Single Technology Appraisal

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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 **from Pfizer**
   a. Company summary of information for patients (SIP)

2. Clarification questions and company responses

3. Patient group, professional group and NHS organisation submissions **from:**
   a. Alopecia UK
   b. British Association of Dermatologists (BAD)*
   ***Royal College of Physicians endorse the submission from BAD**

4. External Assessment Report **prepared by ScHARR**

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 Dermatologists
   b. Dr Nekma Meah, Consultant Dermatologist – clinical expert, nominated by British Association of Dermatologists
   c. Mr Tony Ferguson – patient expert, nominated by Alopecia UK
   d. Lynn Wilks – patient expert, nominated by Alopecia UK

8. Technical engagement responses from stakeholders:
   a. **Alopecia UK**
   b. **British Association of Dermatologists (BAD)**

9. External Assessment Report critique of company response to technical engagement **prepared by ScHARR**

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

Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over [ID4007]

Document B

Company evidence submission

January 2023

File name	Version	Contains confidential information	Date
Ritlecitinib_Company evidence submission_ Doc B	1.3	Yes	12-02-2024

Contents

Tables and figures

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

Alopecia areata (AA) is a chronic autoimmune disorder characterised by non-scarring hair loss and a variable, unpredictable, and relapsing or remitting course.[2] (Section B.1.3.1)

- The exact pathophysiology of AA is not completely understood. It is thought to be caused by immune cells attacking hair follicles which elicits an inflammatory response leading to hair loss.[1,2]
- AA is one of the most prevalent autoimmune disorders.[3] A population-based cohort study reported an overall point prevalence of 0.58% for diagnosed AA among adults.[4] The point prevalence of clinician-adjudicated severe AA and alopecia totalis (AT)/ alopecia universalis (AU) is estimated to be 9 per 10,000 and 4 per 10,000, respectively.[5]
- A UK population-based cohort study found that people of non-white ethnicity were more likely to present with AA ($P < 0.05$), especially those of an Asian ethnicity (incident rate ratio (IRR) 3.32). AA incidence was also associated with social deprivation (IRR most vs. least deprived quintile 1.47 [$P < 0.05$]). In addition, people of higher social deprivation were less likely to be referred for specialist dermatology review.[4]

AA results in a substantial clinical, health-related quality of life (HRQoL) and psycho-social burden on patients, and has an impact on families, caregivers, and society. (Section B.1.3.2)

- Coping with AA is reported to be a daily challenge by 85% of patients.[6] The psychological impact on individuals is reported to lead to an increased risk of anxiety, reduced self-esteem, altered body image, social withdrawal and the breakdown of personal relationships.[3,7–12]
- In a study of patients with various dermatological conditions, a quarter of patients with AA reported having suicidal thoughts; suicidal ideation was significantly more prevalent in patients with AA than in patients with other dermatological conditions, including psoriasis, vitiligo and acne ($P = 0.001$).[13]

- AA also has a large impact on the quality of life of families and caregivers of patients with AA; one cohort study found that 69.9% of family members of adults with AA (aged 17+) and 87.2% of family members of children with AA (aged 4-16) experienced some HRQoL impairment, respectively.[14]
- Market research indicates that the current spending to treat AA by patients is between £50-£150 per month.[15] Although NICE do not typically consider indirect costs in health economic analyses, patients with AA who are of working age are 56% more likely to be issued with certificates for time off work and 82% more likely to be recorded as unemployed than the general population.[16]
- Dermatologists with a specialist interest in hair disorders indicated that patients would only be eligible to receive camouflage treatments such as wigs if they resided in catchment areas which allocate budget for AA management as a result of NHS budget constraints.[17] As people from more deprived backgrounds are more likely to have AA, these patients are likely to be disproportionately affected by out of pocket expenses in catchment areas that do not provide wigs.[4]

Ritlecitinib is being reviewed by the MHRA for the treatment of [redacted]. It is an orally bioavailable small molecule that inhibits JAK3 and the TEC kinase family with very high selectivity over the other three JAK isoforms, as well as over the broader human kinome. (Section B.1.2.1).

There is a significant unmet need for patients with severe AA. (Section B.1.3.3)

- There is currently only one licensed systemic treatment option for adult patients with severe AA, none for adolescents 12 years and over, and none that are currently approved by National institute for Health and Care Excellence (NICE). It has been reported that the likelihood of positive treatment outcomes reduces with severity of hair loss and duration of disease.[7,18,19]
- Current off-licence treatment options for severe AA are limited by their efficacy and safety profile, with no recent advancement.[7,20] Treatment can be uncomfortable, time consuming and have unacceptable side effects.
- There is dissatisfaction among both patients and dermatologists with a specialist interest in hair disorders with the current treatment options.[21,22] An epidemiological study in a UK primary care setting (N=2,634,083) has shown

that 46% of patients with AA did not receive any prescription medication for their AA which may result from the lack of treatment options currently available.[4]

B.1.1. Decision problem

The submission covers the technology's full marketing authorisation. The anticipated marketing authorisation for ritlecitinib ([redacted]) is for [redacted]

We have conducted several research projects to characterise the decision problem for ritlecitinib and support this submission, outlined in Table 1.

Table 1: Studies conducted by Pfizer to support this submission

Study	Purpose
Clinical practice and therapeutic landscape Delphi panel[21] (Therapeutic landscape Delphi panel)	To develop a clear and comprehensive understanding of current clinical practice and the therapeutic landscape for patients with AA in the UK; based on current guidance and the opinions of key clinical experts. Eight consultant dermatologists with a specialist interest in hair disorders took part in the expert elicitation.
Qualitative research in alopecia[22]	To represent the experience of patients with severe AA in the UK, qualitative interviews with 10 representatives from six Patient Advocacy Groups focused exclusively on patients with AA with ≥50%* hair loss living within the UK, as well as a review of the literature were conducted.
Vignette study for utility estimation in Alopecia Areata[23]	To estimate utilities for health states in AA vignettes were designed to describe how key domains of HRQoL are affected by the disease for adult and adolescent patients and caregivers of patients with AA.
Health care resource utilisation Delphi panel[24]	To develop a clear and comprehensive understanding of HCRU as well as direct and indirect costs associated with overall management of AA, based on the opinions of key clinical experts.

Abbreviations: AA, alopecia areata; HCRU, health care resource use; HRQoL, health-related quality of life

*In this submission we refer to patients with >50% scalp hair loss as patients with severe AA.

Given that there were no concluded appraisals in the AA space at the time of submission, we have sought extensive clinical opinion to support our submission including the Clinical Practice and Therapeutic Landscape Delphi panel (Therapeutic Landscape Delphi panel) of eight consultant dermatologists to support understanding of the decision problem for ritlecitinib. The Therapeutic Landscape Delphi panel was conducted in two rounds of surveys followed by supportive 1:1 interviews. The first-round survey was informed by a pragmatic literature review of peer-reviewed literature to identify treatment guidelines and pathways to understand the existing therapeutic

landscape in AA in the UK. The results from the first-round survey informed the second-round survey; data from both survey rounds informed the interviews. Three consultant dermatologists supported the development of our submission in addition to their contribution to the Therapeutic Landscape Delphi panel by providing clinical input and validating assumptions about UK clinical practice in a series of teleconference discussions.[17] Their biographies are provided in Appendix M.

Feedback from the Therapeutic Landscape Delphi panel, in which the majority said they use ≥50% scalp hair loss to define severe AA (7/8 clinicians, 88%), indicates that UK clinicians use ≥50% scalp hair loss as a threshold to define severe AA.[21] Additionally, the 50% scalp hair loss threshold for severe AA has been endorsed by the Committee for Medicinal Products for Human Use (CHMP) and is consistent with the recent research on adult and adolescent patients with AA.[1, 20,25–28] Further literature refers to ≥50% scalp hair loss as extensive hair loss in patients with AA.[28,29] Therefore, 'severe', 'extensive', and '≥50% scalp hair loss' terminology can be used to describe patients with AA and are considered interchangeable definitions. In this submission, we refer to patients with ≥50% scalp hair loss as patients with severe AA.

Table 2: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People aged 12 years and over with severe AA	As per scope.	N/A
Intervention	Ritlecitinib	Ritlecitinib 50mg dose.	The dose proposed for registration for the treatment of AA is the 50 mg once daily dose.
Comparator(s)	Established clinical management without ritlecitinib.	BSC.	BSC is established clinical management.
Outcomes	The outcome measures to be considered include: • severity of alopecia areata • percentage of area affected by hair loss • adverse effects of treatment • health-related quality of life	As per scope.	N/A
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 a NHS and PSS perspective.	As per scope. Consideration has also been given to indirect costs associated with absenteeism & presenteeism in Section B.3.5.4.1 and presented as a scenario in the economic model.	N/A
Subgroups to be considered	If evidence allows, subgroups based on severity and type of alopecia areata (e.g., alopecia totalis (AT) and alopecia universalis (AU)) will be considered.	The differences between each ritlecitinib group and placebo in the proportion of response based on SALT ≤20 at Week 24 were consistent across most pre-specified subgroups (age, BMI, weight, gender,	N/A

		race, region, severity of disease, duration since diagnosis, duration of current episode, prior pharmacological treatment for AA) for all doses. Furthermore, no subgroups have been considered in the economic analyses as N values are too small to draw any conclusions beyond consistency. Outcomes in the trial population split by age (adolescent and adult populations), previous treatment and race are provided in Appendix E.	
Special considerations including issues related to equity or equality	No consideration highlighted in the final scope.	There are inequalities in the characteristics of people diagnosed with AA (females, people of Asian ethnicity, people from deprived economic backgrounds and people diagnosed with Down Syndrome are more likely to be diagnosed with AA. Research suggests there are significant out of pocket expenses associated with an individual patient's management of AA. Considering that research also shows that AA follows an inverse social gradient and therefore, may pose an issue related to equity. See Section B.1.4 for further details.	N/A

Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; BSC, best supportive care; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PSS, personal social services

B.1.2. Description of the technology being evaluated

[redacted]

Table 3: Technology being evaluated

UK approved name and brand name	**Ritlecitinib ([redacted])**
Mechanism of action	Ritlecitinib is an orally administered small molecule that selectively inhibits JAK3 and also inhibits all 5 members of the TEC family (BTK, BMX, ITK, TXK and TEC). Ritlecitinib inhibits JAK3 and TEC by irreversibly blocking the ATP binding site on these proteins. Ritlecitinib has high selectivity for JAK3 and TEC over the other three JAK isoforms, JAK1, JAK2, and TYK2, as well as over the broader human kinome.[30–33] The JAK3 protein is part of a signalling pathway called the JAK/STAT pathway that mediates the transduction of intracellular signals involved in the process of inflammation. TEC kinases play a key role in B- and T-cell antigen receptor signalling. They also regulate development, activation, and differentiation of lymphocytes. Over-activation of the JAK/STAT pathway is involved in AA, as is typical of other autoimmune diseases and proliferative disorders. In AA this results in damage to hair follicles.[31] T-cells infiltrate the epithelial layers of the hair follicle causing an IFN-γ response and upregulation of several cytokines, breaking down hair follicle immune privilege (certain sites of the human body have immune privilege which means that they are able to tolerate the introduction of antigens without eliciting an inflammatory immune response) and promoting the survival and activity of T-cells in the affected skin, this infiltration of cells is known as the 'swarm of bees'.[34,35] Ritlecitinib works by inhibiting these signalling pathways, preventing the breakdown of the immune privilege of the hair follicle.
Marketing authorisation /CE mark status	[redacted] An application was submitted to the EMA on [redacted], with CHMP opinion expected in [redacted] and marketing authorisation anticipated to be granted in [redacted].

Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Anticipated Indication: Ritlecitinib is indicated [redacted]
Method of administration and dosage	Ritlecitinib is taken orally as a monotherapy. The dose of ritlecitinib is 50mg taken orally once daily.
Additional tests or investigations	[redacted]

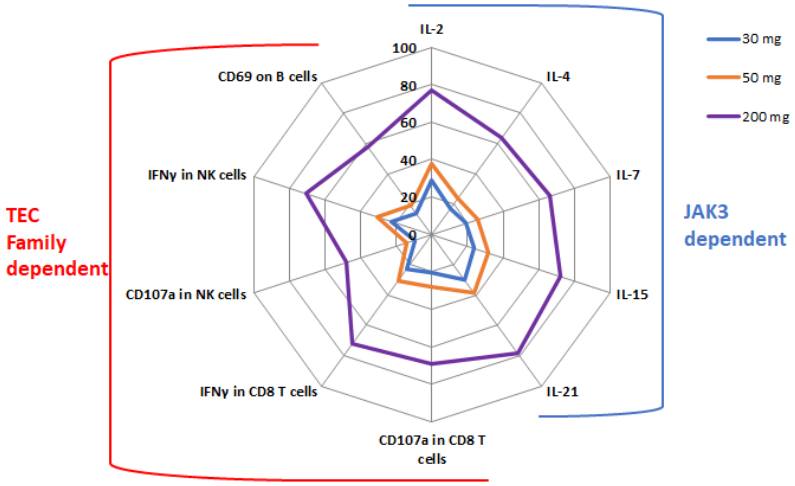
List price and average cost of a course of treatment	List price: £[redacted] per pack of 30 capsules Annual cost: £[redacted] per patient. [redacted]
Patient access scheme (if applicable)	PAS price: £[redacted] Annual cost: £[redacted] per patient.

Abbreviations: ALC, absolute lymphocyte count; ATP, adenosine triphosphate; BTK, Bruton's tyrosine kinase; BMX, bone marrow tyrosine kinase on chromosome X; CHMP, Committee for Medicinal Products for Human Use; EMA, European Medicines Agency; GB, Great Britain; IFN-γ, interferon gamma; ITK, interleukin-2–inducible T cell kinase; JAK, Janus kinase inhibitors; JAK1, Janus kinase 1; JAK2, Janus kinase 2; JAK3, Janus kinase 3; MAA, marketing authorisation application; MHRA, Medicines and Healthcare products Regulatory Agency; PAS, patient access scheme; SmPC, Summary of Product Characteristics; STAT, signal transducer and activator of transcription; TB, tuberculosis; TEC, tyrosine kinase expressed in hepatocellular carcinoma; TXK, tyrosine kinase expressed in T-cells; TYK2, Tyrosine Kinase 2

B.1.2.1. Ritlecitinib mechanism of action

Ritlecitinib is an orally bioavailable small molecule that inhibits JAK3 and the TEC kinase family (BTK, BMX, ITK, TXK and TEC) by irreversibly blocking the ATP binding site with very high selectivity over the other three JAK isoforms, JAK1, JAK2, and TYK2, as well as over the broader human kinome (Figure 1).[30,31, 33,36,37]

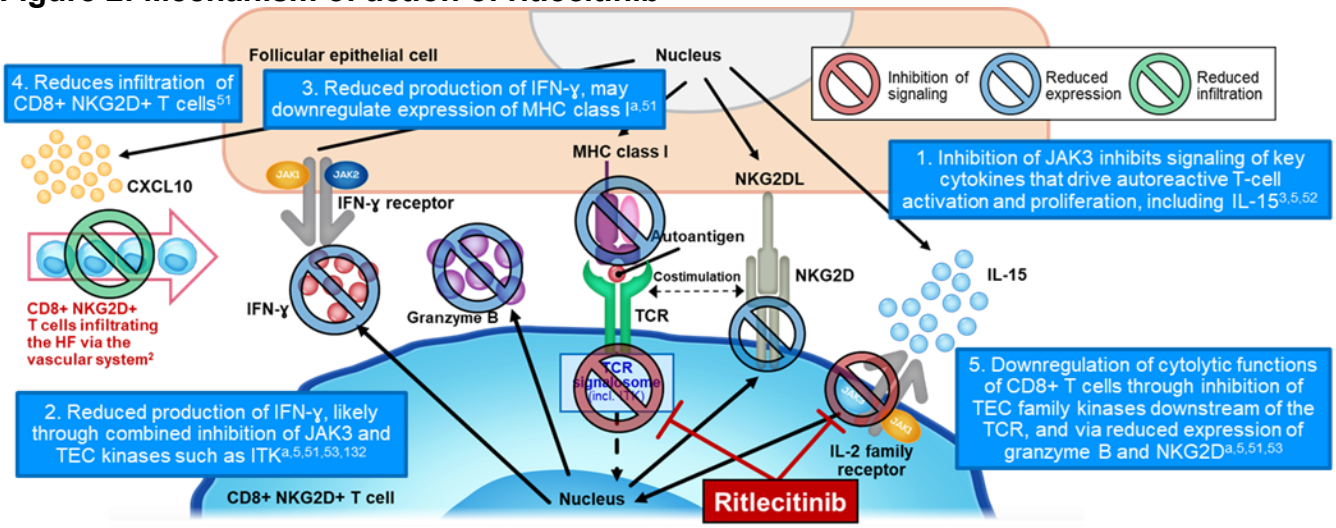
Figure 1: Average daily inhibition of JAK3 and TEC dependent cytokines by ritlecitinib following daily dosing of 200 mg, 50 mg or 30 mg



Abbreviations: CD, cluster of differentiation; IFN-γ, interferon gamma; IL, interleukin; JAK, Janus kinase; mg, milligram; NK, natural killer; TEC, tyrosine kinase expressed in hepatocellular carcinoma

As part of the pathophysiology of AA, T-cells infiltrate the epithelial layers of the hair follicle causing an IFN-γ response and upregulation of several common γ-chain cytokines, breaking down hair follicle immune privilege, and promoting the survival and activity of cytotoxic T-cells in the affected skin, this infiltration of cells is known as the 'swarm of bees'.[35,38] Ritlecitinib inhibits multiple signalling pathways involved in the pathogenesis of AA; such as the cytotoxic functions of CD8+ T cells and NK cells, and the production of IFN-γ by these cells (Figure 2).[39,40] Additionally, ritlecitinib inhibits signalling of the common γc cytokines IL-2 and IL-15, which have also been implicated in the pathogenic pathways of AA.[41–43]

Figure 2: Mechanism of action of ritlecitinib



[a] potential mechanism of inhibition of cytolytic activity by ritlecitinib in AA
Abbreviations: CD, cluster of differentiation; CXCL10, C-X-C Motif Chemokine Ligand 10; HF, hair follicle; IFN-γ, interferon gamma; IL, interleukin; incl. including; ITK, interleukin-2-inducible T-cell kinase; JAK, Janus kinase; MHC, major histocompatibility complex; NKG2D(L), natural killer group 2 member D (ligand); TCR, T-cell receptor; TEC, tyrosine kinase expressed in hepatocellular carcinoma
Source: Adapted from Telliez JB, et al (2016),[31] Xu H, et al (2019),[44] Dai Z, et al (2021),[45] Divito SJ, et al (2014),[46] Smith SEP, et al (2016), [47] and Suchonwanit P, et al (2021).[48]

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

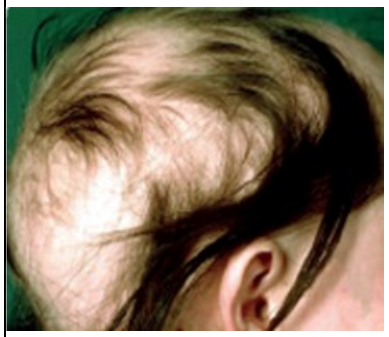
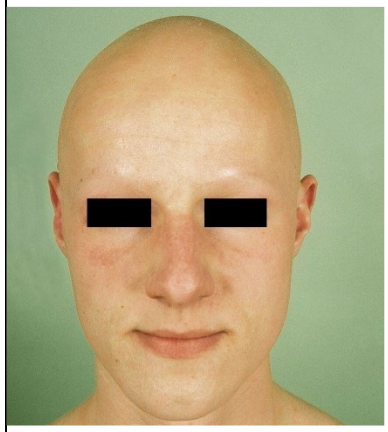
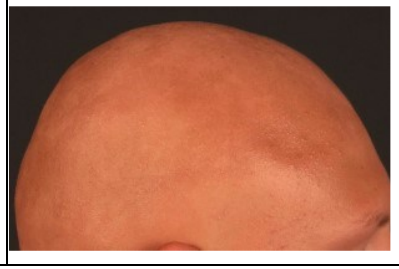
B.1.3.1. Alopecia areata overview

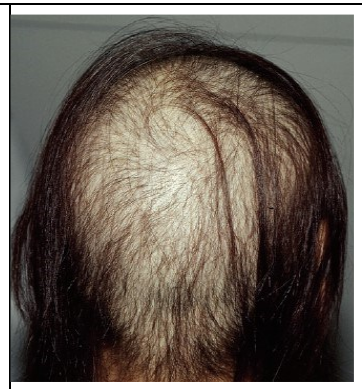
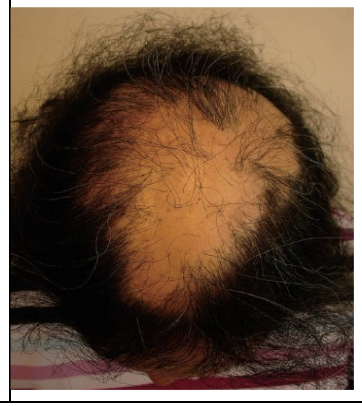
Alopecia areata (AA) is a chronic relapsing autoimmune mediated inflammatory disorder characterised by non-scarring hair loss and a variable, unpredictable, and relapsing or remitting course.[2] Physical examination, patient history and clinical course are often sufficient to make a clinical diagnosis of AA.[7,8]

The clinical presentation of AA can vary by the location of hair loss, pattern and severity.[1,49,50] Hair loss due to AA can range from well-defined patches to diffuse or total hair loss, commonly affecting the scalp but may involve any or all hair-bearing sites.[28]

The most common clinical presentation of AA is bald patches on the scalp.[51] Hair loss typically appears as alopecia focalis (one or more small, circular patches of hair loss). Other patterns of hair loss have been observed including ophiasis alopecia which is a band-like hair loss on the back and side of the scalp and sisaipho alopecia (where hair is retained along the sides and back of the scalp and is lost from the middle and the top of the scalp).[1,49] Complete or nearly complete scalp hair loss is termed alopecia totalis (AT) and complete or nearly complete hair loss on all surfaces of the body that have hair is termed alopecia universalis (AU). Alopecia incognita is also known as diffuse AA and is usually rapidly progressive. The different patterns of hair loss observed in patients with AA are presented in Table 4.

