefits are greatly outweighed by the lack of generalisability of the HRQoL data to the severe AA population in UK clinical practice and therefore the unsuitability of the BRAVE-AA data in capturing the HRQoL gains associated with hair regrowth following baricitinib treatment. As such, the CEM results derived using the BRAVE-AA trial data should be interpreted with caution. Additionally, the Company would like to emphasise that the Adelphi DSP was conducted independently by Adelphi Real World who were responsible for the

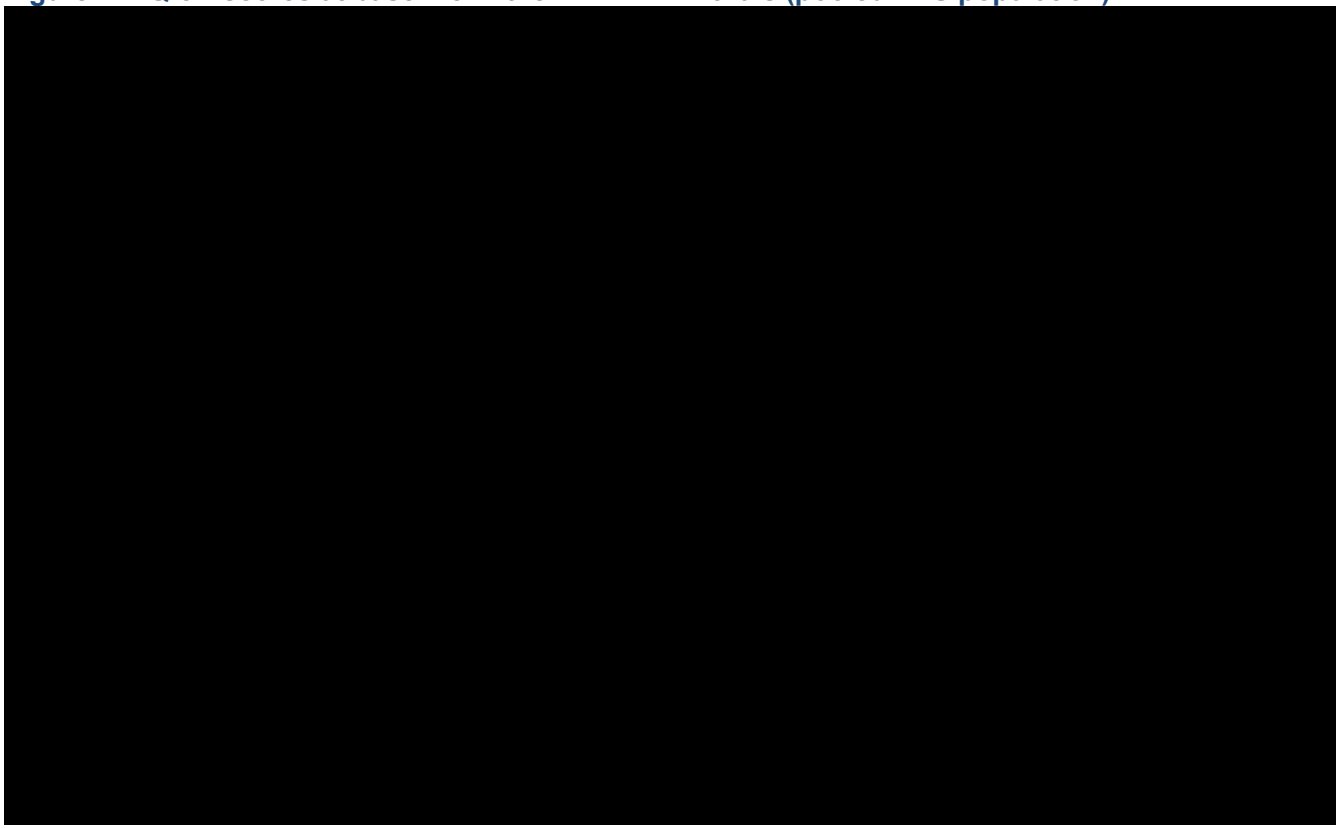
		data collection and data handling and thus, the data provides a robust and non-biased source of evidence to use in the cost-effectiveness analysis. **Figure 2. EQ-5D scores at baseline in the BRAVE-AA trials (pooled FAS population)**  **Abbreviations**: EQ-5D: European Quality of Life-5 Dimensions. **Footnotes:** BRAVE-AA trials (n=1195).

		Figure 3. EQ-5D scores in the Adelphi DSP severe and very severe subgroups  **Abbreviations**: EQ-5D: European Quality of Life-5 Dimensions; DSP: disease specific programme. **Footnotes**: Adelphi DSP severe and very severe subgroup (n=184). As mentioned, the Company sought additional clinical expert opinion post-clarification questions, which provided further support for the Company's response to clarification question B7. When presented with trial and Adelphi DSP utilities at baseline and asked what the most appropriate values for patients with severe AA were based on clinical experience, the clinical expert noted that clinicians would likely not want to put patients with a history of anxiety and depression forward for screening as they are likely to fail the screening process. The clinical expert expressed how this may subsequently skew the trial population away

from those that suffer from the most severe neuropsychiatric consequences of AA.[17] As such, the clinical expert concluded that the utility values derived from the Adelphi DSP are likely to be more relevant and also show a more plausible improvement from baseline associated with response.[17] The skew in the trial participants' baseline HRQoL scores can be demonstrated further by comparing these scores against the mean HADS-A and HADS-D scores for the severe and very severe subgroups of the Adelphi DSP. For both the HADS-A and HADS-D, patients in the Adelphi DSP scored [redacted] which is notably higher than the baseline BRAVE-AA trial participant scores presented in Table 3.

Table 3. Baseline HADS-A and HADS-D scores in the BRAVE-AA trials

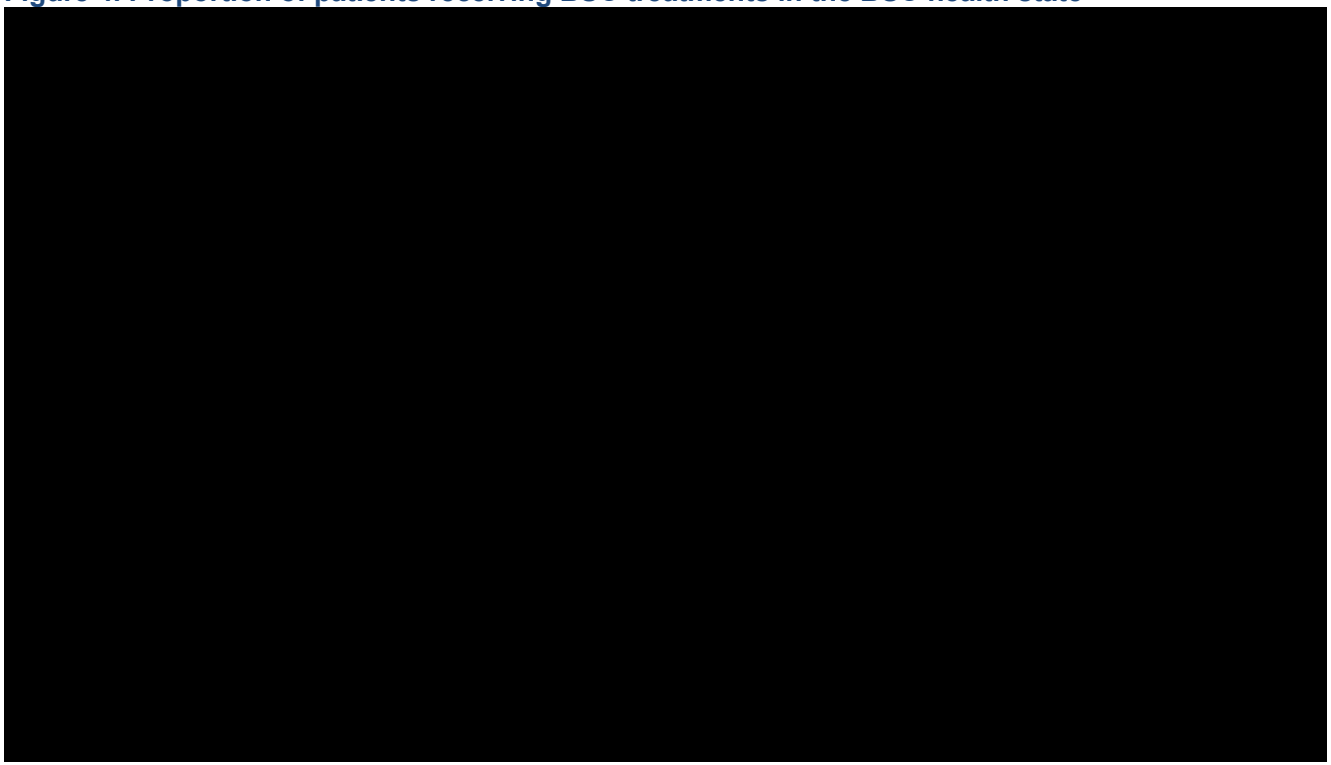
Week 36	BRAVE-AA1			BRAVE-AA2			Adelphi
	PBO (N = 189)	2 mg BARI (N = 184)	4 mg BARI (N = 281)	PBO (N=156)	2 mg BARI (N = 156)	4 mg BARI (N = 234)	Severe/Very Severe (HADS-A: N=166; HADS-D:N=168)
HADS Anxiety							
Mean (SD) baseline score	[redacted]						
HADS Depression							
Mean (SD) baseline score	[redacted]						

Abbreviations: BARI: baricitinib; HADS: Hospital Anxiety Depression Scale-Anxiety
Source: BRAVE-AA1 Clinical Study Report; BRAVE-AA2 Clinical Study Report.[15, 16]

		In the context of the unmet need in this indication, considering the evidence demonstrating the impacts of severe AA on HRQoL in a substantial proportion of patients, and the additional clinical support of the Company's position that the Adelphi DSP is a more appropriate source of utilities, the Company has not chosen to incorporate the EAG's proposed approach to modelling HRQoL in the base case, and invites the Committee to reject the EAG's stance on this issue. **The Company request that Figure 1 be included in the Committee slides to inform their discussions.**
Issue 4: Disease monitoring costs for best supportive care • What happens to patients whose condition does not respond to treatment? • Do patients continue to engage in further treatment? If not, are they discharged from care?	Yes	The EAG advises that if response to baricitinib is not achieved, patients will not engage with further treatment and will be discharged from care. Consequently, the EAG proposes that all disease management costs in the BSC health state should be excluded from both arms of the model. Based on the model structure outlined in Section B.3.2.2 of the Company Submission (CS), patients who do not respond to baricitinib or 'no active treatment' (previously 'watch and wait') transition into the BSC health state, comprised of a basket of treatments. Within this basket of treatments, a large proportion (updated post-clarification to [redacted]) of patients receive BSC treatments such as DPCP, methotrexate and intralesional steroids, and therefore incur disease management costs. The remaining patients ([redacted]) in the BSC health state receive no current treatment, and therefore do not incur any treatment costs. These BSC treatment proportions are informed by the Adelphi DSP, which the Company considers to be the best available evidence for informing the composition of the BSC health state, although the Company accepts that there is some uncertainty associated with these proportions, given that the treatment pathway for severe AA is not well-defined in the UK. Despite this uncertainty, the Company does not find the scenario proposed by the EAG to be plausible since it is unlikely that *all* patients who do not respond would receive no further treatment and be discharged from care. Therefore, the Company maintains that a proportion of non-responders can be expected to seek BSC treatments after not experiencing hair regrowth in the Induction phase, or experiencing treatment failure in the Maintenance phase, although the Company acknowledges that the exact proportion of patients expected to receive BSC treatment remains uncertain. Therefore, the Company have explored the effect of varying the proportion of patients receiving BSC

		treatments versus no treatment in the BSC health state on the cost-effectiveness outcomes for baricitinib, as detailed further below. In addition to this, the Company would like to re-iterate that given the substantial impact of severe AA beyond hair loss and the lack of effective, licensed treatment options currently in this indication, patients with severe AA still often choose to endure these treatments in hope that they induce hair regrowth.[1, 9] In this context, the Company has further considered that it is likely that the receipt of an effective, licensed treatment with a well-studied safety profile would reduce the willingness of patients to further engage with BSC therapies versus patients who have not received such a treatment, i.e. those receiving 'no active treatment'. Similarly, it would be expected that prescribers would be less willing to prescribe BSC treatments following treatment failure with a licensed treatment such as baricitinib. The AA Consensus of Experts, which reported a consensus among 50 international clinical experts that "If all treatments were equally reimbursed, JAK inhibitors would be the ideal choice of systemic therapy in adults", supports this point, given that this preference would likely translate into a reduced willingness of clinicians to prescribe BSC treatments should a JAK inhibitor (such as baricitinib) fail.[18] A number of currently available BSC treatments, such as DPCP and intralesional corticosteroids, are also more burdensome to administer and uncomfortable to receive compared with a once-daily oral tablet, which may further reduce the willingness of patients to tolerate BSC treatments following treatment with baricitinib. For these reasons, the Company anticipate that there would be a reduced incentive for both patients and clinicians to accept BSC treatments following the receipt of baricitinib. This can be modelled as having a lower proportion of patients in the baricitinib arm receiving BSC treatments following non-response compared to the 'no active treatment arm', as shown in Figure 4, which shows a reduction in BSC treatment use of 50% following baricitinib vs 'no active treatment'.