Table 4: Clinical presentations of AA

Types of AA	Description	Figure of clinical presentation	Occurrence
Patchy AA	Hair loss in a single, multiple separate or conjoined (reticular) patches		84.4%[28]
Alopecia universalis (AU)	Complete or nearly complete hair loss on all surfaces of the body that have hair		5.1%[52]
Alopecia totalis (AT)	Complete or nearly complete scalp hair loss		4.4%[52]

Alopecia incognita	Also known as diffuse AA due to diffuse total hair loss; usually rapidly progressive		3.2%[53]
Ophiasis	Hair loss in a band-like shape along the circumference of the head (border of the temporal and occipital bones)		2.7%[53]
Sisaipho	Extensive hair loss sparing the periphery of the scalp		0.2%[53]

Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis

Hair loss in other areas of the body is common, for example eyebrows, eyelashes, facial, nasal and on the full body. Loss of eyelashes, eyebrows and nasal hair are more common in patients with more severe AA.

Black dots and exclamation mark hairs (short broken hairs which are thinner and weakened at the point where they enter the scalp) may be seen around the margin of the hair, or in any part of the patch, during active disease.

The severity of AA can be measured with the Severity of Alopecia Tool (SALT) as detailed in the Alopecia Areata Investigational Guidelines.[54] The SALT score is computed by measuring the percentage of hair loss in each of four areas of the scalp — vertex (40%), right profile (18%), left profile (18%), and posterior (24%) — the total is the composite score.[54] Although calculation of the SALT score is not required in routine clinical practice for AA diagnosis, feedback from the Therapeutic Landscape Delphi panel suggested that the majority of clinicians in the UK use the SALT score to define severity of AA (7/7 clinicians, 100%).[21] It is also commonly used in clinical trials, thereby facilitating collaboration and comparison of data to measure treatment efficacy.

As shown in the diagram (Figure 3), the SALT score is calculated by:

1. Multiplying the percentage hair loss in each of the four quadrants of the scalp by the quadrant surface area
2. Adding the four values together for a composite score

Maximum SALT score = 100 (complete scalp hair loss)

Minimum SALT score = 0 (no scalp hair loss).[55]

This diagram (Figure 3) also allows the evaluator to graph the area(s) of alopecia, if desired, to facilitate the estimate of percent scalp hair loss and to compare the hair loss on subsequent evaluations. This percent hair loss can later be corroborated by image analysis if desired. Figure 4 presents an example of images which can be analysed to help corroborate the percent hair loss.[55] As shown in the examples below, a similar SALT score can have a varied appearance (Table 5).

Figure 3: Visual aid for estimating percentage scalp hair loss, percent growth and SALT score[55]

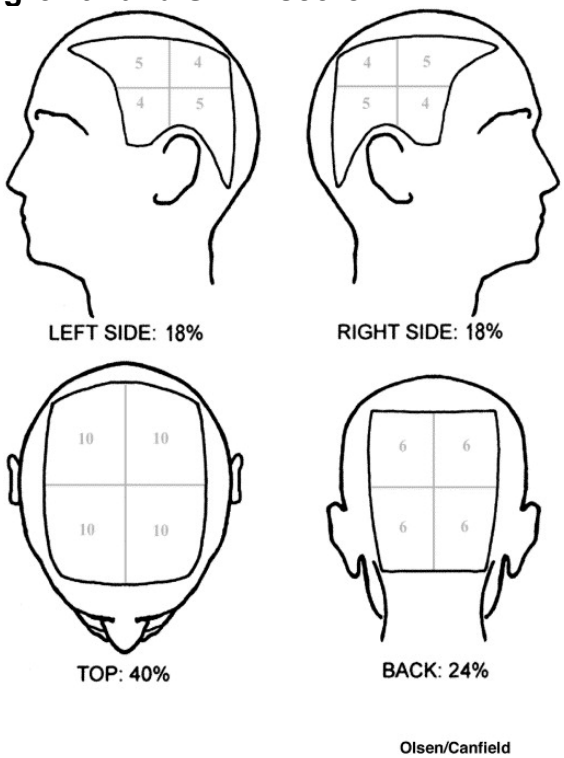
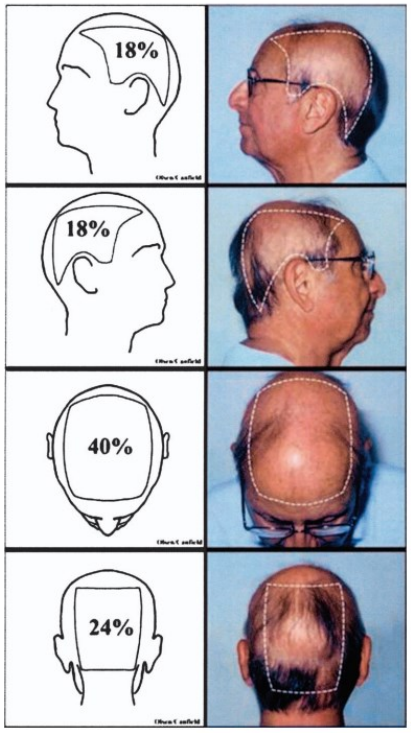
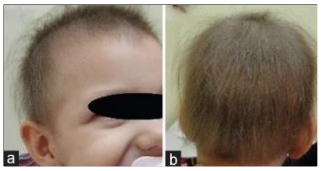
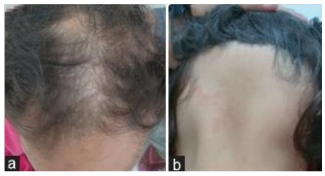
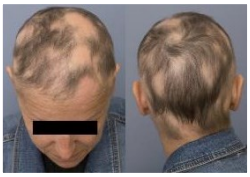
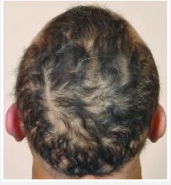
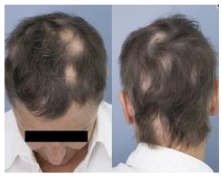


Figure 4: Example photographs taken of the four views[55]



The percentage of hair loss in any one of the four views (areas) of the scalp = the percentage hair loss x percent surface area of the scalp in that area. The SALT score then equals the sum of the scalp hair loss in each area. (a) Top (left side view) = 95% x 0.18 = 17.1; (b) Second (right side view) = 90% x 0.18 = 16.2; (c) Third (top of scalp) = 95% x 0.40 = 38; (d) Bottom (back of scalp) = 55% x 0.24 = 13.2; a+b+c+d = 17.1 + 38 + 16.2 + 13.2 = 84.5% hair loss or SALT 84.5

Table 5. Similar SALT Scores Have Varied Appearance in AA Patients

SALT 50%	SALT 55.2%	SALT 54.8%
		
SALT 10%	**SALT 13.2%**	
		

Abbreviations: AA, alopecia areata; SALT, Severity of Alopecia Tool.
Sources: Pfizer data on file (2022); Kiszewski AE, et al (2018)[56]; Grenier P, Veillette H (2017)[57]; Ibrahim O, et al (2017)[58].

B.1.3.1.1 Pathophysiology

The exact pathophysiology of AA is not completely understood, it is a complex disease determined by genetic and environmental factors and involving autoimmune and inflammatory responses. Immune cells are believed to play an integral role in the development of AA. It is thought that AA is caused by the infiltration of hair follicles by immune cells (known as the 'swarm of bees') that elicit an inflammatory response leading to hair loss.[1,2] Normal hair follicles are immune privileged, meaning that immune cells do not trigger an immune response in the hair follicle, whereas in AA this protective mechanism is disrupted and the immune cells attack the hair follicles causing the hair to fall out. Further information about the immune cells and their action in AA is provided below.

A specific subset of cytotoxic autoreactive CD8+ T-cells are believed to play an integral role in the development of AA.[39] This subset of T-cell infiltrates the epithelial layers of the hair follicle causing an IFN-γ response and upregulation of several proinflammatory cytokines, which can contribute to the collapse of the hair follicle immune privilege, facilitating an assault by immune cells, as well as promote the survival and function of cytotoxic T-cells in affected skin.[35,38] Inflammatory NK cells have also been shown to infiltrate the hair follicle and contribute to the pathogenesis of AA.[59] Interleukins (ILs) such as IL-2 and IL-15 promote the survival and the activation of the cytotoxic CD8+ T-cells.[41,42]

Proinflammatory cytokines stimulate inflammation via the JAK/STAT signalling pathway.[45] Overactivation of this pathway is known to contribute to a number of autoimmune diseases and proliferative disorders, such as AA.[41] JAKs phosphorylate and activate inflammatory cytokine receptors, leading to modulation of downstream targets (including the STAT transcription factors) and activation of immune-response-related activity. JAK1 and/or JAK3 inhibition will lead to modulation of γc cytokine pathways e.g. IL-15, which makes JAK inhibitors an attractive target for pharmacologic manipulation in the treatment of conditions such as AA.[31]

The immune attack on the hair follicle by $CD8^+$ T cells is a hallmark of the disease thought to be mediated by recognition of autoantigens by the T-cell receptor (TCR). Signalling downstream of the TCR involves multiple mediators including members of

the TEC family (tyrosine kinase expressed in hepatocellular carcinoma).[60] In activated T cells from the scalp of patients with AA, a specific TCR signalling signature has been shown, involving multiple mediators including ITK.[47]

B.1.3.1.2 Prognosis

The natural history of AA is not well known. Spontaneous remission varies with the severity of disease.[28] In a long term follow-up study of 191 patients, spontaneous remission ranged from 68% in patients with AA with <25% hair loss to approximately 8% in patients with severe AA while one review found 14-25% of cases of AA may progress to AT or AU, from which a full recovery is less than 10%.[18,28]