Figure 4. Proportion of patients receiving BSC treatments in the BSC health state



Abbreviations: BSC: best supportive care; DSP: disease specific programme

Based on the rationale described above, the Company considers a 50% reduction (halving) in the amount of BSC treatment accepted by patients and clinicians following the introduction of baricitinib to be the most likely scenario. To inform the revised Company base case post Technical Engagement, the Company leaves unchanged the proportion of patients receiving off-label treatments in the Adelphi DSP severe and very severe subgroups (■) as per the post-clarification base case, given that this is a robust source of

		evidence, with this proportion halved for patients entering the BSC state following baricitinib. Under these assumptions, [redacted] of patients in the 'no active treatment' arm would receive BSC treatment, while [redacted] in the baricitinib arm would receive BSC treatment following non-response. Considering this scenario, baricitinib is dominant. Nevertheless, the Company acknowledge that the extent of the reduction in BSC treatments received following baricitinib versus 'no active treatments' is uncertain, and as such, scenarios exploring a relative reduction of 100%, 75%, 50% and 25% in the baricitinib arm versus the 'no active treatment' arm are explored in Table 4 below. It should be noted that the proportions presented in the left-hand side column of Table 4 are in relation to the base case presented in the CS, in which [redacted] of patients in BSC receive BSC treatments, while [redacted] receive no treatment, following both baricitinib and comparator. Therefore, as an example, a 50% relative reduction in BSC treatment use following baricitinib vs the comparator would equate to [redacted] of all patients in the BSC health state incurring the cost of BSC treatments following treatment with baricitinib. Given the uncertainty in determining the absolute proportions receiving BSC treatment in each arm and the consistent pattern in the results across the scenarios in Table 4, the Company invites the Committee to consider whether there is less decision uncertainty in accepting simply that there is likely to be a difference between the treatment arms in the proportion of patients receiving BSC treatment, rather than attempting to determine the exact proportions of patients who would receive BSC treatment following non-response to each of baricitinib and 'no active treatment'. The Company note that the consideration of the difference in proportions between the baricitinib and 'no active treatment' arms rather than the exact proportions receiving BSC treatment, may provide a pragmatic resolution to this issue, and minimise the decision uncertainty as to whether baricitinib is a cost-effective use of NHS resources.

Table 4. Impact of BSC scenarios on the ICER

Proportion of patients receiving BSC treatments after non-response to 'no active treatment', as a percentage of the base case[a]	Relative reduction in BSC following baricitinib vs following 'no active treatment'			
	100% reduction *(baricitinib BSC proportion relative to 'no active treatment' base case value)*[b]	75% reduction *(baricitinib BSC proportion relative to 'no active treatment' base case value)*[b]	50% reduction *(baricitinib BSC proportion relative to 'no active treatment' base case value)*[b]	25% reduction *(baricitinib BSC proportion relative to 'no active treatment' base case value)*[b]
100%	Dominant *(0.0%)*	Dominant *(25.0%)*	Dominant *(50.0%)**	Dominant *(75.0%)*
90%	Dominant *(0.0%)*	Dominant *(22.5%)*	Dominant *(45.0%)*	Dominant *(67.5%)*
80%	Dominant *(0.0%)*	Dominant *(20.0%)*	Dominant *(40.0%)*	Dominant *(60.0%)*
70%	Dominant *(0.0%)*	Dominant *(17.5%)*	Dominant *(35.0%)*	Dominant *(52.5%)*
60%	Dominant *(0.0%)*	Dominant *(15.0%)*	Dominant *(30.0%)*	Dominant *(45.0%)*
50%	Dominant *(0.0%)*	Dominant *(12.5%)*	Dominant *(25.0%)*	Dominant *(37.5%)*
40%	Dominant *(0.0%)*	Dominant *(10.0%)*	Dominant *(20.0%)*	£8,166 *(30.0%)*
30%	Dominant *(0.0%)*	Dominant *(7.5%)*	Dominant *(15.0%)*	£21,389 *(22.5%)*
20%	Dominant *(0.0%)*	Dominant *(5.0%)*	£16,626 *(10.0%)*	£34,611 *(15.0%)*
10%	Dominant *(0.0%)*	£29,849 *(2.5%)*	£38,841 *(5.0%)*	£47,834 *(7.5%)*

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio.
Footnotes: [a]These proportions are calculated from the base case value of patients incurring treatments costs as part of BSC, as defined in the CS ([redacted]) [b] The values in parentheses relate directly to the values on the most left-hand side column of the table. * This result reflects the company's preferred base case following Technical Engagement, as presented in the model structure diagram. **Key**: Dominant: ; ICER < £30,000: ; ICER > £30,000

		The Company request that Table 4 be included in the Committee slides to inform their discussions.
Other issues identified by the NICE technical team (not included in the EAR): Issue 5: Where would baricitinib and its comparators typically be commissioned in NHS practice?	No	Baricitinib is anticipated to be commissioned in secondary care only; this assumption is based on the Company's understanding of the treatment pathway for patients with AA, in which patients with mild to moderate AA are treated in primary care and patients with severe AA are treated in secondary care, following referral to a specialist dermatologist. Given the population of relevance to this submission is adults with severe AA, secondary care is the most relevant setting for the dispensing of baricitinib. For BSC treatments, the Company has assumed that the majority would be commissioned in primary care, as it is anticipated that patients would be discharged from secondary care following non-response to baricitinib. The only exceptions to this assumption are DPCP and intralesional corticosteroids, which require specialist administration and would therefore be anticipated to delivered in secondary care.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 5 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	**Does this response contain new evidence, data or analyses?**	**Response**
Additional issue 1: Inclusion of adverse events and associated unit costs	EAG Report, Page 99	No	Given the very low incidence of serious adverse events (AEs) observed in the BRAVE-AA trials (less than 2.5% of patients had at least one serious AE), their impact was negligible in the results and therefore the model does not include any specific serious AE. However, treatment-emergent adverse events (TEAE) have been considered in the economic model for transparency. TEAEs are defined as an untoward medical occurrence that emerges during a defined treatment period, having been absent pre-treatment, or worsens relative to the pre-treatment state, which does not necessarily have a causal relationship with this treatment. Both upper respiratory tract infection (URTI) and nasopharyngitis were assigned the cost of one GP visit (£39 based on PSSRU costs). Acne was assigned a cost of £171.53 based on the cost of a dermatology outpatient visit (Weighted average of WF01A–D and WF02A–D – Dermatology). Headache was assigned a cost of £206.34 based on the cost of a neurology outpatient attendance from the NHS Reference Costs (Service Code 400).

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 6 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Issue 1: Definition of the comparator	'Watch and Wait'	'No Active Treatment'	+ £3,212
Issue 2: Definition of treatment response at Week 36	$SALT_{50}$	SALT≤20	- £1,001
Issue 4: Disease monitoring costs for best supportive care	All patients receive BSC therapies following non-response	[redacted] of patients receive BSC treatment following 'No Active Treatment' [redacted] of patients receive BSC treatment following baricitinib	Dominant
Long-term discontinuation in baricitinib arm	10.00%	8.84%	- £1,779
Removal of non-pharmacological psychological support costs	Include non-pharmacological psychological support costs	Exclude non-pharmacological psychological support costs	+ £607
Wig use in the induction phase	Two wigs assumed in induction phase of the model for both arms	One wig assumed in induction phase of the model for both arms	-£4

Company's base case following technical engagement (or revised base case)	Deterministic incremental QALYs: [redacted]	Deterministic incremental costs: [redacted]	Dominant

Abbreviations: BSC: best supportive care; QALYS: quality-adjusted life year; SALT: severity of alopecia tool.

Sensitivity analyses around revised base case

Table 7. Revised Company base case cost-effectiveness results (probabilistic)

	Total cost	Total QALYs	Incremental Cost	Incremental QALYs	ICER
No Active Treatment	[redacted]	[redacted]	-	-	-
Baricitinib	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Abbreviations: ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life year.

Probabilistic sensitivity analysis (PSA)

A probabilistic sensitivity analysis (PSA) was run with 1,000 Monte Carlo simulations in order to assess the uncertainty associated with model input parameters. Use of 1,000 iterations was deemed appropriate based on the results of an ICER convergence test, shown in Figure 5.

Figure 5. Convergence plot for NMB

Abbreviations: CI: confidence interval; NMB, net monetary benefit; PSA: probabilistic sensitivity analysis; QALY: quality-adjusted life year.

A visual representation of the PSA results comparing baricitinib 4mg and placebo is provided in the cost-effectiveness plane (see Figure 6 below). Each dot represents one Monte Carlo simulation where the input parameters are sampled from the distributions in a total of 1,000 loops. The results of the cost-effectiveness plane show moderate uncertainty with regards to the extent of the additional costs for baricitinib 4 mg compared with 'No Active

Treatment', but little uncertainly with regards to the existence of these additional costs, as most dots fall on the North quadrants of the plane. On the effectiveness side, PSA results show moderate uncertainty with regards to the extent of additional benefits.

Figure 6. Cost-effectiveness plane for baricitinib 4 mg compared with 'No Active Treatment'

Abbreviations: QALY, quality-adjusted life year.

The cost-effectiveness acceptability curve shows a [redacted]% probability of baricitinib 4 mg being cost-effective compared with 'No Active Treatment' at a cost-effectiveness threshold of £30,000/QALY (Figure 7).

Figure 7. Cost-effectiveness acceptability curve for baricitinib 4mg compared with placebo

Footnotes: "placebo" in the figure represents "No Active Treatment".

Deterministic sensitivity analysis (DSA)

A deterministic one-way sensitivity analysis (OWSA) has been performed and baricitinib remained dominant in all scenarios tested. The tornado diagram is not presented here, due to the challenges associated with interpreting negative ICERs.

Scenario Analyses

Various scenario analyses were explored (please see the list below). Baricitinib is dominant in all scenarios tested and therefore full results are not presented.

- **Scenario 1:** Starting population with SALT 50–94 (severe population)
- **Scenario 2**: Starting population with SALT 95–100 (very severe population)
- **Scenario 3:** Response based on SALT75
- **Scenario 4:** Response based on SALT50
- **Scenario 5:** Utilities based on EQ-5D data from the BRAVE-AA trials
- **Scenario 6:** Utilities based on HADS data from the BRAVE-AA trials

References

1. Messenger AG, McKillop J, Farrant P, et al. British Association of Dermatologists' guidelines for the management of alopecia areata 2012. British Journal of Dermatology 2012;166:916-926.
2. Lamb RC, Young D, Holmes S. Retrospective review of diphencyprone in the treatment of alopecia areata. Clin Exp Dermatol 2016;41:352-8.
3. Olsen EA, Roberts J, Sperling L, et al. Objective outcome measures: Collecting meaningful data on alopecia areata. J Am Acad Dermatol 2018;79:470-478.e3.
4. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. January 2022. Available at: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741. Last accessed: April 2022.
5. Rencz F, Gulácsi L, Péntek M, et al. Alopecia areata and health-related quality of life: a systematic review and meta-analysis. Br J Dermatol 2016;175:561-71.
6. Balieva F, Kupfer J, Lien L, et al. The burden of common skin diseases assessed with the EQ5D™: a European multicentre study in 13 countries. Br J Dermatol 2017;176:1170-1178.
7. Titeca G, Goudetsidis L, Francq B, et al. 'The psychosocial burden of alopecia areata and androgenetica': a cross-sectional multicentre study among dermatological out-patients in 13 European countries. Journal of the European Academy of Dermatology and Venereology 2020;34:406-411.
8. Aldhouse NVJ, Kitchen H, Knight S, et al. "'You lose your hair, what's the big deal?' I was so embarrassed, I was so self-conscious, I was so depressed:" a qualitative interview study to understand the psychosocial burden of alopecia areata. J Patient Rep Outcomes 2020;4:76.
9. Davey L, Clarke V, Jenkinson E. Living with alopecia areata: an online qualitative survey study. Br J Dermatol 2019;180:1377-1389.
10. Liu LY, King BA, Craiglow BG. Health-related quality of life (HRQoL) among patients with alopecia areata (AA): A systematic review. J Am Acad Dermatol 2016;75:806-812.e3.
11. Abedini R, Hallaji Z, Lajevardi V, et al. Quality of life in mild and severe alopecia areata patients. Int J Womens Dermatol 2018;4:91-94.
12. Dalgard FJ, Bewley A, Evers AW, et al. Stigmatisation and body image impairment in dermatological patients: protocol for an observational multicentre study in 16 European countries. BMJ Open 2018;8:e024877.