This finding is supported by the Therapeutic Landscape Delphi panel in which 75% of dermatologists with a specialist interest in hair disorders said that few patients (<20%) with severe hair loss (including AT/AU) would experience spontaneous remission without treatment (6/8 clinicians, 75%) and half of the dermatologists with a specialist interest in hair disorders stated that <10% of these patients would experience spontaneous remission (4/8 clinicians, 50%).[21] Regrowth potential is lifelong but typically occurs after a minimum of 3 months and up to a few years after initial onset.[28] Almost all patients will experience more than one episode of the disease;[61,62] with estimates suggesting that 85% of patients experience relapse and when the observation period exceeded 20 years, 100% of patients showed relapse.[61] Psychological stress and/or stressful lifetime events can exacerbate or trigger AA in patients.[22,63–65] In a study of patients with AA, over 20% of patients reported more stress before the onset and/or exacerbation of symptoms.[66]

The strongest predictors of long-term outcome of treatment are location and severity of hair loss, age at presentation, and duration of disease.[7,18,19] Poor prognostic factors are:

- Family history of AA[8]
- More severe disease at onset[7,18,19]
- Disease duration ≥1 year and/or a long episode of hair loss prior to treatment[67]
- Younger age at onset (particularly before the age of six)[1]

- Ophiasis[1]
- Nail involvement[8]
- A history of atopy or other autoimmune diseases[8]

Dermatologists with a specialist interest in hair disorders have suggested that treatment at an early stage could influence long-term outcomes.[51] In accordance with this opinion, a study found that patients with an episode of AT or AU of longer than 10 years in duration were less likely to respond to treatment.[52]

B.1.3.1.3 Epidemiology

AA is one of the most prevalent autoimmune disorders and the second most prevalent hair loss disorder after androgenic alopecia (male/female pattern baldness).[3] It affects around 2% of the general population at any point during their lifetime and, based on a meta-analysis of 93 studies (published before September 2018), has an estimated global population-based prevalence of approximately 0.75% (N=301,173,403).[28,68] A population-based cohort study conducted using UK electronic primary care records from the Oxford-Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) network database between 2009-2018 (N=2,634,083) reported an overall diagnosed point prevalence of 0.58% among adults with AA.[4] The estimate is based on combined data from both clinic- and population-based studies and is in line with the global population-based prevalence of AA for Europe (0.58%).[68]

Based on combined data from both clinic- and population-based studies, the global prevalence of AA has been reported to have increased three-fold over time when comparing studies published before 2000 and studies published after 2009 ($P < 0.001$).[68] This increase may simply be due to detection bias from improved disease awareness, increased self-consciousness in the age of social media meaning more patients seek treatment for AA or improvements in study methodology over time. Alternatively, this finding may be due to environmental factors such as dietary deficiencies (vitamin D, zinc, folate) and psychological distress.[68]

There is uncertainty around the proportion of patients with severe AA in the UK, however Pfizer has conducted a web-based survey in the United States (US; N=45,016) to estimate the point prevalence of severe AA, and AT or AU combined.[5]

The point prevalence of clinician-adjudicated severe AA and AT/AU was 9 per 10,000 and 4 per 10,000, respectively.[5] Based on this study and Pfizer's internal analysis, AT/AU appears to represent a subpopulation of approximately 3–5% of the total population of patients with AA.[69] At first presentation, approximately 42% of patients with AA present with severe AA.[3]

AA affects people of all ages, races, and sex; most studies report no significant differences in the age of onset between those of different sex or ethnic groups;[28] however, in the meta-analysis of 94 clinic-based and population-based studies the prevalence of AA has been shown to be higher in studies of children/adolescents than in studies of the adult population ($P < 0.001$).[68] The reported prevalence of AA in children in the study was 1.92%.[68] During discussions with clinical experts, however, they stated that they do not believe the prevalence of AA varies between adolescents and adults.[17] AA incidence has been shown to peak between 25–29 years of age in the UK, with the median age at diagnosis being 31 for males and 34 for females.[4]

People with particular demographic characteristics are more likely to be diagnosed with AA. In a population-based cohort study in the UK, incidence of AA in females was 19% higher than in males when adjusted for age and other sociodemographic characteristics ($P < 0.05$).[4] AA was more likely in non-white ethnicity groups than white ethnicity groups ($P < 0.05$), with Asian ethnicity groups 232% more likely to develop AA than white ethnicity groups ($P < 0.05$).[4] Additionally, incidence in those with the highest levels of social deprivation was 47% higher than the least deprived quintile ($P < 0.05$), and the incidence among urban was 23% higher than rural dwellers ($P < 0.05$).[4]

There are a number of risk factors which have been identified to increase the risk of AA. These risk factors are as follows:[70]

- Chromosomal disorders such as Down Syndrome
- Polyglandular autoimmune syndrome type 1
- Other autoimmune conditions such as vitiligo and thyroid disease
- A family history of AA

- Certain susceptibility genes

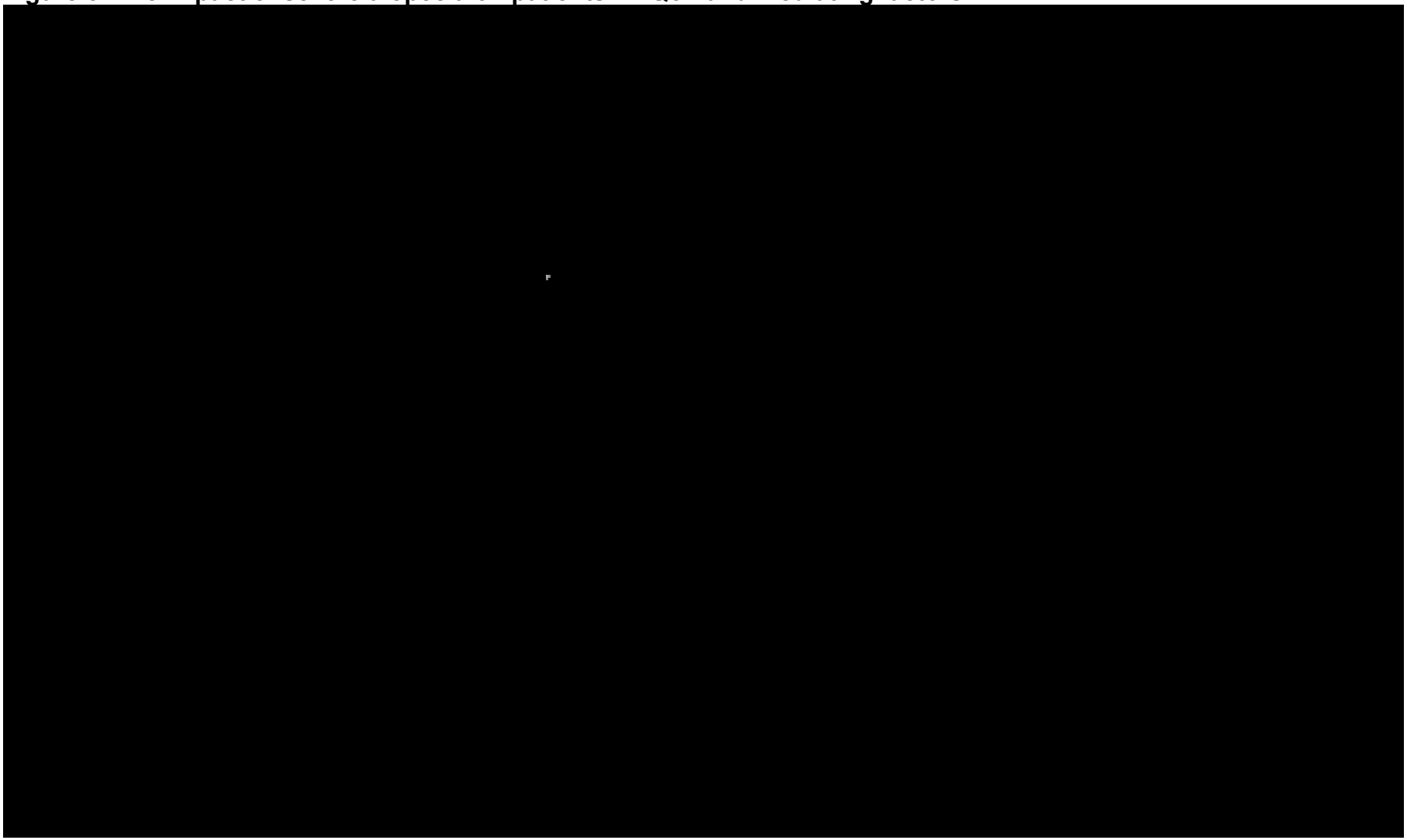
B.1.3.2. Burden to patients, carers, and society

Literature was identified to characterise the burden of AA, however none of the publications focused on patients with ≥50% hair loss from a UK perspective. Therefore, to contextualise the burden of severe AA to patients in the UK, we discussed concepts from published literature and confirmed the relevance via qualitative interviews with representatives from the UK AA community. Ten representatives from six Patient Advocacy Groups (PAG) focused exclusively on patients with AA with ≥50% hair loss (i.e., severe AA) living within the UK confirmed the relevance of the burden to patients in the UK (Figure 5).[22] The burden of severe AA to patients is categorised into two themes: health-related quality of life (HRQoL) and mediating factors. Each category was then divided into sub-themes to explicate the different facets of the disease experience for UK patients with severe AA.

Quality of life can be split into two overarching themes: physical and psychosocial impacts of AA which are further described in sections B.1.3.2.1 and B.1.3.2.2 respectively. Psychosocial impacts are further split into emotional and functional impacts. Mediating factors, include disease management and coping mechanisms which are considered further in Section B.1.3.3.

Throughout Section B.1.3.2, the burden of AA is described with a specific emphasis on patients with severe AA given the licensed indication of ritlecitinib.

Figure 5: The impact of severe alopecia on patients HRQoL and mediating factors



B.1.3.2.1 Physical impacts of AA

Hair loss and associated symptoms

As described in Section B.1.3.1, the clinical presentation of AA can vary by the severity, location and pattern of hair loss. In adult and adolescent patients, scalp hair loss is described as the most burdensome symptom of AA, followed by loss of eyebrow and eyelash hair.[71,72]

It is reported in the literature that severe AA causes dry and flaky skin, itchiness, tingling, burning, or pain with disease activity in the affected areas.[20] Hair regrowth can initially present as patchy and fine white hair and trigger other skin problems such as an itchy scalp.[11,28] Patchy hair regrowth is, however, generally considered less burdensome than hair loss.[73] Among adolescents, scalp itchiness, patchy hair regrowth and headaches due to sun exposure are regarded as the most bothersome symptoms after scalp hair loss.[73] Sunburn was also found to particularly affect individuals with more severe hair loss and individuals with AT/AU.[6,71]

The loss of eyebrow and eyelash hair lead to practical impairments, including sweat and debris in the eye. Whereas, loss of nasal hair can in turn lead to frequent sneezing and runny nose.[6, 72,74,75]

Nail changes including pitting (shallow or deep holes), splitting, brittleness, or trachyonychia (rough accentuated linear ridges) appear in 7–66% of patients with AA, and are more common in patients with severe AA, AT and AU.[76] When pain is experienced it is often related to areas of hair loss (i.e. the skin and scalp) but can also be due to nail weakness.[11,77,78]

Following severe scalp hair loss, symptoms such as itching, burning, pain, irritation and bleeding are associated with more severe symptomatic burden according to the Dermatology Life Quality Index (DLQI; a self-administered questionnaire designed to measure the HRQoL of adult patients suffering from a skin disease)[79] and Skindex-29 (a measure of the effects of skin diseases on patients' HRQoL).[80,81]Indeed, patients with AU had the worst scores on the symptom domain of the Skindex-29 and symptoms and feelings domains of the DLQI compared to all other levels of hair loss. A similar trend was observed by level of body hair loss.

The physical symptoms of severe alopecia, secondary to hair loss, as summarised in Figure 5, were also noted by PAG representatives and the most commonly cited physical symptoms were pain, itching or tenderness of the scalp, eye irritation, a runny or streaming nose, sensitivity to temperature, nail damage, susceptibility to illness or infection, sunburn, and impaired hearing.[22]

Scalp sensitivity, one of the most common physical symptoms of severe AA, was noted to have often occurred prior to or during hair loss.[22] It was also suggested that scalp sensitivity can be related to the use of wigs, which are more likely to be used by people with greater hair loss. Eye irritation was associated with a lack of protection because of loss of eyelashes, or irritation from sweat because of a loss of eyebrows. A runny nose was linked to the lack of nasal hair and sensitivity to temperature, related to both hot and cold temperatures. In patients with AT/AU, the loss of eyebrows and eyelashes was reported as being particularly impactful by all PAG representatives asked.[22] For example, one PAG representative said:

[redacted]

Comorbidities

A high prevalence of comorbidities in patients with AA is reported in the literature. Disease burden in AA is frequently compounded by comorbidities including asthma, allergic rhinitis, atopic dermatitis (AD), and autoimmune diseases such as thyroiditis and vitiligo.[63] A retrospective cohort study in the UK found that patients diagnosed with AA had a higher prevalence of atopic (37.2% vs 26.7%; p≤0.001) and autoimmune conditions (11.5% vs 7.9%; p≤0.001) compared with controls.[82] Similarly, a global meta-analysis of 87 studies indicated that patients with AA were more likely to have AD, allergic rhinitis and anaemia than controls without AA (Table 6).[83]

Table 6: Meta-analysed OR and prevalence of comorbidities in patients with AA

Comorbidity	Prevalence			OR			
	Studies, N	AA, N	Mean prevalence, % (95% CI)	Studies, N	AA, N	Control, N	OR (95% CI)
Atopic dermatitis	16	11,423	9.6 (6.2, 14.4)	5	7,191	872,218	2.36 (1.80, 3.09)

Allergic rhinitis	8	8,283	17.7 (14.1, 21.9)	5	7,081	785,008	1.33 (1.19, 1.47)
Iron deficiency anaemia	2	655	7.5 (5.7, 9.8)	2	655	243	2.78 (1.23, 6.29)
Lupus erythematosus	3	5,143	0.8 (0.3, 2.5)	1	4,334	784,158	4.73 (3.70, 6.10)
Thyroid disease	23	15,188	8.0 (5.9, 10.7)	6	5,265	784,684	1.99 (1.29, 3.06)
Psychiatric disorder (any or unspecified disease subtype)	6	5,941	49.4 (17.8, 81.5)	1	5,117	20,468	1.35 (1.20, 1.52)
Anxiety	13	10,257	27.1 (17.7, 39.2)	7	5,945	20,878	2.00 (1.51, 2.65)
Obsessive–compulsive disorder	2	5,148	1.0 (0.2, 5.8)	1	5,117	20,468	1.72 (1.06, 2.77)
Depression	14	10,331	18.9 (10.9, 30.8)	8	5,976	20,918	2.27 (1.52, 3.37)

The OR of each comorbidity presented was significant between AA and control groups (general population without AA)
Abbreviations: AA, alopecia areata; CI, confidence interval; N, number of studies/participants included in group analyses; OR, odds ratio
Source: Lee et al, (2019)[83]

Psychiatric disorders are also more common in patients with AA compared to the general population (Table 6),[84] as further described in Section B.1.3.2.2.[83]

The findings in literature were supported by the PAG representatives who noted that people with severe alopecia tend to have a variety of comorbidities, including other autoimmune conditions and dermatological conditions.[22] PAG representatives offered estimates that the prevalence of comorbid conditions was as high as 90%.[22]

B.1.3.2.2 Psychological impacts of AA

In a cross-sectional study (N=216), coping with AA is reported to be a daily challenge for 85% of patients: issues faced by patients included worries that others will not find them attractive (78%), feeling anxious in social situations (69%), low self-esteem (71%) and being treated differently by others (58%).[6]

The extent of psychological impacts were also explored in the PAG study. Participants were asked to rank the three most bothersome symptoms or impacts they felt people with AA experienced;[22] overall, those ranked first and second were always emotional and psychological impacts (Figure 6). A summary of findings for PAG representatives with experience with adults (n=9) can be found in Figure 6. Those impacts ranked as the most bothersome were anxiety [redacted], impact upon appearance [redacted], loss of identity [redacted], lack of confidence [redacted], uncertainty/unpredictability [redacted], and

distress [redacted]. This suggests that PAG representatives conceptualise the burden associated with severe alopecia as being primarily emotional and psychological, rather than physical.

Figure 6: PAG representative reports of the three most bothersome symptoms and impacts of severe alopecia in adults

Abbreviations: PAG, patient advocacy group

Emotional impacts

It has been reported that AA can have a significant psychological impact on individuals, which can lead to an increased risk of anxiety, reduced self-esteem, altered body image, social withdrawal and the breakdown of personal relationships.[3,7–12] Depression and anxiety have been closely linked to AA; a meta-analysis of patients with AA (N=6,010) and controls (N=20,961) demonstrated a positive association between AA and anxiety ($P < 0.05$) or depression ($P < 0.05$).[85] While a high proportion of patients are diagnosed with a psychiatric disorder prior to their AA diagnosis,[65,86] AA can lead to the development of new onset mental health conditions and exacerbate existing ones.[16]

The lifetime prevalence of psychiatric disorders such as depression, anxiety, obsessive-compulsive disorder and alexithymia (the inability to recognise or describe

one's own emotions) in patients with AA has been estimated at 66–74%,[83] with depression and anxiety having a lifetime prevalence of 38–56% and 39–62%, respectively.[3,87–90] Additionally, relapse of hair loss increases the frequency of negative psychological traits such as paranoia and obsessive-compulsive tendencies.[65]

In extremis, the sense of hopelessness patients with AA feel result in suicidal thoughts (Figure 5). In a study conducted in the UK including patients with AA (N=45), 9.1% of patients with AA often wished that they were dead.[91] In another study of patients with AA (N=64), 12.8% of patients were at risk of committing suicide.[92] In a further study of patients with various dermatological conditions (N=300), a quarter of patients with AA reported having suicidal thoughts, making suicidal ideation significantly more prevalent in patients with AA than in patients with other dermatological conditions, including psoriasis, vitiligo and acne (P = 0.001).[13] The study also found that the risk of suicide increases with the severity of hair loss.[13]

There is conflicting evidence on the extent to which the burden of AA varies between adolescents and adults with AA,[13, 16, 66,92–95] however some evidence suggests the psychological burden of AA is particularly evident in adolescent patients. A reported case series of the suicides of four adolescent males within one year of AA diagnosis, all of whom had no preceding psychological disorders.[95] In addition, a US study found that young adolescents with AA (n=33) were significantly more anxious, depressed, withdrawn, aggressive or delinquent (P < 0.01), compared to young adolescents without AA (n=30).[94]

The themes from the literature were supported by qualitative interviews with PAG representatives. The most frequently cited psychological impact of severe AA was anxiety.[22] Anxiety was described in two forms: "social anxiety" and "uncertainty". Social anxiety reflected patients' concern about how they would be viewed by others. The emotional toll of living with an unpredictable condition (i.e., "uncertainty anxiety") was also felt to be burdensome for many patients: particularly the lack of control around continued progression of hair loss and uncertainty over whether and when hair would regrow. Participants also described depression or low mood as a typical experience of the patients they met in their advocacy group. Closely linked to depression,

participants also noted feelings of distress, devastation or heartbreak. These reports were often described in terms of “grief” or “bereavement”.

Another important impact identified by participants was a loss of identity associated with hair loss, whereby patients no longer feel like or recognise themselves.[22] For example, one participant said:

[redacted]

This loss of identity was noted to be particularly impactful for those with AU. The loss of eyebrows and eyelashes was felt to exacerbate the loss of identity associated with severe alopecia, having the effect of “wiping out” an individual’s appearance and were felt to be more difficult to conceal than scalp hair. In addition, PAG representatives reported a loss of femininity or masculinity associated with severe alopecia. For women, this was usually related to the loss of scalp hair, whereas for men this was associated with the loss of body or facial hair.

PAG representatives also described that patients with severe alopecia may have feelings of guilt associated with alopecia (due to mistaken perception by others of being a cancer patient); patients who were not open about their condition felt guilty about being disingenuous when discussing hair with others.[22] Feelings of guilt were also associated with the fact that patients with severe alopecia were physically well.

Functional impacts

The emotional impacts of AA leads to patients limiting or functioning differently as a consequence. It has been reported that people with AA who are of working age (18–65 years) are 56% more likely to be issued with certificates for time off work and 82% more likely to be recorded as unemployed than the general population, contributing to the economic burden of AA described in Section B.1.3.2.5.[16] Patients take time off work due to social withdrawal or anxiety related to their alopecia or due to general worries of returning to a place of work following an absence (such as COVID-related lockdowns).[22] This was reflected by PAG representatives, who raised specific concerns with attending the workplace such as the reaction of others, including gossip

or feeling pressure to cover up hair loss and physical discomforts associated with the work environment, such as wearing a synthetic wig when working in a hot kitchen.[22]

The emotional implications of AA also lead to increased absenteeism and the loss of education in adolescents, which will have lifelong implications; of patients attending school in a cross-sectional study (N=47), 51% reported missing time from school because of their AA.[6] Similarly, PAG representatives reported impacts upon school or college for adolescents with severe alopecia.[22] Participants reported a reluctance to attend school or college and the need for time off for some adolescents with severe alopecia.[22] Bullying was mentioned as an impactful issue on education by PAG representatives. Participants noted a lack of support or leniency from schools around the need to wear a head covering or make-up; adolescents' requests for taking time off school were frequently rejected. One participant said:

[redacted]

PAG representatives reported that stigma and bullying, which could take the form of staring or nasty comments, including via social media, physical harassment, the misgendering of women, or being mistaken for a cancer patient, impacts social activities for patients with severe alopecia.[22] Additionally, a US study (N=69) found that 13% of patients with AA aged 12–14 and 40% aged 15–19 had experienced bullying associated with their AA.[27]

The social impact of AA was found to be comparable to other skin disorders such as psoriasis, AD and chronic spontaneous urticaria in a study of patients with AA in France (n=60).[96] According to the Global Burden of Disease study (2010), the disability weighting of AA (0.035) was similar to other skin conditions in which there are available licensed therapies, including eczema (0.038), cellulitis (0.035) and urticaria (0.031).[97]

In addition to the topics raised in literature, PAG representatives noted broader functional impacts to patients with AA. For example, alopecia can negatively impact

upon the relationships people have with their friends, families or partners.[22] Although partners were generally described as supportive, it was reported that patients found it difficult to accept that their partner could still view them in the same way as they had prior to the onset of alopecia. Similarly, challenges associated with dating were mentioned by participants, including struggles patients had with deciding if and when to inform dating partners of their condition and past negative experiences upon doing so. In addition, PAG representatives noted a more general burden was placed upon patients to explain, justify or "come out" with alopecia.