Technical engagement response form

13. Edson-Heredia E, Aranishi T, Isaka Y, et al. Patient and physician perspectives on alopecia areata: A real-world assessment of severity and burden in Japan. J Dermatol 2022;49:575-583.
14. FDA. The Voice of the Patient: Alopecia Areata. Public Meeting: September 11, 2017. Report Date: March 2018. Available at: https://www.fda.gov/files/about%20fda/published/Alopecia-Areata--The-Voice-of-the-Patient.pdf. Last accessed: February 2022.
15. Eli Lilly. Clinical Study Protocol: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA1). June 2021.
16. Eli Lilly. Clinical Study Protocol: A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA2). June 2021.
17. Eli Lilly. Clinical Expert Input for Technical Engagement: Meeting Minutes. Meeting date: 24th October 2022.
18. Meah N, Wall D, York K, et al. The Alopecia Areata Consensus of Experts (ACE) study: Results of an international expert opinion on treatments for alopecia areata. J Am Acad Dermatol 2020;83:123-130.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing your 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. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR (section 1). You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

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. 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.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

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 redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **Thursday 1 December 2022.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

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.

Part 1: Treating severe alopecia areata and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Abby Macbeth
2. Name of organisation	Norfolk & Norwich University Hospitals NHS Trust (main employer) and on behalf of the British Association of Dermatologists
3. Job title or position	Consultant Dermatologist & Service lead (NNUH)
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 severe alopecia areata? ☐ A specialist in the clinical evidence base for severe alopecia areata or technology? ☐ 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 did not submit one, I do not 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. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	
8. What is the main aim of treatment for severe alopecia areata?	

(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. 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)	
10. In your view, is there an unmet need for patients and healthcare professionals in severe alopecia areata?	
11. How is severe alopecia areata currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • 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.) • What impact would the technology have on the current pathway of care?	
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	

13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	
15. Will the technology be 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)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen	

may be more easily administered (such as an oral tablet or home treatment) than current standard of care	
18. 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? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
20. Do the clinical trials on the technology reflect current UK clinical practice? • 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?	
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	
22. How do data on real-world experience compare with the trial data?	
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any	

potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Issue 1: Definition of the comparator Which of the following is considered to be standard of care in the treatment of severe alopecia areata? • 'Watch and wait' comprising no active treatment and frequent monitoring? Treatment with diphenylcyclopropenone (DPCP)? • Treatment with a basket of low-effectiveness non-DPCP therapies including systemic corticosteroids and immunosuppressants?	It is difficult to define "standard of care" for severe alopecia areata, as topical steroids and intralesional steroids are usually reserved for less than 50% scalp involvement. There are no evidence-based treatments routinely available at all geographic sites in the NHS. In my clinical practice, I have access to Contact Immunotherapy (DCPC) and so this would be an effective comparator for Baricitinib but many dermatology departments do not have access to contact immunotherapy. I also would use systemic immunosuppressants, such as methotrexate or ciclosporin, at this point in a patient pathway and these drugs are also available widely across UK dermatology centres. I am unsure of the frequency of use for Alopecia areata outside of tertiary centres but this would be a more widely available comparator.

• No active treatment and discharge from care?	Systemic corticosteroids are reserved in my practice for rapidly progressing disease, or induction of disease whilst starting another immunosuppressant such as methotrexate, and as such should not be used as a comparator as a monotherapy. If DCPC and systemic immunosuppressants have failed or are contraindicated, 2 annual wig prescriptions with annual follow-up may be used or patients may wish to try other available immunosuppressants. On balance, I feel that a systemic immunosuppressant (e.g. Methotrexate) would be the best comparator for severe alopecia, for a patient who is requesting active treatment.
Issue 2: Definition of treatment response at Week 36 What is considered a clinically meaningful change in the Severity of Alopecia Tool (SALT)? • SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair) • $SALT_{50}$: at least a 50% improvement from baseline SALT score • $SALT_{75}$: at least a 75% improvement from baseline SALT score	I consider $SALT_{50}$ (at least a 50% improvement from baseline SALT score) to be a clinically meaningful response. Achieving an absolute SALT score of 20 or less is very difficult for a patient who may have started with 95-100% hair loss, however if they achieved an absolute SALT score=50, this may have a significant impact on the patient's quality of life. I would be in support of the use of a 50% reduction in surface area of hair loss for this reason (SALT50.) Surface area regrowth only forms a part of assessment of response, however. I explain this to patients in clinical practice and this forms the basis of a risk benefit discussion. I also use the opinion of the patient on their general satisfaction with hair growth (considering distribution, hair quality, hair colour) and an overall impression of improvement, which could be considered to be a patient global assessment of sorts. [redacted]

	[redacted] Outcomes should also consider a patient-reported measure, given that the predominant reason for treatment is impact on the patient. However, I do understand that clear and measurable criteria will be required.
Issue 3: Source of utilities in the model • For the majority of people with severe alopecia areata, would their health-related quality of life be similar to the general population? • Which utility value best represents health-related quality of life for people with severe alopecia areata during the induction phase of treatment? Value from Adelphi DSP study (EAR, Table 29) or value using	Alopecia areata has a significant impact on quality of life and is associated with both Depression and Anxiety both prevalent, at the point of diagnosis, and new onset within the year after diagnosis. (Macbeth *et al*. 2022 *Br J Dermatol*.) 2 large, systematic reviews have demonstrated poorer HRQoL than control populations (Lui *et al.* 2016 *J Am Acad Dermatol*., Frenz *et al* 2016. *Br J Dermatol*.) Lui *et al* concluded that those with AA had comparable HR-QoL to those with eczema and psoriasis. EQ-5D-5L is not a desirable measure for Alopecia areata, as this tool does not capture the impact of hair loss and some of the domains are not relevant to this condition. Multiple tools have been developed for specific use for Alopecia areata but few are validated. The Alopecia areata Symptom Impact Scale (AASIS) has been partially validated.

pooled data from BRAVE-AA1 and BRAVE-AA2 (EAR, Table 30)?	The use of HADs score may have more relevance, due to the impact of AA on mental health and the likelihood of those patients with AA scoring more highly in the anxiety/depression domain of the ED-5D. I am obviously aware that the HADs is not a quality of life tool. I understand that EQ-5D and SF-36 are preferred tools. I do not feel able to comment regarding the utility value in tables 29/30, as this lies outside of my expertise.
Issue 4: Disease monitoring costs for best supportive care • What happens to patients whose condition does not respond to treatment? • Do patients continue to engage in further treatment? If not, are they discharged from care?	Some patients want to try all available treatments and so remain under active follow-up every 3 months with periods of systemic treatments (e.g.- methotrexate, ciclosporin- with appropriate drug monitoring bloods) or DCPC contact immunotherapy. If hair loss falls below 50% (i.e. there is some hair growth), patients may attend every 8 weeks for intralesional steroid injections. Some patients reach a point of acceptance of their chronic hair loss, if there are no signs of regrowth, and decide not to continue with treatments and request to be discharged from secondary or tertiary clinics. Some patents opt to wear wigs and are eligible for wig prescriptions in our geographic area if they have 50% or more hair loss. 2 wigs are prescribable per year by a Hospital Appliances prescription- with the cost of each wig x 2 and the cost of the annual follow up needed in some geographic areas to remain eligible for wigs (our area included.) Cost of wigs varies significantly, and usual NHS wig prescriptions are for 2 acrylic wigs per year but if there is a history of contact allergy, human hair wigs may be required on NHS prescription at greater cost to the NHS. Wig entitlements vary across the UK. Patient costs should also be considered, as patients often self-fund dietary/ vitamin supplements, scalp tonics, shampoos, microblading of eyebrows, tattooing for scalp or beard, and cosmetic cover-up preparations such as colour matched sprays and hair

	fibres. Some patients are ineligible for NHS wigs due to geographic variation and self-fund wigs for their lifetime at significant cost. These patient costs are difficult to quantify but are likely to be significant. The additional patient loss of earnings must also be considered. Individuals with alopecia areata are more likely to be issued with time off work certificates compared with the general population in the UK (adjusted Hazard Ratio 1·56, 95% CI 1·43-1·71); and to be recorded as unemployed (aHR 1·82, 95% CI 1·33-2·49.) (Macbeth et al. 2022- *Br J Dermatol*.)
Other issues identified by the NICE technical team (not included in the EAR): **Issue 5: Where would baricitinib and its comparators typically be commissioned in NHS practice?**	Oral Jak inhibitors for Alopecia areata would be used at the same point as Contact immunotherapy (DCPC) or oral immunosuppressants. These are usually used for alopecia areata of 50% or greater involvement, with the current episode of hair loss lasting for at least 6 months duration. Also, patch-type alopecia at non-scalp sites of culturally- or psychologically- impactful sites, for example at the beard site for religious or cultural reasons, should be considered. Current long waiting lists for secondary care assessment and treatment may suggest that patients may wait much longer than 6 months with severe alopecia areata in the community, and so commissioning in primary care could also be considered.
Are there any important issues that have been missed in EAR?	Additional information: Some patients are currently self-funding oral Jak inhibitors privately for AA in the UK. This is further leading to health inequalities, as those unable to self-fund do not have the opportunity to receive treatment. At patient engagement events, I am asked frequently to deliver presentations on Jak inhibitors and the likelihood of availability in the UK, as patients are aware of available research and there is a growing voice of patients hoping for treatment.

	Patients are now asking to remain “on the list” rather than being discharged from my secondary and tertiary clinics, so that they can access Jak inhibitors more quickly should they be approved. I am sure you will hear much more of the patient voice from the patient representatives.

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Alopecia areata has a significant impact on quality of life and increases the likelihood of anxiety, depression and loss of employment.
There is a significant unmet, clinical need for a safe, effective and approved medication for people with moderate-to-severe AA and Baricitinib appears to demonstrate efficacy.
There is no agreed standard of care beyond topical steroids and intralesional steroids (for <50% scalp involvement), but contact immunotherapy (DCPC) or systemic immunosuppression (e.g. Methotrexate), which is available at all secondary care sites, would be an appropriate comparator.
SALT50 (50% scalp hair regrowth from baseline) is my preferred option to determine treatment response, however patient-rated outcomes must also be considered.
EQ-5D-5L does not capture the impact of alopecia areata on quality of life, as the domains are not all directly relevant to the condition, and indirect patient costs should also be included in any analysis to capture the full impact on those affected, if possible.

Thank you for your time.

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

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing your 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. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR (section 1). You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

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. 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.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

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 redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **Thursday 1 December 2022.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

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.

NICE National Institute for Health and Care Excellence

Part 1: Treating severe alopecia areata and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Matthew Harries
2. Name of organisation	Northern Care Alliance NHS Foundation Trust
3. Job title or position	Consultant Dermatologist
4. Are you (please tick all that apply)	x An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with severe alopecia areata? ☒ A specialist in the clinical evidence base for severe alopecia areata or technology? ☐ 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 did not submit one, I do not 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. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	
8. What is the main aim of treatment for severe alopecia areata? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	

9. 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)	
10. In your view, is there an unmet need for patients and healthcare professionals in severe alopecia areata?	
11. How is severe alopecia areata currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • 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.) • What impact would the technology have on the current pathway of care?	
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	
13. Do you expect the technology to provide clinically meaningful benefits compared with current care?	

• Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	
15. Will the technology be 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)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	

18. 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? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
20. Do the clinical trials on the technology reflect current UK clinical practice? • 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?	
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	
22. How do data on real-world experience compare with the trial data?	
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this	

treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Issue 1: Definition of the comparator Which of the following is considered to be standard of care in the treatment of severe alopecia areata? • 'Watch and wait' comprising no active treatment and frequent monitoring? Treatment with diphenylcyclopropenone (DPCP)? • Treatment with a basket of low-effectiveness non-DPCP therapies including systemic corticosteroids and immunosuppressants?	There is a difference between "Best supportive care (BSC)" and "Standard of care". I agree that many people with extensive AA will receive no active treatment and be discharged from follow-up. However, whether these patients were truly informed of the treatment options or realistically could receive BSC due to geographic and access reason is less clear. In my practice (specialist clinic with access to DPCP) I would rate topical immunotherapy as the best comparator for baricitinib and represents current BSC in the UK. The reason for this is that it does regrow hair (and is recommended in BAD guidelines), response rates (albeit based on lower quality observational studies (see Lee et al. JAMA Derm 2018)) show not dissimilar regrowth rates to baricitinib for extensive disease, and it is safe and generally well tolerated (I disagree with the statement that it causes "severe

• No active treatment and discharge from care?	adverse events" in the EAG summary – these are rare). However, I agree access is limited to many as it is only available in certain specialist centres. For our clinic our referral rates are high, suggesting an unmet need for effective treatment in severe AA. We also frequently use immunosuppression (particularly ciclosporin), which I think can work well at regrowing hair (and is the best (IMO) out of all the "non-DPCP" options). The issue is the lack of robust evidence and the inability to use longer-term due to side effects. Many patients a willing to try multiple options for their hair loss. For example, from our clinic we retrospectively reviewed 50 consecutive AT/AU patients (i.e. SALT 100), with over half of them having received 3 or more treatment options (MH unpublished). I appreciate this population is skewed, being more severely affected and/or a more motivated population to get to the clinic in the first place, so may not be representative of everyone with extensive AA in the UK, but I do hear from many who attend the clinic that they just were not aware of the options available. I suspect many people with extensive AA do fall into the "no active treatment and discharge" category (reasons discussed above). However, in this group there are potential on-going costs: 1) virtually all of these patients (particularly females) will receive long-term wig prescription, and these can be human hair wigs (if acrylic allergic or intolerant) with higher cost to the NHS; 2) co-morbidities may be identified on assessment or blood screening; and 3) epidemiology studies in primary care show higher rates of anxiety, depression and unemployment in AA (see MacBeth et al. BJD 2022) requiring support.
Issue 2: Definition of treatment response at Week 36	Absolute SALT <20 (see Wyrwich et al. BJD 2020) and SALT 75 are most appropriate response outcome and will likely represent a meaningful improvement for patients. In

What is considered a clinically meaningful change in the Severity of Alopecia Tool (SALT)? • SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair) • $SALT_{50}$: at least a 50% improvement from baseline SALT score • $SALT_{75}$: at least a 75% improvement from baseline SALT score	fact, we have used SALT 75 as a critical criteria (based on patient feedback) in assessing the evidence for updating the BAD AA guidelines.
Issue 3: Source of utilities in the model • For the majority of people with severe alopecia areata, would their health-related quality of life be similar to the general population? • Which utility value best represents health-related quality of life for people with severe alopecia areata during the induction phase of treatment? Value from Adelphi DSP study (EAR, Table 29) or value using pooled data from BRAVE-AA1 and BRAVE-AA2 (EAR, Table 30)?	From experience, and the recent epidemiology literature, it feels counter-intuitive that people with AA have the same HRQOL to the general population. As mentioned, certain co-morbidities are more common, and anxiety and depression levels higher than a matched population in primary care. In my clinic over 1/3 have significant depression scores on PHQ-9 screening when they first attend. However, we do not have data to support the HRQOL data in a wider population with severe AA. I don't think I can comment on the utility value used. My impression is that the Adelphi DSP results would align more with my experience in managing AA patients (i.e. more impacted than the general population). One caveat - I have not seen the Adelphi DSP data – is it available in the supplied documentation? If so, it is not easy to find. I do wonder whether the BRAVE trial investigators may have excluded those more severely affected emotionally. When you look at other phase 2/3 JAKi trials run around this time, severe depression / psychological disease are down as exclusion criteria, which may have influence the investigators choice to put someone forward. Although I

	appreciate this exclusion was not in the BRAVE exclusion criteria (except identifying those actively suicidal) and this comment is just conjecture……
Issue 4: Disease monitoring costs for best supportive care • What happens to patients whose condition does not respond to treatment? • Do patients continue to engage in further treatment? If not, are they discharged from care?	As mentioned above, my patients are willing to try multiple therapies to regrow their hair. Once the decision is made to stop active treatment, most will still be supported by long-term wig prescription. Ongoing clinic visits and blood tests are usually not necessary, and they will be discharged. The wig service will continue (usually via the local orthotics department) and wig cost come from the dermatology department budget.
Other issues identified by the NICE technical team (not included in the EAR): **Issue 5: Where would baricitinib and its comparators typically be commissioned in NHS practice?**	Secondary care dermatology. Dermatologists are experienced in managing JAKi as they are increasingly used in atopic dermatitis management.
Are there any important issues that have been missed in EAR?	Typo in background section (p23) – it says SALT = 0 represents complete hair loss and SALT=100 is a full head of hair. This is incorrect – SALT = 100 is complete hair loss. Need for long-term pharmacovigilance – it would be good for NICE to encourage new baricitinib patients to take part in a national AA disease safety register. This is particularly important as new drugs come into the market and as these drugs start being used in children.

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

There is high unmet need for treatment in the UK.
Best supportive care does not necessarily equal standard care (due to access issues and advice), and may not be available to many in the UK
Patients may undergo multiple unsuccessful treatments before either regrowing hair or stopping active therapy due to inefficacy, resulting in costs to the NHS and potential side effects to the patient.
Most not on active treatment will still receive long-term wig prescriptions (including human hair wigs for those intolerant of acrylic) (via dermatology department budgets) and wider support for co-morbidities (via their GP).
Need for long-term prospective safety data as these drugs are introduced into UK dermatology practice.

Thank you for your time.

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

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholde r responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with severe alopecia areata or caring for a patient with severe alopecia areata. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR (section 1).

A patient perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

Patient expert statement

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

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. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **Thursday 1 December 2022.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

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.

Part 1: Living with this condition or caring for a patient with severe alopecia areata

Table 1 About you, severe alopecia areata, current treatments and equality

1. Your name	Lynn Wilks
2. Are you (please tick all that apply)	☒ A patient with severe alopecia areata? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with alopecia areata? ☒ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	Alopecia UK
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☒ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☒ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☒ I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	☒ I am drawing from personal experience ☒ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: Contacted patients taking JAKs on behalf of Alopecia UK ☐ I have completed part 2 of the statement **after attending** the expert

	engagement teleconference ☒ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with severe alopecia areata? **If you are a carer (for someone with severe alopecia areata) please share your experience of caring for them**	All my head hair was lost in Spring 2020, and by Autumn 2020 I was AU with no scalp, face or body hair. I had also lost my eyelashes and eyebrows. I had suffered all over head hair thinning 22 years before, probably caused by underactive thyroid & that recovered within 12 months In October 2018 I had suffered a brain haemorrhage, which needed surgery and six weeks in hospital – I was immobile, had cognitive difficulties, poor sight and in great pain. But with physiotherapy, an eye operation, exercise and taking tablets I improved over time in all aspects. With my hair loss – I felt 'why me, why now', I lost all confidence, did not want to go outside or socialise, I was depressed. My husband felt useless as he could not say or do anything to make me 'feel better'. Loss of all body hair added to the trauma, I felt very cold. And losing eyelashes and eyebrows and nasal hair was devastating, as even with a wig I had visible differences which people did stare at and ask about It is a journey – for me, I am thankful I found Alopecia UK for peer-to-peer support and to help me manage the grief of losing my hair and finding acceptance of wig wearing
7a. What do you think of the current treatments and care available for severe alopecia areata on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	7a. I am sad, angry, disappointed that healthcare professionals don't seem to 'care' about alopecia – seeing it as 'just cosmetic'. You have to fight to be referred and even to get a full range of blood tests from your GP. I wish there was a cure for alopecia and failing that a safe & effective treatment readily accessible. I am aware of the treatments on the treatment pathway – but only steroid cream was offered from the NHS and some private dermatologists. I finally saw a dermatologist with an interest in alopecia and was offered a JAK privately, I could not afford £10K per year. It is a post code lottery – I see and hear that on the Alopecia UK social media groups and a friend with Alopecia in Windsor was offered dithranol, then

	cyclosporin, then Methotrexate – all from the NHS; though nothing has worked for her. 7b: I see and hear on the Alopecia UK public & private social media groups and also did some focus groups for developing a wigs charter - many people (of our 10,000+ community) feel the same as me – little available, little offered and a continuous battle to receive care and treatment for alopecia. Then few treatments work so no hope and continued psychosocial impacts to life.
8. If there are disadvantages for patients of current NHS treatments for severe alopecia areata (for example, how they are given or taken, side effects of treatment, and any others) please describe these	For me as AU – I understand from the limited studies, BAD review and Cochrane review that even by the limited clinical studies, many of the treatments work in very limited numbers of patients with % hair growth often being limited. I am thinking Cyclosporin and methotrexate. The biggest issue is limited dermatologists who will even offer these treatments.
9a. If there are advantages of baricitinib over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does baricitinib help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	9a: The advantage of baricitinib is that in the phase lll trials the results look positive for numbers of patients who see hair regrowth and the % hair regrowth. Also, people seem to see regrowth of eyelashes and eyebrows. So, the main benefit is it works and hair regrowth is considerable to total. I see/hear positive stories on social media channels of people taking JAK inhibitors. My goal, as a wig wearer, would be not to need a wig in order to go out and socialise. I always wear a wig as I do not want the staring and feeling of being different. It is uncomfortable wearing a wig and there is always the fear of it being knocked/blowing off. 9b: Having real hair again so I could act as 'normal'. 9c: Trial and word of mouth results suggest that baricitinib works in high % numbers of patients and high % hair regrowth – including eyebrows and eyelashes. Also works well in severe alopecia so for someone like being AU for 3 years.
10. If there are disadvantages of baricitinib over current treatments on the NHS please describe these. For example, are there any risks with baricitinib? If you are concerned about any potential side effects you have heard about, please describe them and explain why	My understanding is that there are some possible side effects with baricitinib and need for regular blood monitoring, but that is the same with cyclosporin and methotrexate. My understanding is that there are less side effect risks with baricitinib than cyclosporin and methotrexate. Also, Baricitinib will be licensed for alopecia.

11. Are there any groups of patients who might benefit more from baricitinib or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	I understand that there can be a higher % of alopecia in some Asian and African heritage people. And we hear the stigma that these people can suffer so they could benefit more. It really is psychosocial impact of hair loss that is devastating. So, I would ask that baricitinib be available on the NHS for those who are really suffering from psychosocial impacts and hence have decreased quality of life. I hope it is prescribed for patients who are suffering severe psychosocial impact e.g. not going to work, not going out, severe anxiety.
12. Are there any potential equality issues that should be taken into account when considering severe alopecia areata and baricitinib? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	As answered above in question 11. Please consider ethnic populations, where the stigma of hair loss may be greater. And please consider men – male pattern hair loss may be common and hence baldness in men normalised. But I see and hear how much some men with severe alopecia suffer.
13. Are there any other issues that you would like the committee to consider?	Please take seriously psychosocial impact of severe alopecia. Alopecia may not progress physiologically to death, as cancer or other serious diseases. It may not affect an EQ5D score or QOL measures in terms of mobility, cognition dexterity and self care. But take it from someone who, following a brain haemorrhage and craniotomy, had difficulty with staying awake, eating, talking, memory, sight and mobility. Lost my career and driving licence. The impact to my mental health and

	quality of life was much greater and more severe when I suffered total head and body hair loss – and at the moment, I am not being offered any treatments or have any hope to get my hair back. Please consider the NHS goal of 'free at the point of treatment'. We see and hear of people taking out loans or even re-mortgaging homes in order to access and pay for JAK inhibitors privately. We had one situation recently where a person did this to access a JAK for her 20 year old daughter who could not go out and continue her life without hair. Please consider psychosocial impacts and the benefits that baricitinib could provide.

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement

Issue 1: Definition of the comparator Which of the following is considered to be standard of care in the treatment of severe alopecia areata? • 'Watch and wait' comprising no active treatment and frequent monitoring? Treatment with diphenylcyclopropenone (DPCP)? • Treatment with a basket of low-effectiveness non-DPCP therapies including systemic corticosteroids and immunosuppressants? • No active treatment and discharge from care?	We consider patient perspectives may particularly help to address this issue. I really do not like or accept your term 'no active treatment and discharge from care' While this is common once a person with alopecia has been referred to and seen a dermatologist (even just once), I feel this is neglect. It is mentioned in the EAG file that 'patients choose this option' I strongly disagree with this comment (P25). I was sent away after one secondary care dermatology appointment with the 'go and get yourself a nice wig dear' This, to me was not good patient care. I felt abandoned. To me 'watch and wait' with even annual monitoring and blood tests is valuable, and annual dermatology appointments are required in England to enable annual wig

Patient expert statement

	prescriptions. Many people with alopecia have atopy and comorbidities so need monitoring of these conditions. Please also consider the psychosocial impacts and support needed. While, as the EAG comments, there is little true psychological support available (through psychology or counselling), an annual appointment with a dermatologist can provide that support, hope and help to acceptance. That time to talk is often all that is needed. In my first year I had 7 GP appointments, 1 NHS dermatology appointment (after 9 months wait) and two private dermatology appointments (at 6 months and 15months) and one trichology appointment (at 8 months) – from the social media sites, I suggest that is the real situation and cost to the NHS in healthcare professional time.
Issue 2: Definition of treatment response at Week 36 What is considered to be clinically meaningful in the Severity of Alopecia Tool (SALT)? • SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair) • $SALT_{50}$: at least a 50% improvement from baseline SALT score • $SALT_{75}$: at least a 75% improvement from baseline SALT score	We consider patient perspectives may particularly help to address this issue. I am AU and from my perspective would need sufficient hair regrowth to enable me not to need a wig, so for me, less/equal to 20% scalp hair loss (80% scalp coverage) would be the only outcome worthwhile. AU changing to 50% scalp cover when that would most likely be patchy coverage would not be a good outcome for me. 75% improvement if it was 75% head/scalp coverage would also be good for me – if I felt I did not have to wear a wig.
Issue 3: Source of utilities in the model • For the majority of people with severe alopecia areata, would their health-related quality of life be similar to the general population?	We consider patient perspectives may particularly help to address this issue. We need an appropriate HRQoL model for alopecia. In terms of mobility, cognition, dexterity etc – then yes, I am sure the majority of people would have QOL similar to general population.