The PAG representatives also reported significant impacts upon the social lives of the patients with whom they interacted.[22] For example, one participant said:

[redacted]

Participants reported a persistent fear of scalp hair loss being "exposed" in public should a wig or head-covering fall off. As such, participants mentioned that patients with severe alopecia tend to avoid social events and to withdraw from public-facing life.[22] Even when patients did engage in social activities, their behaviour was heavily influenced by considerations around concealing their hair loss. For example, patients are more likely to avoid specific activities that risked "exposure" of their alopecia such as going swimming, to the gym or to theme parks.[22] In addition, PAG representatives mentioned that getting ready for the day involved a lot of effort and time for people with severe alopecia, such as skin care and make up routines.[22] This was noted to be particularly true of those with AU who use fake eyebrows and eyelashes:

[redacted]

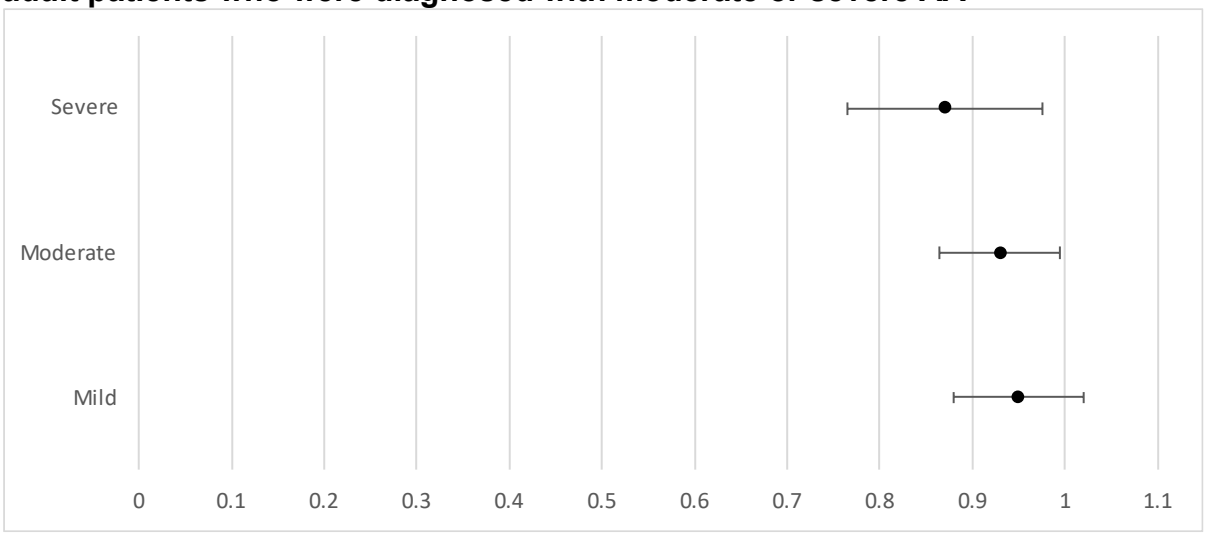
Moreover, it was noted by participants that activities of self-care (such as washing and brushing hair) may be avoided out of concern for exacerbating hair loss.

B.1.3.2.3 HRQoL burden

AA is a complex heterogenous condition that impacts across physical, functional and emotional domains of HRQoL. The impact of AA on HRQoL is particularly evident in adolescent patients. A US study (N=69) found that 67% patients with AA aged 12-14 and 75% aged 15-19 reported an impact on HRQoL.[93] The extent to which the heterogeneity impacts on the measure of HRQoL across the full population of patients with AA is unclear. The question remains how this potential source of noise impacts on the interpretation of results in the literature and how this aligns with individual patient and clinical experiences.

A real-world study from the US was conducted to understand how the severity of AA affects patients' HRQoL.[98] The sample consisted of adult patients who were diagnosed with moderate or severe AA or with a history of moderate/severe AA. The study reported EQ-5D-5L utility values by physician-rated AA severity categories of mild (n=56), moderate (n=140) and severe (n=65). Severity was rated subjectively by the physician based on the patient's medical history and an assessment during the consultation. Values of 0.95 (± 0.14), 0.93 (± 0.13) and 0.87 (± 0.21) were observed for patients with mild, moderate, and severe AA, respectively, as presented in Figure 7.

Figure 7: EQ-5D-5L utility values by physician-rated AA severity categories of adult patients who were diagnosed with moderate or severe AA



Abbreviations: AA, alopecia areata

The degree of difference between each group was small and values were very close to US population norms.[99] This is inconsistent with the broader literature that suggests HRQoL impact in patients with AA increases with more severe hair loss and suggests that the EQ-5D may not capture the full impact of AA.[3, 28,100,101]

Another study conducted in Iran explored the impacts of AA on SF-36 scores (n=100) and showed no significant differences in severity of hair loss (categories: <25%, 26-50%, 51-75%, 76-100%).[102] One other study using the SF-36 did not observe an association between severity of hair loss and HRQoL and only reported SF-36 domain scores for the whole sample but not by severity of hair loss.[80] This again suggests that the extent to which generic measures of HRQoL capture the full burden of AA may not align with the impact in patients with AA.

The question then remains how well disease relevant patient reported outcome measures (PRO) better capture the HRQoL burden for patients with AA; a detailed discussion on the strengths and limitations of HRQoL instruments is discussed in Section B.3.4.3. It is thought that patients with severe AA experience more severe psychological effects and poorer HRQoL than those with mild to moderate AA.[64] A review from the European Academy of Dermatology and Venerology found that across different HRQoL instruments (AAQ and DLQI) individuals with severe AA and AT/AU had poorer HRQoL than those with less severe forms of AA.[85,103] In addition, the majority of studies using the DLQI showed that more severe hair loss was related to poorer DLQI scores across all HRQoL sub-domains.[75, 80, 102,104–107] Data from alopecia-specific instruments (Alopecia Areata Patient Priority Outcomes [AAPPO], AA Quality of Life [AAQ], AA Quality of Life Index [AA-QLI]), which capture more relevant domains for patients with AA, as well as DLQI, which has limitations (but is potentially more favourable compared to EQ-5D and SF-36), suggests that the psychological burden that AA has on patients affects their HRQoL.[96, 101,108]

B.1.3.2.4 Caregiver burden

The psychosocial impact of caregiving is a key issue in the field of AA; one cohort study of 229 family members of patients with AA found that 69.9% of family members of adults with AA (aged 17+) experienced some HRQoL impairment and 14%

experienced a very large or extremely large effect (measured using the Family Dermatology Life Quality Index [FDLQI]). For children with AA (aged 4-16) the impact was more widespread with 87.2% of family members experiencing some HRQoL impairment and 31.4% a very large or extremely large effect.[14]

Based on a systematic review of 11 English-language studies measuring caregiver burden using the Paediatric Quality of Life Family Impact Module (PEDSQoL) caregivers of patients with AA have scored lower than controls on psychosocial and physical health domains;[100] furthermore, AA is found to be worse than other chronic skin conditions (such as atopic dermatitis, psoriasis and urticaria) in terms of caregiver impairment as measured using the FDLQI and the Impact on Family scale (IOF).[109]

There is evidence that the psychosocial strain on caregivers increases with severity of disease.[110] In a prospective study conducted in the US of 153 paediatric patients with AA, significant mild-to-moderate negative correlations were found between SALT scores and both FDLQI and Quality of Life in a Child's Chronic Disease Questionnaire (QLCCDQ) scores.

Evidence of the impact to caregivers from the literature was supported by the PAG representatives; of the PAG representatives consulted who had experience with caregivers of adolescents with severe alopecia, [redacted] reported impacts upon the caregivers themselves.[22] Participants mentioned that the type of support that adolescents and adults needed from PAGs could be different.[22] Caregivers of adolescent patients expressed feeling helpless when confronted with an illness with no effective treatment; they wanted to help their child but did not know how to do so. For example, one participant said:

[redacted]

PAG representatives also reported that caregivers worried about the future of their children. Participants noted that parents face challenges in getting their child to attend school.

Other less common impacts specifically reported by PAG representatives included frustration and upset or distress.[22] Uncertainty associated with the unpredictable and uncontrollable hair loss and a strain on family dynamics were also reported.

B.1.3.2.5 Economic burden

AA is a chronic condition that requires long-term treatment and as such is a major economic burden to society. Direct costs to the NHS include outpatient visits, prescriptions costs, therapy and headwear such as wigs.[111]

The burden on AA extends to individuals; although not thoroughly understood, the economic burden of AA has been the subject of recent research. Most patients (57%) rated their financial burden as moderately or severely burdensome in a patient survey.[112] Market research indicates that the current spending of patients with AA to treat AA is between £50 - £150 per month.[15] Direct costs to patients include out-of-pocket (OOP) costs such as headwear or cosmetic treatment (e.g. scarves, hats, wigs, makeup, artificial eyebrows or eyelashes), hair appointments, travel to appointments, vitamins/supplements and medications purchased over the counter.[98,111,112] One dermatologist with a specialist interest in hair disorders also highlighted that due to the lack of treatment options available and unmet need, patients with AA are frequently purchasing JAK inhibitors from overseas against the advice of their prescribing physician.[17] One PAG representative noted that more severe hair loss is associated with greater financial expense because of the need for wigs, fake eyebrows and eyelashes.[22]

Indirect costs of AA include absenteeism, presenteeism and overall work and activity impairment (as described in Section B.1.3.2.2).[98,111] A real world study conducted in the US found that the overall work and activity impairment due to AA increased with increasing severity.[98]

B.1.3.3. Clinical pathway of care

Few therapies have been assessed for the treatment of AA in randomised controlled trial (RCT) studies.[7] There is no established standard of care for severe AA and no licensed therapies available within the UK.

There is no therapy which reliably produces long-term remission of disease off drug nor is there a therapy convincingly demonstrated to have the efficacy and safety appropriate for long-term management required to maintain hair regrowth.[47, 49,50,113,114] Treatment can be uncomfortable, time consuming and potentially toxic for patients and relapse, either during or even after initially successful treatment, is common and some patients find it difficult to cope with this experience.[115,116] Patients may consider the side effects from these treatments combined with their unpredictable results unacceptable,[117] and so opt for the non-pharmacological therapy option.[118] In the absence of any licensed therapies, there is a lack of consensus in expert opinion regarding the optimal management of AA. A recent consensus study involving 50 international hair and scalp experts found that a consensus of only 33% was achieved for treatment specific questions and concluded that there is a need for robust research in AA therapeutics to support clinical decision making.[119]

B.1.3.3.1 Treatment pathway

There is currently no up to date pathway of care for patients with AA in the UK. The most recent guidelines for the treatment of AA, published by the British Association of Dermatologists (BAD) in 2012, provide recommendations according to disease severity (for limited or severe hair loss and AT/AU).[7]

The most frequently recommended treatments are corticosteroids: topical, intralesional and systemic therapy. Contact immunotherapy, is also recommended for some individuals with chronic AA; and it is recommended that patients are given the choice to remain untreated and to consider a wig or hair piece.[7] Given the limited evidence base, all the BAD recommendations are to a strength of C (where A and D are the strongest and weakest recommendations, respectively). Furthermore, there is very limited information about how treatments might be sequenced.

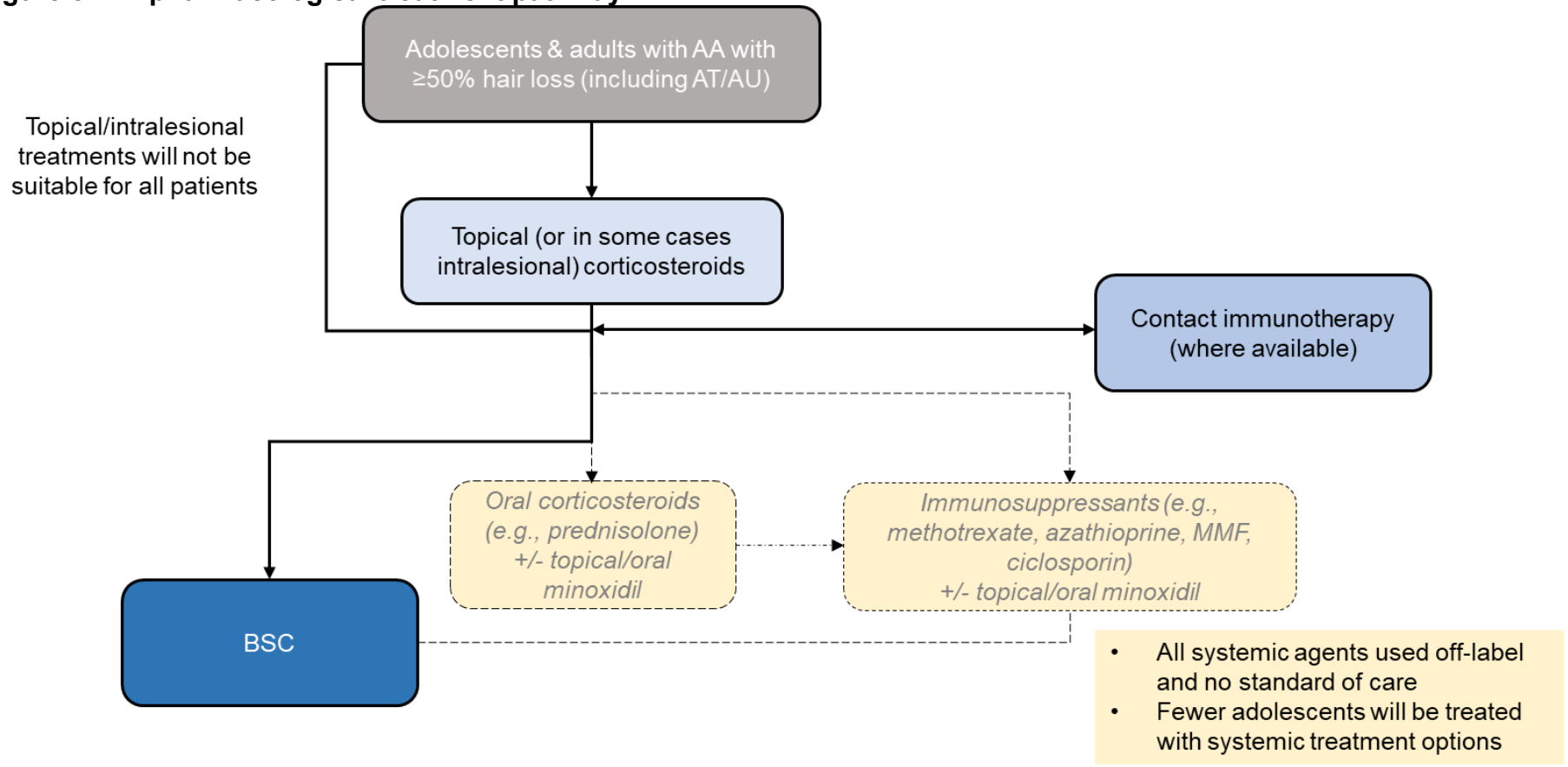
The BAD guidelines were developed over ten years ago in the absence of any approved treatments which specifically target the underlying cause of AA. Since then, there has been very limited advances in treatment for patients with AA. Given the limited innovation and progress in this area, the BAD guideline review which was scheduled initially for 2017 is understood to be postponed ensuring that recent clinical

trials and Marketing Authorisation Applications for novel treatments in AA are reflected.

PAG representatives report poor experiences with healthcare services and practitioners, which may be a reflection of a lack of new or up to date guidance. .[22] PAG representatives described typically negative experiences of obtaining a diagnosis of alopecia.[22] Feelings of frustration due to the lack of information and understanding from GPs and lack of treatment options were frequently described. GPs were felt to lack sufficient knowledge and often struggled to differentiate between different subtypes of alopecia, used incorrect terminology, or they even misdiagnosed the condition. PAG representatives also reported that members often mentioned that they had received minimal or no psychological support from their GPs, and that they often felt dismissed or "brushed off." Participants also noted dissatisfaction with the limited opportunities for referral to dermatology specialists, or long waiting times, if referred. All of these factors contribute to the emotional burden patients face, as described in Section B.1.3.2.2 and contrasts with the commitment to improve mental health services available to patients with skin conditions.[120]

A pharmacological treatment pathway for adults and adolescents with severe AA (including AT/AU) was developed through discussion and iteration with clinical experts from the UK in the Therapeutic Landscape Delphi panel;[21] the final pathway is shown in Figure 8.

Figure 8. AA pharmacological treatment pathway*



Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; BSC, best supportive care; MMF, mycophenolate mofetil
**75% of dermatologists with a specialist interest in hair disorders agreed with an earlier version of the diagram and then following three further interviews the diagram was optimised and finalised.*

Insights from the clinician validation are that typically dermatologists outside of a specialist centre would not prescribe systemic treatment and patients would go straight to being treated with best supportive care (BSC).[17] In addition, dermatologists suggested that there is no consistent choice of systemic treatments for patients with severe AA. Only dermatologists with a specialist interest in hair disorders would consider prescribing systemic treatment, but this is off-licence and not consistently available for all patients with severe AA.

When off-licence treatments are used, it was noted by dermatologists with a specialist interest in hair disorders that the treatment options vary by patient and are particularly dependent on age (i.e., adolescent or adult).[21] In the Therapeutic Landscape Delphi panel, dermatologists with a specialist interest in hair disorders agreed that treatment options used for adolescents would differ to those used for adults (7/8 clinicians, 88%), with the majority stating that there is a reluctance to use systemic therapies in this age group (4/7 clinicians, 57%) due to the fact that the immune system in younger patients is still developing. When further explored, all dermatologists with a specialist interest in hair disorders agreed that adolescents are less likely to be prescribed off-licence systemic therapies such as oral corticosteroids or conventional immunosuppressants (8/8 clinicians, 100%) and they would only use them in some instances. Such instances included for older adolescents (i.e., aged 16-17 years who have experienced puberty), those with a significant psychological effect due to the disease, those with no comorbidities, those with a rapidly progressive disease and patients with *"the most severe loss e.g., those potentially heading towards AT/AU"*. Other therapies indicated for use in adolescents include topical treatments such as topical corticosteroids or contact immunotherapy.

Other factors which dermatologists with a specialist interest in hair disorders noted would influence the treatment pathway were the severity and location of hair loss.[21] Generally, the dermatologists also indicated that topical or intralesional corticosteroids are likely to be used at the lower end of the disease severity spectrum (i.e., <50% scalp hair loss [4/8 clinicians, 50%]), and contact immunotherapy or off-licence systemic treatments are more likely to be used at the higher end (i.e., ≥50% hair loss including AT/AU [5/8 clinicians, 63%]). Most dermatologists with a specialist interest in hair disorders stated that the use of treatments varies depending on the type of hair

loss (e.g., scalp, eyebrow, beard [6/8 clinicians, 75%]). Two of these dermatologists with a specialist interest in hair disorders stated a reluctance to use intralesional corticosteroids on the face (2/6 clinicians, 33%) with one mentioning a *"risk of atrophy"* (i.e., decreased thickness of the skin).

It was clear that treatment choice is also based on a dialogue with the patient and their preferences and other factors such as comorbidities and duration of disease.[21]

B.1.3.3.2 Pharmacological therapies

There are several off-licence pharmacological therapies available described below but they are rarely used for patients with severe AA due to broad concerns on their safety, limited evidence for their efficacy, and not being widely available. Therefore, the vast majority of patients with severe AA are not receiving any pharmacological treatment.

Topical and intralesional steroids are rarely used for patients with severe AA, with intralesional steroids associated with risks of atrophy. Contact immunotherapy is time-intensive and dermatologists in less specialist practices may not be equipped in its application; therefore, it is not widely available in the UK. Finally, systemic treatment is off-licence and limited to provision from dermatology specialists; there is limited evidence for their efficacy and broad concerns on their safety. These limited treatment options can be uncomfortable, time consuming and have unacceptable levels of toxicity for the patients.