	But PLEASE consider psychosocial impact – depression, anxiety, loss of confidence & self-worth, agoraphobia.
Issue 4: Disease monitoring costs for best supportive care • What happens to people whose condition does not respond to treatment? • Do most people continue to engage in further treatment? If not, do they stop treatment?	We consider patient perspectives may particularly help to address this issue. Your EAG report suggests it is the person with alopecia who is 'choosing' what to do, to take treatment and whether or not to engage. The reality is that the person has to nag/fight/ persuade for a referral to secondary care and to be prescribed treatments! I was not offered any treatments from the NHS So yes, as your report suggests people with severe alopecia will re-engage with their GPs and dermatologists in order to receive baricitinib, which they consider a safe and effective treatment which will result in hair regrowth. There is excitement in the Alopecia UK social media groups and the two JAK inhibitor social media groups. When a treatment does not work, after approx. 6 months, the person is sad and disappointed. I know from friends with alopecia some just give up trying and others want to try anything else (which means ongoing nagging for repeat referrals to secondary care and the endless wait for a dermatology appointment).
Other issues identified by the NICE technical team (not included in the EAR): **Issue 5: Where do people with severe alopecia areata receive treatment? From their GPs or consultant dermatologist?**	We consider patient perspectives may particularly help to address this issue. GP first point of call and we see/hear the challenges that many people see, of receiving little support, blood tests, referral to dermatology. We know from BAD data that approx. 1:3 or 1: 4 people referred. In England people who want a wig prescription often need annual dermatology appointments to receive a prescription and this is a chance for monitoring the hair loss and mental health, as well as the patient 'hoping' to try a new treatment which will work. A dermatologist is often where patients go to for psychosocial impact support.
Are there any important issues that have been missed in EAR?	No – I think the EAR is comprehensive and strong in the three areas questioned: comparator, treatment response and utilities.

Patient expert statement

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Alopecia Is NOT just cosmetic, it is an autoimmune condition, with many people having other autoimmune conditions
- Please consider, it is not just % hair loss that matters but the psychosocial impact to the person with alopecia, living with a non-curable and unpredictable visible difference
- This treatment gives hope – the promise of an effective treatment, licenced for alopecia – this can bring my hair back and let me feel ‘normal’ again – no longer having a visible difference that results in stigma and prejudice and affects my mental health
- Quality of life is much more than dexterity, mobility, cognition and self-care (take it from a severe stroke/SAH survivor!) Please consider the psychosocial impacts of severe alopecia and the hope Baricitinib offers
- I don’t opt out! I want to be offered NHS care, treatments and support for alopecia

Thank you for your time.

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 NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder r responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with severe alopecia areata or caring for a patient with severe alopecia areata. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR (section 1).

A patient perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

Patient expert statement

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

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. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

NICE National Institute for Health and Care Excellence

The deadline for your response is **5pm** on **Thursday 1 December 2022.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

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.

Part 1: Living with this condition or caring for a patient with severe alopecia areata

Table 1 About you, severe alopecia areata, current treatments and equality

1. Your name	Sue Schilling
2. Are you (please tick all that apply)	☒ A patient with severe alopecia areata? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with alopecia areata? ☒ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	Alopecia UK
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☒ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☒ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☒ I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	☒ I am drawing from personal experience ☒ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: ☒ I have completed part 2 of the statement **after attending** the expert engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the

	expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with severe alopecia areata? **If you are a carer (for someone with severe alopecia areata) please share your experience of caring for them**	I was diagnosed with Alopecia Areata in primary school, with repeated spells of patchy loss through school and into adulthood. In my 40s it developed to Totalis and to Universalis over a few weeks, and I have been bald for the last 6 years. Temperature regulation is problematic without body hair; I am often cold and struggle to cool down in hot weather. My eyes are less protected than others without lashes and I have been to A&E in the past because of debris getting in my eyes. The weeks of shedding were awful. I was terrified about what was happening to my health generally. I was ashamed that somehow, I have not been able to predict what was going to happen and stem the hair loss. Remember, my experience for nearly 40 years had been small manageable patches that always grew back, this was on a much bigger scale. I was traumatised by the woman I saw looking back at me in the mirror, and once my lashes and eyebrows went then I could not recognise myself. I experienced a complete loss of identity. I have been verbally abused for being visibly different on the street. As a result of the trauma and fear of being ridiculed for looking “weird” I stopped going out socially, I stopped going to my place of work, I reduced my family network to my immediate family. I cried a lot. I lost my self-esteem. I lost my sense of femininity. I was constantly anxious. In the 18 months after hair loss, I spent something in the region of £2,500 on camouflage products like root sprays, wigs, brow microblading, fake lashes. I was in part-time self-employed work at this time, and this was a huge financial commitment, but I was focussed on anything to make me look more socially normal. Wigs are very uncomfortable for me, and I soon stopped wearing them. I learned to look for chemo headwear on websites because regular hats are too itchy. I got used to people who saw me without a wig asking how my cancer treatment was

	going. This was particularly gruelling, every time I had to say, "it's not cancer, it's alopecia", I felt like I was giving away a bit of my personal life to a stranger. It was intrusive. Alopecia occurred alongside a flare-up and diagnosis of another autoimmune condition, which was treated with oral medication. I knew that alopecia was different and that there was no known treatment, and so I felt powerless. It was clear that something was going more wrong than usual, but I didn't know what. When I am singled out as looking different, it can still trigger the anxiety I felt at the height of my alopecia-related trauma. And whilst, after a lot of psychological support, I have grown to like myself again, there is no doubt that alopecia has impacted my mental health, confidence, relationships, and sense of my own identity.
7a. What do you think of the current treatments and care available for severe alopecia areata on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	In my role as CEO at Alopecia UK I hear a lot from the patients in our community. Common themes that people are palmed off with a message that alopecia areata is only a cosmetic issue. This makes me furious; it is an autoimmune condition and should be treated with as much care as any other. People report that it is hard to get in front of a dermatologist and finding one who is a hair specialist is tough, it is a postcode lottery! Because many dermatologists are not hair experts, I hear stories of people who are fighting for treatments that they are seeing used in their peer network. The off-label nature of the currently available treatments seems to suggest that some dermatologists are willing to treat alopecia with a range of technologies and others are not. Access to treatments that require a lot of management like intralesional steroids and DPCP are the hardest to come by. It is inequitable and frustrating to have to fight for every bit of care, this requires energy, often time off work, and the mental capacity to stay committed while you are in psychological distress. It is exhausting. And, to layer this on top of the fatigue that comes for many from an autoimmune flare-up is simply cruel, people with alopecia deserve better.

8. If there are disadvantages for patients of current NHS treatments for severe alopecia areata (for example, how they are given or taken, side effects of treatment, and any others) please describe these	A person can progress from mild to severe alopecia in a short period of time. As I understand it, many of the current treatments are about arresting the shedding phase or stimulating the growth phase, so if your loss has developed too far your treatment options are reduced. If like me, you have been AU for years then I understand that the chance of a treatment working is limited. To reiterate my point above, perhaps the biggest challenge is finding a dermatologist who has the knowledge and confidence to prescribe off-label treatment at the time you need it.
9a. If there are advantages of baricitinib over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does baricitinib help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	The science looks positive and, alongside a high percentage of hair regrowth, people seem to be getting their lashes and brows back. As I understand it this is not likely to happen in other non-oral treatments. As above, for me and many others, the loss of facial features is reported as more distressing than the loss of scalp hair, so this is an important advantage. On social media and blogs, people who are taking JAKS already are reporting fabulous progress in hair recovery and confidence. This requires fewer clinic visits than contact immunotherapies, which, from our patient groups, I know can be time-consuming, incur travel costs and time away from work can be career (life-outcomes) limiting. The percentage of hair regrowth is the most important factor, if I could get back to a normal amount of hair then I would. I am not alone, for many of us it would mean a life feeling “normal” and not isolated. Baricitinib is licensed, which will make it immediately more accessible to a wider community as more dermatologists will be confident to prescribe it. I understand that the side effects of baricitinib are not as severe as with other immunosuppressants.
10. If there are disadvantages of baricitinib over current treatments on the NHS please describe these. For example, are there any risks with baricitinib? If you are concerned about any potential side effects you have heard about, please describe them and explain why	I understand that people with baricitinib will need to have regular monitoring to check for possible side effects. But I believe this is also true for other immunosuppressants.

11. Are there any groups of patients who might benefit more from baricitinib or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	In some minoritised ethnicities there is a social and religious stigma associated with hair loss. One of our Sikh ambassadors explains that "For a Sikh, you have 5 K's, and 'Kesh' (the practice of allowing one's hair to grow naturally) is one of them, so losing my hair almost felt like losing some of my Sikh identity." I understand that there is a higher prevalence of alopecia in South Asian communities, in urban areas and in those who are socially deprived, so these groups will benefit. I have heard Prof. Andrew Messenger describe this as a young person's disease, And I understand that the peak age of onset is in the late 20's. So, these people, who are at the stage of developing their career paths will benefit; they will not have their early lives disrupted by extreme psychosocial distress! There is a practical inequality between men and women insofar as it is hard for men to get access to wigs and appropriate camouflaging products, with this in mind I suspect baricitinib will be particularly well received by men. I have recently heard of a clinician whose practice monitors the mental health of men with alopecia particularly closely because he considers them to be at an especially high risk. The mental health ramifications of alopecia can be debilitating and for many life limiting; people who are experiencing mental ill health will benefit. In terms of those who will benefit less, I am concerned about how severe alopecia is measured. I would like to see this being a combination of percentage of hair loss + extent of distress. And because we know that facial features are important, I would like to see that is a patient has loss of facial hair then this is considered more severe.
12. Are there any potential equality issues that should be taken into account when considering severe alopecia areata and baricitinib? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil	As well as the issues covered in question 11, I would also like to observe something that I have picked up from my conversations with the leaders of other charities. I am told that there is specialised psychological care provision for many other conditions where a patient lives with a visible difference. I believe this to be true for many congenital issues that are discovered at birth. If this is true, then all patients with alopecia are disadvantaged. Hair growth, through the use of baricitinib will lessen the impact.

partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	
13. Are there any other issues that you would like the committee to consider?	I am concerned that the quality-of-life comparators do not allow patients like us to truly share the impact of living with this disease. The psychosocial impacts are well documented and well understood and whilst alopecia doesn't directly kill you, I know of people who have taken their own life because of how their alopecia triggered a terrible decline in their mental health. I believe that treatments like this one provide hope of living a normal life, not just for patients, but for their families and children. I also hope that this is the first stage of innovation and will drive more research and simplify treatments and support more people with other types of alopecia. JAK inhibitors, including baricitinib, are already available to buy, and people are already getting into debt to do so. At Alopecia UK we recently received a call from someone re-mortgaging their home to do so, such is their distress at living with alopecia. We have also heard of others paying for JAKs treatments with credit cards/loans. Please seriously consider how the NHS goal of 'free at the point of treatment' will improve the psychosocial wellbeing of patients with alopecia.

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement

Issue 1: Definition of the comparator Which of the following is considered to be standard of care in the treatment of severe alopecia areata? • 'Watch and wait' comprising no active treatment and frequent monitoring? Treatment with diphenylcyclopropenone (DPCP)? • Treatment with a basket of low-effectiveness non-DPCP therapies including systemic corticosteroids and immunosuppressants? • No active treatment and discharge from care?	As I mention above DPCP is not freely available to most patients, and I don't think it can be considered a fair comparator. I also understand that there is a risk of developing vitiligo following DPCP, and this is of particular concern for non-white patients. I understand why "watch and wait" is the right option for those who have a small patch that is not visible. I understand that 30% to 50% of patients spontaneously recover within 1 year of diagnosis, and that less than 25% progress to more severe AA. I don't like "watch and wait" as a treatment for someone with continued episodes. In my 30's I would use my hair loss as a trigger to go to the GP to get my bloods done as I knew it was likely that my hypothyroidism was also flaring. My alopecia flares were an accurate predictor of other challenges in my body. And they have proved to be into my 40's as I was diagnosed with SCLE around the same time as developing Universalis. I would like to see more tests to assess the immune system and swifter progression to secondary care for those with continued episodes, or growing patches. I am not alone, as I understand it a high percentage of people with alopecia areata have other autoimmune conditions.

Patient expert statement

	I am concerned about the tone in the EAG surrounding “no active treatment and discharge from care”. At Alopecia UK we run a private peer-support Facebook group with over 12,000 members. The patients there describe feelings of abandonment after being sent away. They describe a lack of treatment options and a lack of empathy. They are being forced to discharge rather than choosing to. There are some who choose to discharge, this may be because they have had access to an excellent dermatologist who has educated them, supported their distress and they have experienced a number of treatments. I suspect many people registered as having no active treatment have not been given a choice. I also hear that many of the patients in our community must go to their clinic to collect their wig prescription, are they really signed off active care if they are still receiving a prescription for an orthotic?
Issue 2: Definition of treatment response at Week 36 What is considered to be clinically meaningful in the Severity of Alopecia Tool (SALT)? • SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair) • $SALT_{50}$: at least a 50% improvement from baseline SALT score • $SALT_{75}$: at least a 75% improvement from baseline SALT score	I think it depends on the extent and location of your hair loss at the start of treatment. The aim for me and for many others is to look socially normal. When I was in the shedding phase 75% improvement would have made me have hair that looked close to my natural self and have enough to style my hair to camouflage any rogue patches on the sides or back of the head. If patches remain on the front or top of the head in a location that is hard to cover with a style, then this is unlikely to feel like success for some patients. To that end, someone with a very visible patch might be happy with a 50% improvement if it meant they could manage the remainder with cosmetic root coverage. Now with AU, I think I would need 80% recovery to look socially normal, but it would still very much depend on where any hair loss remained. For instance, if 20% loss is not sustained across the scalp but instead is one large area, it will be difficult to cover. I want to reiterate the importance of the regrowth of lashes and brows, and patients who regrow these may view success differently. While I would like a full head of hair, having eyelashes and eyebrows would make a huge difference to my self-esteem, even now!
Issue 3: Source of utilities in the model • For the majority of people with severe alopecia areata, would their	In my view the quality-of-life comparators do not reflect the impact of living with this disease. We need a better measure, the absence of one should not stop the progress of this important treatment. The psychosocial impacts are well documented and well understood and whilst

health-related quality of life be similar to the general population?	alopecia doesn't directly kill you, I know of people who have taken their own life because of how their alopecia triggered a terrible decline in their mental health. At Alopecia UK every day we hear about people not going to school or work, or so anxious that they can't leave the house, and more. And for those people who can go to work there is a stigma, there is research that shows people are less likely to employ someone with alopecia - varied by alopecia severity and whether alopecia is believed to be a medical condition. The life outcomes for lonely, isolated, disadvantaged people are not similar to the general population.
Issue 4: Disease monitoring costs for best supportive care • What happens to people whose condition does not respond to treatment? • Do most people continue to engage in further treatment? If not, do they stop treatment?	People want to be offered the best treatment at the time it is needed, sadly people with alopecia often have to battle busy GPs to get to secondary care. I think few people "stop" treatment through choice. When I went got to secondary care, I was already AU and was offered a wig prescription, I do not recall being offered medical treatment for my alopecia. I would like to see more people referred to NHS mental health services if clinic cannot support. Many people in our Facebook community are excited about the change that baricitinib can bring to their lives. They consider baricitinib to be safe and have seen photographic diaries of good regrowth in those taking it. I expect people with severe and even less severe alopecia will reengage with their dermatologists to get it.
Other issues identified by the NICE technical team (not included in the EAR): **Issue 5: Where do people with severe alopecia areata receive treatment? From their GPs or consultant dermatologist?**	First step is to go to the GP, where I understand that some GPs prescribe first line treatments, like mild topical steroids. Where more treatment is required, we hear many people fight for referrals to dermatology. Approx. 1 in 4 people are referred on. The dermatologist diagnoses, provides treatments, and if available wig provision. I suspect the dermatologist is the primary source of NHS mental health care for many patients. Where patients are not getting good care on the NHS they will fund privately, or some turn to trichologists, which as an unregulated sector can has some extremely qualified practitioners, but also some charlatans. Some vulnerable people are hoodwinked by unscrupulous salespeople into buying treatments with no scientific underpinning.
Are there any important issues that have been missed in EAR?	I don't think so.

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Alopecia areata (AA) is an inflammatory disease, it is not just a cosmetic issue, for many patients there are other associated autoimmune co-morbidities.
- There is currently on-label treatment available in the NHS, access to treatment options varies widely, baricitinib would go some way to standardising treatment for those with severe AA.
- The percentage of scalp hair loss is only one element of living with AA, loss of lashes, brows, nasal and body hair matters too, baricitinib seems to provoke hair growth everywhere.
- The psychological and social impacts of having AA are debilitating and can be life limiting; the quality-of-life measures do not adequately represent this; a lack of meaningful measure must not stop progress.
- Many patients are dismissed from NHS care without receiving treatment options and without hope of living a socially normal life, baricitinib would offer a solution to those who progress to the most severe forms of this disease.

Thank you for your time.

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.

Patient expert statement

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating severe alopecia areata [ID3979]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

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.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. 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.

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 redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Thursday 1 December 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

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.

Technical engagement response form

About you

Table 1 About you

Your name	[redacted] on behalf of the British Association of Dermatologists' Therapy & Guidelines sub-committee and [redacted] on behalf of the British Hair and Nail Society and BAD guideline development group for managing people with alopecia areata.
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	British Association of Dermatologists (the BAD)
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Issue 1: Definition of the comparator Which of the following is considered to be standard of care in the treatment of severe alopecia areata? • 'Watch and wait' comprising no active treatment and frequent monitoring? Treatment with diphenylcyclopropenone (DPCP)? • Treatment with a basket of low-effectiveness non-DPCP therapies including systemic corticosteroids and immunosuppressants?	No	• 'Watch and wait' would not be an option in severe alopecia areata. • Treatment with diphenylcyclopropenone (DPCP) is used in severe alopecia areata but this treatment only addresses the hair loss on the scalp. It would not address the loss of eyebrows/eyelashes or body hair. DPCP is also not readily available in all regions. For patients with severe alopecia areata, i.e. extensive scalp loss and loss of eyelashes or eyebrows or facial/body hair, we often have to opt for systemic corticosteroids and/or immunosuppressants such as ciclosporin/methotrexate/azathioprine and mycophenolate mofetil. These have very limited evidence in the context of alopecia areata and are not licensed treatments for this. There is low effectiveness or high rate of severe relapse. We therefore have to weigh carefully their side-effects profile versus their benefits for alopecia patients. These agents also require blood test monitoring weekly initially and then less frequently as well as clinic reviews in secondary care frequently therefore these costs should be factored in. • Very few patients would opt for 'no treatment' and the main reason for this is due to the lack of reliably efficacious therapies or lack of access to DPCP. Even if they choose no pharmacotherapy for their alopecia areata, they may still need wig prescriptions (the frequency of this varies) which

• No active treatment and discharge from care?		has to come from dermatology, therefore we would not discharge them from our care.
Issue 2: Definition of treatment response at Week 36 What is considered a clinically meaningful change in the Severity of Alopecia Tool (SALT)? • SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair) • $SALT_{50}$: at least a 50% improvement from baseline SALT score • $SALT_{75}$: at least a 75% improvement from baseline SALT score	No	The outcome scores set by the trial was for SALT≤20: scalp hair loss of no more than 20% (or at least 80% scalp coverage with hair). What is considered clinically meaningful change to patients can vary as they may find meaningful change as being able to conceal their hair loss or not needing to wear a wig. However, a patient could have less than 20% hair loss but if the patches are in highly visible areas, then this will have significant psychosocial impact. Also, the lack of eyebrows and eyelashes can have a huge detriment on patients despite having achieved a SALT score of ≤20%.
Issue 3: Source of utilities in the model • For the majority of people with severe alopecia areata, would their health-related quality of life be similar to the general population? • Which utility value best represents health-related quality of life for people with severe alopecia areata during the induction phase	No	Patients with alopecia areata have significant disease burden which has been found to be higher than psoriasis and melanoma patients on a global scale (Hay R *et al.*, Global Burden of Skin Disease, *J Invest Dermatol* 2014). Their quality of life is severely impacted with higher rates of depression/anxiety, suicidal ideation and sickness from employment compared with controls (Macbeth A *et al.*, BJD Feb 2022). Therefore, their health-related QoL is not similar to the general population.

of treatment? Value from Adelphi DSP study (EAR, Table 29) or value using pooled data from BRAVE-AA1 and BRAVE-AA2 (EAR, Table 30)?		
Issue 4: Disease monitoring costs for best supportive care • What happens to patients whose condition does not respond to treatment? • Do patients continue to engage in further treatment? If not, are they discharged from care?	No	Patients who may not respond to baricitinib, despite optimisation, could be trialled on another JAK inhibitor. There are reports of patients having poor or incomplete response to one JAK inhibitor but show a better response with a different agent. We have certainly seen this in the eczema cohorts as well in the UK. Patients may reconsider the logistics/option to travel further away to trial DPCP as an alternative. If they fail these agents, it is likely these patients will require wig prescriptions for life, and these would need to be done by dermatology. Therefore, they would not be discharged unless they wish to not have any hair piece support. Wigs are very expensive, so patients do require input from hospital orthotics.
Other issues identified by the NICE technical team (not included in the EAR): Issue 5: Where would baricitinib and its comparators typically be commissioned in NHS practice?	No	DPCP and immunosuppression for alopecia areata takes place in secondary care dermatology clinics. Baricitinib is already being prescribed by dermatologists for atopic dermatitis, so they are familiar with this treatment and its monitoring requirements.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

Baricitinib for treating severe alopecia areata [ID3979]

Technical engagment response

December 2022

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135648.

1 Introduction

This document contains the Evidence Assessment Group's (EAG's) critique of the company's response to technical engagement (TE) for the single technology appraisal (STA) of baricitinib for treating severe alopecia areata (AA) [ID3979]. Table 1 summarises the company's position and EAG's critique for each issue identified for TE, and the EAG responds to each unresolved issue in the TE report in Section 2. The company's updated base case analysis is presented in Table 2 and the EAG's preferred analysis following TE is reported in Section 3.

Table 1. Issues for TE and EAG critique

Key Issue		Company approach	EAG approach
1	Definition of the comparator.	The company have removed monitoring costs included in the Induction and Maintenance health states for the comparator arm, in-line with the EAG's preference.	Resolved. Company is aligned with the EAG.
2	Definition of treatment response at Week 36.	Updated the definition of response to SALT≤20, in-line with the EAG's preference.	Resolved. Company is aligned with the EAG.
3	Source of utilities in the model.	Provided further justification for the greater validity of the Adelphi DSP EQ-5D data over the BRAVE-AA trial data and provided further comments on the lack of face validity of EQ-5D data for AA in general.	Unresolved. The EAG continues to prefer the EQ-5D utility data directly recorded in the BRAVE-AA clinical trials.
4	Disease management and monitoring costs for best supportive care.	The company has provided additional scenarios based on the proportion of patients receiving BSC treatments after non-response to treatment, and the relative reduction in BSC following baricitinib vs following 'no active treatment'.	Unresolved. The EAG continues to prefer the assumption of no disease management and monitoring costs in the BSC health state for both arms of the model.

Abbreviations: AA, alopecia areata; BSC: best supportive care; EAG: Evidence Review Group; SALT, Severity of Alopecia Tool; TE: technical engagement.

In addition to resolving key issue 1 and 2, the company accepted the following EAG preferred assumptions:

- Long-term all-cause discontinuation based on Week 36-52 data for baricitinib 4 mg.
- Removal of non-pharmacological psychological support costs.
- One wig assumed in the induction phase for both arms of the model.

The company's updated base case post technical engagement is presented in Table 2. The change that has the biggest, positive impact on the incremental cost-effectiveness ratio (ICER) is the company's assumption of reduced disease management costs for baricitinib patients in the best supportive care (BSC) health state (discussed in Section 2.1.1 of this report).

Results reported include the company's proposed patient access scheme (PAS); a fixed pack price of [redacted]. Confidential medicines unit (CMU) and Drugs and pharmaceutical electronic market information tool (eMIT) prices are available for medicines included in BSC, as such the EAG has produced a confidential appendix for this document.

Table 2. Company's base case results post-technical engagement

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
'No active treatment'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	Dominant
Probabilistic results							
'No active treatment'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	Dominant

Abbreviations: ICER, incremental cost effectiveness ratio, LYG, life year gained; QALY, quality adjusted life year.

2 Unresolved issues post-technical engagement

2.1 Key Issue 3: Source of utilities in the model

2.1.1 Source of EQ-5D data

In response to TE, the company have reaffirmed their preference for the Adelphi Disease Specific Programme (DSP) EQ-5D data over the EAG's preferred source of utility data: EQ-5D measured directly from patients in the BRAVE-AA trials. Compared to the BRAVE-AA trial data (baseline EQ-5D utility across the baricitinib 4 mg and placebo arms: [redacted]), the Adelphi DSP sample size reported a lower EQ-5D utility score for adults with severe or very severe AA [redacted].

In the Evidence Assessment Report (EAR), the EAG considered the BRAVE-AA trial utility data to be more robust and more suitable for decision making than the Adelphi DSP data because:

- The BRAVE-AA trials (N=860 severe or very severe patients in the baricitinib 4 mg or placebo arms) had [redacted] the sample size of the Adelphi DSP study (N=[redacted] severe or very severe patients);
- The EAG assessed the BRAVE-AA trials to be high-quality trials, but the company did not provide sufficient methodological detail about the Adelphi DSP study to enable a detailed quality assessment;
- The BRAVE-AA trials reported the within-patient change in EQ-5D following SALT response, which matches the structure of the economic model, whereas the Adelphi DSP study only reported between-patient differences in EQ-5D;
- While a small and heterogenous proportion of the patient population may be greatly affected by their severe AA, at the population level this may not lead to large HRQoL gains following treatment with baricitinib 4 mg because:
    - Only a minority of patients treated with baricitinib 4 mg achieved a SALT≤20 response in the BRAVE-AA trials;
    - Many patients have high baseline EQ-5D scores;
    - Treatment with baricitinib is non-curative;
    - Hair regrowth does not necessarily lead to greatly improved HRQoL, especially if other sources of reduced HRQoL, e.g. depression and/or anxiety, are not directly treated.

In the company response to TE, the company restated their arguments that:

- Because of a ceiling effect in the BRAVE-AA trial data, little improvement in HRQoL was possible for patients in the BRAVE-AA trials and as such the Adelphi DSP data should be preferred;
- Patient HRQoL in the BRAVE-AA trials was not representative of UK patients because patients with a history of anxiety or depression may have not be put forward for screening for the BRAVE-AA trials;
- Evidence of impaired HRQoL in AA, in-line with the Adelphi DSP data, can be found in the literature;
- The EQ-5D descriptive system lacks validity for capturing HRQoL changes associated with AA.

The EAG responds to each of these arguments below.

2.1.1.1 A ceiling effect in the BRAVE-AA trial EQ-5D data makes these data unsuitable for decision making

The company presented new data, requested by the EAG, that showed that while [redacted] of patients in the BRAVE-AA trials had "perfect-health", i.e., a score of 11111 across the EQ-5D, only [redacted] of patients in the Adelphi DSP reported a score of 11111. The EAG does not consider a larger proportion of patients reporting perfect health in the BRAVE-AA trial EQ-5D data to make these data less suitable for use in economic modelling than the Adelphi DSP data, because the trial data are an accurate reflection of the average EQ-5D utility score in the BRAVE-AA patient population, which is also linked to the treatment outcomes used in the model. Rather than being an artefact of the trial design, the EAG considers the utility data reported from the BRAVE-AA trials to appropriately capture the variable impact of severe AA on patients' quality of life, with a substantial proportion of patients reporting high EQ-5D-5L scores.

2.1.1.2 The Adelphi DSP patient population is more representative of adults with severe AA in UK clinical practice than the BRAVE-AA trial populations

The company claimed that patients with a history of anxiety or depression may not have been put forward for screening for the BRAVE-AA trials and, because of this, the Adelphi DSP patient population is more representative of adults with severe AA in UK clinical practice than the BRAVE-AA trial populations.

The EAG accepts that some patients with a history of anxiety or depression may have been less likely to be put forward for screening in the BRAVE-AA trials than patients without such a history, however the EAG notes that:

- The company has only provided clinical expert input, rather than data, to support this, and;
- [redacted] because of significant uncontrolled neuropsychiatric disorders.

The EAG considers it implausible that the number of patients not put forward for the BRAVE-AA trials due to having anxiety/depression would have been large enough to make the BRAVE-AA trial population not representative of UK clinical practice, to a large degree. To explore this, the EAG calculated the number of patients who would have not been put forward for screening in the BRAVE-AA trials in order to reduce the BRAVE-AA baseline EQ-5D utility value [redacted] to match that of the Adelphi DSP study [redacted] under a variety of assumptions about the missing patients' average EQ-5D scores (Figure 1). Even at the most extreme scenario considered by the EAG, i.e., the average of missing patients EQ-5D total utility scores was 0.50, 390 patients would be needed to be added to the baricitinib 4 mg or placebo arms of the BRAVE-AA trials in order to produce a baseline EQ-5D equal to the Adelphi DSP trial, i.e. an increase in the size of the trials by 45%. Not only does the EAG consider this scenario implausible — both in the number of missing patients and the required average EQ-5D score of "missing" patients — the EAG notes that if such a large number of patients were missing from the BRAVE-AA clinical trials, this would raise serious concerns about the suitability of the BRAVE-AA trial data to inform decision making.

Figure 1. The number of patients needed to be added to the BRAVE-AA trial baricitinib 4 mg and placebo arms to reduce baseline EQ-5D to the Adelphi DSP level, for different mean EQ-5D values for the additional patients

Abbreviations: DSP: disease specific programme; EQ-5D: EuroQol-5 dimensions

In contrast, the EAG does consider the BRAVE-AA trial data to be suitable to inform decision making and considers it plausible that a number of biases in the Adelphi DSP study likely caused the Adelphi DSP data to underestimate the HRQoL of patients with severe AA in UK clinical practice. Specifically, in the Adelphi DSP study:

- Patients were recruited by their expert dermatologists, meaning that the majority patients with severe AA who were not being routinely seen by an expert dermatologist were unlikely to have been included in the study;
- Patients had to opt into to completing a survey response about the effects of AA on their QoL, which is a source of selection bias;
- When completing the Adelphi DSP questionnaire, patients answered questions on their AA history and symptoms prior to completing the EQ-5D measure, which may lead to different responses than if patients were only asked to rate their HRQoL;

- Patients could not be involved in a clinical trial at the time of the survey. As three large clinical trials in patients with severe AA were ongoing at the time of the Adelphi DSP study (BRAVE-AA1,[1] BRAVE-AA2[1] and ALLEGRO-LT,[2] with a combined N=2360 and estimated completion dates after 2024) a substantial number of severe AA patients were ineligible for the Adelphi DSP studies, not accounting for other ongoing clinical trials. The EAG notes that if the company's argument concerning the relationship between patient HRQoL and the likelihood of entering clinical trials were to hold, then the opposite bias would hold in the Adelphi DSP data: patients not put forward for clinical trial screening, but who were actively engaging with care, would be overrepresented in the Adelphi DSP data. These patients, *per* the company's argument, would be expected to have lower EQ-5D scores than the "average" severe AA patient.

As such, the EAG considers there to be a larger potential for selection and response bias within the Adelphi DSP study than the BRAVE-AA trials, and therefore does not consider there to be evidence that the Adelphi DSP population is more representative of serve AA patients in UK clinical practice than the BARVE-AA population.

2.1.1.3 Additional literature references and HRQoL SLR

The EAG notes that the company performed a systematic literature review (SLR) to identify relevant HRQoL studies in Appendix D of the original submission, but did not present the results further in the main text. The EAG therefore presents the EQ-5D overall score data from the SLR in Table 3, alongside the EQ-5D data from BRAVE-AA1, BRAVE-AA2, the Adelphi DSP study, and a further reference provided by the company at TE, Edson-Heredia *et al*.[3] Notably, all additional sources reported higher EQ-5D overall scores than the Adelphi DSP data used in the CS, with one company-sponsored study of HRQoL in AA in a real-world USA sample providing data directly in-line with the BRAVE-AA trial. The EAG is concerned that such data were available to the company at the time of the submission, but that these were not reported in any detail.

Table 3. EQ-5D overall score data identified from the company's HRQoL SLR, presented alongside the BRAVE-AA1 and BRAVE-AA2 trial data, in addition to the Adelphi DSP data.

Study	EQ-5D Overall Score	N severe or very severe	Location	Design	EAG comment
BRAVE-AA1[a]	[redacted]	470	International	Randomised controlled-trial	

BRAVE-AA2[a]	[redacted]	390	International	Randomised controlled-trial	High-quality, Phase 3 international randomised placebo-controlled trials.
Burge *et al.*	0.87[b]	NR	US	Cross-sectional survey	Company sponsored HRQoL in AA survey, performed in the USA. Uses Adelphi-DSP US data.
Edson-Heredia *et al.*	0.79[b]	85	Japan	Cross-sectional survey	Company sponsored HRQoL in AA survey, performed in Japan. Uses Adelphi-DSP Japanese data.
Adelphi DSP	[redacted]	[redacted]	EU5	Cross-sectional survey	Company's preferred source of the EQ-5D utilities, assessed by EAG to be at high risk of bias.

[a]Data from the baricitinib 4 mg and placebo arms only

[b]Severe or very severe patients only

Abbreviations: AA: alopecia areata; DSP: disease specific programme: EAG; evidence assessment group; HRQoL: health-related quality of life; SLR: systematic literature review

Sources: CS Appendix D Table 27

The company also presented a narrative overview of three studies to provide evidence supporting a HRQoL impairment associated with AA. However, the EAG considers the data from the three additional studies reported by the company, Balieva *et al.*,[4] Titeca *et al.*[5] and Rencz *et al.*,[6] to be at high risk of bias relative to the current appraisal and unsuitable to inform decision making. A critique of these studies is presented in the Appendix.

2.1.1.4 Validity of the EQ-5D in assessing HRQoL in AA

In the response to TE, the company reasserted that there may be limitations in the suitability of the EQ-5D instrument in assessing HRQoL in AA. The EAG's clinical experts agreed that the HRQoL impact of severe AA may only manifest in the anxiety/depression domain of the EQ-5D, as the condition may not affect, for example, self-care or mobility. While the EAG recognises that some domains of the EQ-5D may be relatively unaffected in patients with severe AA, the EAG considers high scores in these domains are likely to accurately reflect patients' HRQoL in clinical practice, and notes that the EQ-5D is the utility method preferred for technology appraisal by NICE.[7] The EAG maintains that the company has not provided sufficient evidence that the EQ-5D performs poorly on tests of construct validity or responsiveness in the severe AA population. The EAG notes the contrast between the company's statement that, "*[generic] questionnaires are not fit for purpose to assess*

HRQoL in patients with severe AA", and the company's continued use of EQ-5D utility data from the Adelphi DSP in the economic models.

2.1.2 Within- versus between-person changes in HRQoL

In addition to considering the BRAVE-AA trial utility data to be more robust and more representative of UK clinical practice than the Adelphi DSP data, the EAG highlights how the two sources provide different data:

- The BRAVE-AA trials provide data on the within-person change in HRQoL following different degrees of hair regrowth and baricitinib 4 mg or placebo treatment;
- The Adelphi DSP provides data on the between-person difference in HRQoL of patients with different severities of AA.

The EAG considers the BRAVE-AA trial data to appropriately reflect the structure of the economic model, whereas the Adelphi DSP data must make the additional assumption that the between-person difference in HRQoL of patients with different AA severities is equivalent to the within-person change in HRQoL when their AA severity changes, which may not hold because of:

- Differences between individuals who have experienced severe AA and those who have not, including disease severity;
- Treatment with baricitinib is not curative and HRQoL in severe AA patients who achieve a SALT≤20 response of life may continue to be affected by: i) still having the underlying condition; and ii) the possibility that hair regrowth might be lost;
- Hair regrowth to SALT≤ 20 may not necessarily improve all aspects of a patient's reduced HRQoL.

2.1.3 ALLEGRO trial data

The EAG also notes that a small amount of EQ-5D data, in the form of a qualitative statement from the ALLEGRO trial clinical study report synopsis, have also been made publicly available. The ALLEGRO Phase 2b/3 trial evaluates the safety and effectiveness of ritlecitinib for treating severe AA.[8] In ALLEGRO, the overall response SALT≤20 response rate was similar to the BRAVE-AA trials, and, [redacted], the company reported that, "*From Weeks 4 to 24… EQ-5D-5L in adults… did not change*".

2.1.4 Summary

In summary, the EAG considers the BRAVE-AA trial data to provide more robust utility data that is more representative of UK clinical practice and is associated with lower risk of bias than the Adelphi DSP data. The EAG considers that Adelphi DSP data to be at high risk of selection bias, and notes that the company provided the EAG with insufficient information on the study to perform a thorough risk of bias assessment. In addition, the EAG considers the BRAVE-AA trial data to be directly linked to the treatment response used in the economic model, whereas assumptions about the likely consistency of within- and between-person data must be made when using the Adelphi DSP data — assumptions the EAG considers unlikely to hold.

The EAG recognises that some patients do experience large HRQoL deficits associated with their severe AA, and this was highlighted by the EAG's clinical experts. However, when combined with the low overall response rate to baricitinib treatment, the EAG consider the expected HRQoL of baricitinib treatment at the population level to be small, and to be more appropriately captured by the BRAVE-AA trial data than the Adelphi DSP data.

2.2 Key Issue 4: Disease management costs for best supportive care

In the model, when patients fail to achieve treatment response (now defined as SALT≤20 in the company's base case post technical engagement), they transition to the best supportive care (BSC) health state and are assumed to receive a basket of treatments (presented in Table 32 of the EAR). As such, costs associated with drug acquisition and monitoring were included in the BSC health state. Patients in the BSC state remain there until the end of the model time horizon or death. Therefore, the contribution of costs incurred in the BSC health state make up the bulk of the total costs in the model.

As noted in the EAR, the EAG's clinical experts considered that a range of treatments may be given to patients but that these are not very effective. Additionally, in the company's own Adelphi DSP study, it was estimated that the majority of severe/very severe patients were treatment experienced ([redacted]). Therefore, it is likely that non-responders will not engage with further treatment (as these would have likely been exhausted) or will not be followed up (effectively patients are discharged from care). It should be noted that even though frequency of treatments based on their summary of product characteristics (SmPCs) were included for treatments such as DPCP, annual drug acquisition costs were applied in the model. Combined with the lifetime horizon and the fact that once patients

enter the BSC health state, they remain there until death, it is clinically implausible that treatments (which are deemed to have limited effectiveness) in the BSC health state will be given for such a long duration of time. Additionally, the EAG considers that a significant proportion of patients may not take up treatment and instead opt for using wigs or hair removal to manage their hair loss.

Given the issues around further treatment for non-responders, the EAG's preferred approach was to exclude disease management costs in the BSC health state (that is, the costs of drug acquisition and monitoring and ongoing management costs) from both arms of the model. In the company's response to technical engagement, they considered that the EAG's preferred approach to exclude disease management costs from the BSC health state was not plausible as it is unlikely that all non-responders would not go on to receive further treatment. The company did consider that for those patients who failed on baricitinib, a proportion would not be prescribed further treatment and as such, updated their base case to include an assumption of a reduction of 50% in the use of treatments included in the BSC health state for baricitinib patients. As a reminder, in the company's base case post clarification, it was assumed that in both arms of the model, [redacted] of patients in the BSC health state would engage with further treatment and [redacted] do not go on to have further treatment, based on data from the Adelphi DSP study, and these percentages remain unchanged for the comparator arm ('no active treatment').

The company's new base case assumption of a reduction in the use of treatments and ongoing disease management included in the BSC health state for baricitinib patients removes a substantial amount of costs, resulting in baricitinib being cost-saving and thus reducing the ICER from £18,072 to dominant. Additionally, the company provide a range of scenarios exploring varying reductions in the use of BSC treatments for both arms of the model (Table 4 of the company's technical engagement response). Though, the EAG notes that the company's reduction in BSC costs for the baricitinib arm is relative to the proportion included in the 'no active treatment' arm, such that the proportion of baricitinib patients incurring BSC costs is always less than the comparator.

The EAG considers that for patients who do not achieve hair growth, as defined by the primary outcome of SALT≤20, irrespective of whether they had baricitinib or not, the decision to engage with further treatment is likely to be same. Thus, the EAG maintains its position that if after Week 36, if the outcome of SALT≤20 is not achieved, for either arm of the model, patients are effectively discharged from care (except for the provision of wigs and orthotics) and do not incur disease management costs. The EAG does not agree with the company that assuming baricitinib patients

incur less costs due to the reduced use of BSC treatments and on-going management is associated with less decision risk, as this is the assumption that has the greatest, positive impact on the ICER in favour of baricitinib. The EAG considers that assuming use of BSC treatments and on-going management is removed or reduced equally between both arms limits the decision risk. The EAG has supplied alternative scenarios exploring equal reductions in the use of BSC costs for both arms of the model, presented in Section 3.1.

2.3 Adverse events

In the EAR, the EAG requested further details on the estimation of adverse event (AE) rates used in a scenario supplied by the company during the clarification stage. In their TE response, company confirmed that AEs included in the scenario supplied in response to question B21 of the clarification response were based on treatment-emergent adverse events (TEAEs) based on data from the BRAVE trials. Costs used in the scenario were sourced from NHS reference costs.[9]

The EAG still unclear what threshold was used to determine the TEAEs to be included (i.e. ≥2% of patients in either arm) or the number of patients informing the proportions presented in Table 28 of the EAR. In Table of 28 the company submission, TEAEs affecting ≥2% of patients in either arm are presented and include a greater range of AEs than that used in the economic model.

Nonetheless, the EAG considers that the impact of AEs on costs should be included in the cost-effectiveness analysis and ran a scenario including costs of AEs on the company's updated base case (presented in Section 3.1). As the impact on the ICER was minimal and it is methodically correct to include AEs in the analysis, the EAG has included this assumption in the EAG base case for the committee to consider, presented in Section 3.2. However, the inclusion of AE costs can be considered illustrative as details informing the proportions for AEs in the model are still lacking.

3 Additional economic analysis undertaken by the EAG

3.1 EAG scenario analysis

Table 4 presents the deterministic results of the EAG exploratory analyses described in Section 2. Results reported include the company's proposed patient access scheme (PAS); a fixed pack price of [redacted]. Confidential medicines unit (CMU) and Drugs and pharmaceutical electronic market information tool (eMIT) prices are available for medicines included in BSC, as such the EAG has produced a confidential appendix for this document.

Table 4. EAG scenario analyses

	Results per patient	Baricitinib 4 mg	'No active treatment'	Incremental value
0	**Company base case post-technical engagement**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			Dominant
1	**Proportion of patients receiving BSC treatments - 0% (EAG preferred assumption)**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			61,052
2	**Proportion of patients receiving BSC treatments - 25%**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			50,569
3	**Proportion of patients receiving BSC treatments - 50%**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			40,085
4	**Inclusion of AE costs**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			Dominant
5	**SALT≤20 baseline and CFB utility from BRAVE trials**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			Dominant

Abbreviations: AE, adverse event; BSC, best supportive care; EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year.

3.2 EAG preferred assumptions

As mentioned in Section 1 of this report, the company accepted several the EAG's preferred assumptions, but the following remain outstanding and are still included in the EAG's preferred base case:

- Baseline and change from baseline utility values associated with achieving SALT≤20 at Week 36 from the BRAVE trials.
- Removal of disease management costs in the BSC health state for both arms of the model.

In addition, the EAG has included adverse event costs in the EAG preferred base case.

Table 5 presents the impact of each assumption on the ICER and Table 6 presents the EAG's deterministic and probabilistic base case results. Table 7 presents the severity subgroup analysis around the EAG base case but it should be noted that probabilistic subgroup results could not be obtained due to a problem with the probabilistic sensitivity analysis (PSA) function in the model.

In the EAG base case probabilistic analysis, an incremental quality-adjusted life-year (QALY) gain of [redacted] over 'no active treatment' along with additional costs of [redacted] for the baricitinib 4 mg, generates an ICER of £544,945 per QALY. The net monetary benefit (NMB) using the £30,000 threshold is [redacted] and the net health benefit (NHB) is [redacted]. The EAG considers that the ICERs are highly sensitive due to the small incremental costs and quality-adjusted life-year (QALY) gain, such that small changes can cause a substantial impact.

Additionally, as mentioned in Section 4.2.8.1 of the EAR, the EAG acknowledges that there is a small, but heterogenous, patient population that is more adversely affected in terms of health-related quality of life (HRQoL) but that the demographics of this population are difficult to identify clinically and consistently, and it is beyond the scope of this assessment to identify that group. Nonetheless, the EAG ran two scenarios around the EAG base case and severity subgroup analysis to identify the QALY gain needed for the ICER to reach the £20,000 and £30,000 cost-effectiveness threshold and these are presented in Table 8. The results of the threshold analysis demonstrate that for the overall population, a QALY gain of [redacted] to [redacted] is needed for the ICER to be within the £20,000 to £30,000 threshold. Thus, the EAG advises the committee to consider if the estimated QALY gain needed for baricitinib 4 mg to be cost-effective is plausible for the condition under consideration.

Table 5. EAG's preferred model assumptions - FAS population

Preferred assumption	Incremental costs	Incremental QALYs	Deterministic ICER (£/QALY)	Cumulative ICER (£/QALY)
Company base case post-technical engagement	[redacted]	[redacted]	**Dominant**	**-**
SALT≤20 baseline and CFB utility from BRAVE trials	[redacted]	[redacted]	Dominant	Dominant
Removal of disease monitoring costs in the BSC health state for both arms of the model	[redacted]	[redacted]	61,052	423,775
Inclusion of AE costs	[redacted]	[redacted]	Dominant	425,532
EAG preferred base case	[redacted]	[redacted]	**425,532**	**-**

Abbreviations: AE, adverse events; BSC, best supportive care; CFB, change from baseline; EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

Table 6. EAG's base case post-technical engagement

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
'No active treatment'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	425,532
Probabilistic results							
'No active treatment'	[redacted]	22.60	[redacted]	-	-	-	-
Baricitinib 4 mg	[redacted]	22.60	[redacted]	[redacted]	0.00	[redacted]	544,945

Abbreviations: ICER, incremental cost effectiveness ratio, LYG, life year gained; QALY, quality adjusted life year.

Table 7. Deterministic scenarios around the EAG base case

	Results per patient	Baricitinib 4 mg	'No active treatment'	Incremental value
0	**EAG base case post-technical engagement**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			425,532
1	**Severe subgroup - baseline SALT 50-95**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			408,951
2	**Very severe subgroup - baseline SALT 95-100**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			458,365

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

Note: the same baseline utility ([redacted]), change from baseline ([redacted]) and treatment discontinuation rate ([redacted]) have been used for the subgroups as for the base case as the relevant data were not available by severity.

Table 8. Threshold analysis on QALY gain needed for £20,000 to £30,000 cost-effectiveness threshold

Population	QALY gain - £20,000 threshold	QALY gain - £30,000 threshold
Full analysis set - baseline SALT 50-100	[redacted]	[redacted]
Severe subgroup - baseline SALT 50-94	[redacted]	[redacted]
Very severe subgroup - baseline SALT 95-100	[redacted]	[redacted]

Abbreviations: EAG, Evidence Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; SALT, Severity of Alopecia Tool.

4 References

1. King B, Ohyama M, Kwon O, Zlotogorski A, Ko J, Mesinkovska NA, et al. Two Phase 3 Trials of Baricitinib for Alopecia Areata. *The New England Journal of Medicine* 2022.
2. ClinicalTrials.gov. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT), 2022. Available from: https://clinicaltrials.gov/ct2/show/NCT04006457. Date accessed: December 2022.
3. Edson-Heredia E, Aranishi T, Isaka Y, Anderson P, Marwaha S, Piercy J. Patient and physician perspectives on alopecia areata: A real-world assessment of severity and burden in Japan. *J Dermatol* 2022; 49: 575-83.
4. Balieva F, Kupfer J, Lien L, Gieler U, Finlay AY, Tomás-Aragonés L, et al. The burden of common skin diseases assessed with the EQ5D™: a European multicentre study in 13 countries. *Br J Dermatol* 2017; 176: 1170-8.
5. Titeca G, Goudetsidis L, Francq B, Sampogna F, Gieler U, Tomas-Aragones L, et al. 'The psychosocial burden of alopecia areata and androgenetica': a cross-sectional multicentre study among dermatological out-patients in 13 European countries. *Journal of the European Academy of Dermatology and Venereology* 2020; 34: 406-11.
6. Rencz F, Gulacsi L, Pentek M, Wikonkal N, Baji P, Brodszky V. Alopecia areata and health-related quality of life: a systematic review and meta-analysis. *Br J Dermatol* 2016; 175: 561-71.
7. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. Process and methods [PMG36]. 2022. Available from: https://www.nice.org.uk/process/pmg36/ Date accessed: Sept 2022.
8. Pfizer. B7981015 Clinical Study Report Synosis: Public Disclosure Synopsis, 2022. Available from: https://www.pfizer.com/node/554011. Date accessed: December 2022.
9. National Health Service (NHS). 2020/21 National Cost Collection data. 2020. Available from: https://www.england.nhs.uk/national-cost-collection/.
10. Dalgard FJ, Gieler U, Tomas-Aragones L, Lien L, Poot F, Jemec GBE, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. *J Invest Dermatol* 2015; 135: 984-91.
11. de Hollanda TR, Sodre CT, Brasil MA, Ramos ESM. Quality of life in alopecia areata: a case-control study. *Int J Trichology* 2014; 6: 8-12.
12. Gulec AT, Tanriverdi N, Duru C, Saray Y, Akcali C. The role of psychological factors in alopecia areata and the impact of the disease on the quality of life. *Int J Dermatol* 2004; 43: 352-6.
13. Dubois M, Baumstarck-Barrau K, Gaudy-Marqueste C, Richard MA, Loundou A, Auquier P, et al. Quality of life in alopecia areata: a study of 60 cases. *J Invest Dermatol* 2010; 130: 2830-3.
14. Masmoudi J, Sellami R, Ouali U, Mnif L, Feki I, Amouri M, et al. Quality of life in alopecia areata: a sample of tunisian patients. *Dermatol Res Pract* 2013; 2013: 983804.

5 Appendix

5.1 EAG critique of company's narrative overview of reduced HRQoL in AA

The company presented three studies suggesting patients with AA have reduced HRQoL compared to healthy controls. The EAG noted that two of the studies Balieva *et al.*[4] and Titeca *et al.*[5] reported on the same cohort of patients, and as such are not independent data sources. These studies reported on 37 (Titeca *et al.*) and 31 (Balieva *et al.*) AA patients from the Dalgard *et al.* cohort,[10] which was a case-control study assessing adult dermatology outpatients (cases) and hospital employees (control). From these studies, the EAG notes that:

- The severity of AA is not reported;
- The median EQ-5D of the AA patients (Figure 3c of Titeca *et al.*) is around 0.88, [redacted];
- For case-control comparisons, there is likely residual confounding;
- As the number of cases of AA are low, the odds ratio of 4.19 for EQ-5D depression/anxiety in cases vs controls is highly uncertain, even excluding likely confounding.

The other study cited by the company, a systematic review, Rencz *et al.,*[6] reported significant impairment in HRQoL for AA patients compared to controls on a variety of dermatology specific questionnaires and SF-36. In Rencz *et al.*, four studies reporting deficits in SF-36 were reported:

- A study comparing 37 Brazilian AA patients (of which 14 had ≥50 hair loss) with 49 voluntary blood donors at the same hospital;[11]
- A study comparing 52 AA patients with hospital employee controls;[12]
- A letter comparing 60 French AA patients (median scalp surface involvement 77%) to age- and sex-matched controls, but with no covariate adjustment;[13]
- A study comparing 50 newly diagnosed Tunisian AA patients (20% with ≥50% scalp involvement), with 50 healthy control patients, but with no covariate adjustment.[14]

Hence, in contrast to the company's suggestion that the Rencz *et al.* meta-analysis, "*demonstrated that SF-36 outcomes were significantly poorer among AA patients than the general population*", it instead is a meta-analysis of four case-control studies with unique control groups. Notably, Rencz *et al*. did not report a risk of bias assessment for any of the included studies.

While the EAG recognises the general paucity of data concerning AA, especially regarding treatment patterns and HRQoL, the EAG does not find the literature cited by the company to be acceptable to guide decision making.