Contact immunotherapy

If available, contact immunotherapy would typically be used prior to systemic treatments because, according to clinical experts, once a patient is on the path of immunosuppression, it is not logical to go back to contact immunotherapy. Contact immunotherapy involves applying diphenylcyclopropenone (DPCP) (liquid formulation) to the scalp once a week. The concentration of the solution is increased at each treatment session until an allergic reaction occurs, this allergic reaction stimulates the immune system in the skin surface which is thought to divert the immune system away from the hair follicle, allowing the follicle to recover and start to grow hair again. When used, contact immunotherapy is regarded as a viable treatment option and there is evidence that a proportion of patients respond on treatment.[121] However,

some patients may have variable response rates and relapse rates, combined with adverse effects (such as eczematous reaction consisting of erythema, itching, and scaling at the site of application as well as swelling of cervical and/or occipital lymph nodes,[122] and lymphadenopathy)[121] that may lead to discontinuation. In addition, the length of time patients must remain on the treatment varies, which means a tailored approach that considers patient expectation for regrowth is needed to improve therapeutic adherence and outcome satisfaction.[121]

Corticosteroids

Intralesional corticosteroids are a simple treatment, and their effectiveness can be enhanced when given under occlusion (i.e., absorption is increased as the patient is wrapped in bandages).[21] However, they are rarely used for patients with severe AA given the extent of hair loss and the number of injections that would be required. Dermatologists with a specialist interest in hair disorders commented that they might use intralesional injections for the frontal hairline, eyebrows or potentially beard loss but this would still depend on a patient's preference for injections, especially due to clinician reluctance to use intralesional injections on the face.[21]

Older immunomodulators/suppressants

The BAD guidelines mention that broad immunomodulators or immunosuppressants (including methotrexate and ciclosporin) may be used in treatment refractory patients.[7] These treatments can be included in dermatologist's armamentarium based on experience in uncontrolled studies; however, limited evidence is available in AA patients.[119,123]

Use of immunosuppressants in the UK is mixed; methotrexate (MTX) was the most common immunosuppressant, mentioned by seven dermatologists with a specialist interest in hair disorders who use them in between 5% and 60% of patients with severe AA (7/8 clinicians, 88%). Mycophenolate mofetil (MMF), azathioprine, ciclosporin were used although to a lesser degree (each mentioned by two of the eight clinicians). However, interviews suggested that the preferred immunosuppressant varies by clinician. Unlike oral corticosteroids, there was no definite and consistent timeframe

given for treatment duration for immunosuppressants if a patient is responding to treatment and this varies by the specific immunosuppressant given.

Methotrexate can be effective in the treatment of a range of inflammatory and autoimmune disorders. In the management of AA, methotrexate can be used either as an adjunct for low-risk maintenance therapy after initiation with corticosteroids, or as monotherapy. In a meta-analysis of 16 studies, the pooled complete regrowth of hair in adults treated with methotrexate (n=361) was 44.7% (95% CI: 32.9%, 57.1%); 69.3% of the patients (n=219) had at least ≥50% regrowth (95% CI: 59.5%, 77.7%).[124]

Despite widespread use of ciclosporin for treatment of AA, there is little evidence in support of its efficacy. Studies of small numbers of cases indicate that patients with severe AA initially respond to treatment with ciclosporin, however there is limited evidence to support the durability of the response to treatment.[113,125] In a double-blind, randomised, placebo-controlled trial of adults with severe AA (N=36), no significant difference was reported between the ciclosporin and placebo treatment groups for ≥50% reduction in SALT score (p=0.07).[126] When ciclosporin was combined with the systemic corticosteroid methylprednisolone in patients with severe AA (N=43), complete and partial regrowth was observed in 10.9% and 71.7% of patients, respectively;[113] 23.7% of patients who had complete or partial response relapsed within 12 months of treatment cessation.

Systemic treatment

UK dermatologists with a specialist interest in hair disorders interviewed confirmed they use severe AA as a criterion for the use of systemic treatment,[21] though this use is off-licence. In discussion they supported that the vast majority of patients seen by clinicians have severe AA.[17] From the Therapeutic Landscape Delphi panel, it was identified that oral corticosteroids and immunosuppressant therapies are the most commonly used off-licence systemic treatment options .[21] All three dermatologists with a specialist interest in hair disorders specified that oral corticosteroid use is limited to <3 to 6 months given the toxicity profile; therefore, patients are commonly tapered off oral corticosteroids whilst an immunosuppressant is initiated (in combination with the tapering corticosteroid). Immunosuppression treatment is continued for a longer term (3/3 clinicians, 100%). Oral corticosteroids followed by conventional

immunosuppressants would therefore describe the off-licence use of systemic treatment for adult patients with severe AA that tends to be limited to prescription by dermatologists with a specialist interest in hair disorders.

Those consulted in the Therapeutic Landscape Delphi panel described that corticosteroids can be useful to determine whether the patient will respond to anything else (i.e., if they do not respond to oral corticosteroids, they are unlikely to respond to anything else).[21] The most commonly used oral corticosteroid by clinicians was reported to be prednisolone (5/8 clinicians, 63%).

Minoxidil

Topical minoxidil is approved for treatment of androgenetic alopecia (male or female pattern baldness), but it is also used off-licence (topical and oral) for other hair loss conditions. Minoxidil's mode of action for stimulation of hair regrowth is not fully understood (vasodilation and angiogenesis are among the mechanisms that have been postulated), but it does not have a direct immunomodulatory effect and thus is not expected to alter the course of AA or induce remission.

Although minoxidil is not recommended as a standalone treatment by clinicians to treat patients with AA due to a lack of efficacy, dermatologists with a special interest in hair disorders highlighted that they are aware of patients conducting their own research in treatment and buying minoxidil over the counter against the advice of their prescribing physician.[17]

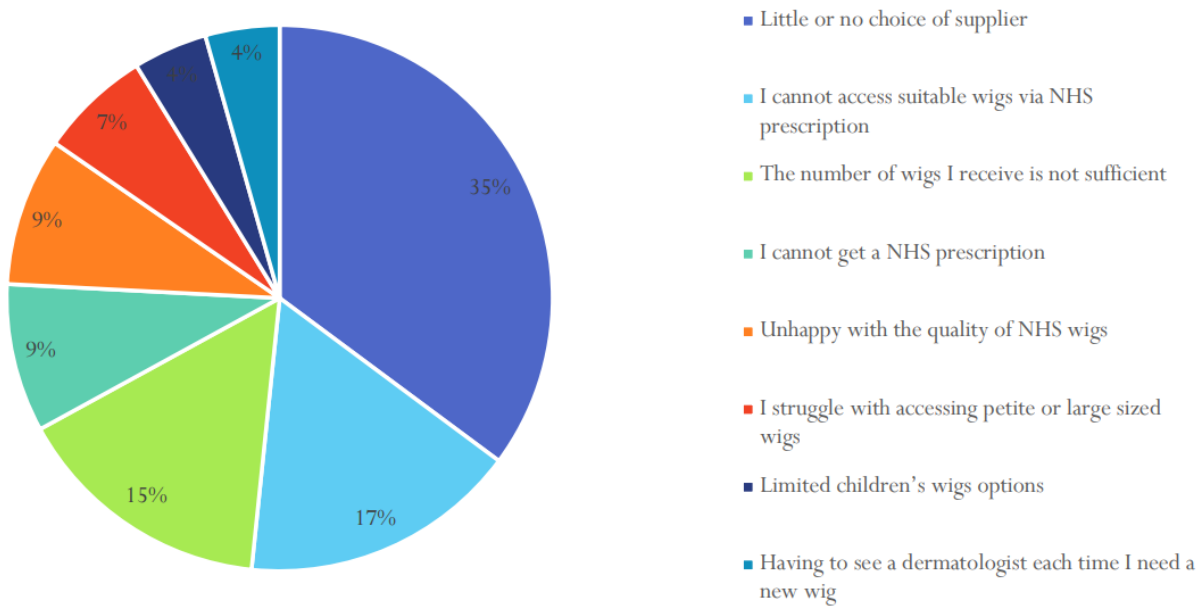
B.1.3.3.3 Non-pharmacological treatment options

For patients with severe AA, non-pharmacological prosthetics are a mainstay of treatment. Prosthetics can be used to mask hair loss, including the use of a wig, hairpiece, head cover or hat, false eyelashes and semi-permanent make-up. They may be used alongside pharmacological treatment.

Wigs are very commonly used. They can either be bought privately or are subsidised on the NHS for a subset of patients. Patients qualify for free wigs or a wig-voucher for reasons such as if they are under 16 years old, 16-18 years old and in full-time education, a hospital inpatient, receiving income support or valid for an NHS tax credit

exemption certificate.[127] Given that waiting lists to see a dermatologist can be long, many patients are unable to access prescription wigs or unaware they are entitled to prescription wigs. The quality of prescription wigs can often be poor;[105,106] PAG representatives reported that the provision of wigs was variable and becoming more difficult for patients.[22] In a survey conducted in the UK which asked what the biggest issue with NHS wig provision is, 35% of the 91 responders said the lack of supplier choice was the biggest challenge, as such in Figure 9. [128]

Figure 9: Community responses - biggest issue with NHS wig provision



Source: Johnson and Montgomery (2017)[128]
Abbreviations: NHS, National Health Service

There are regular reports of people who struggle to afford wigs privately; therefore if the NHS provides no support with wig provision, patients may have no access to wigs.[128] Market research conducted by Pfizer found that the cost of wigs sourced privately per patient varies from £50 to £130 per month (median: £80 per month).[15]

Prosthetics may provide acceptable cosmetic results for some, but appear to be inadequate or unacceptable for many patients;[129] some patients may have difficulty applying such prosthetics or fear that they may be discovered by others.[7] A mixed methods survey of social anxiety, anxiety, depression, and wig use in patients living with alopecia in the UK (N=338, including mainly patients with AA (N=114), AU (N=106) and AT (N=59), but also patients with other forms of alopecia (N=59) was

conducted. The results showed that 43% of participants reported that wearing a wig can have a negative impact on confidence during social situations; 47% worried about others knowing about their wig and 39% were concerned about the wig coming off or discomfort.[130] Wearing a wig also led to reduced activity in 41% of participants, in particular sports were avoided due to concerns about having to take off the wig. The majority of participants (65%) worried about affording new wigs and only 23% reported that wearing a wig improved their confidence/self-esteem in everyday life, showing that for the majority of patients wigs are not a viable solution to the burden of AA, including its psychological effects.[130]

B.1.3.3.4 BSC

BSC is the management of patients with AA in the absence of pharmacological treatment and includes non-pharmacological therapy alongside disease management. Dermatologists with a specialist interest in hair noted that it is difficult to define exactly what this includes as there is a variance between specialist centres.[17] An example provided by one of the dermatologists interviewed stated that they can only prescribe wigs to patients who live in their catchment area, and cannot prescribe wigs to patients who have travelled from outside of the catchment area.[17]

B.1.3.3.5 Current treatment unmet need

There is a substantial treatment unmet need in adult and adolescent patients with severe AA, as discussed below.

There are currently no licensed systemic treatment options for adult and adolescent patients with AA.

To date, there have been no advances in treatment options with current off-licence treatment options for severe AA limited by their efficacy, safety profile and tolerability.[7,20] Treatment can be uncomfortable, time consuming and have unacceptable levels of toxicity for the patients. The treatments also alter the patients' attitudes to their hair loss; relapse, either during or even after initially successful treatment, is common and some patients find it difficult to cope with this experience.[115,116] Patients may consider the AEs of these treatments combined with their unpredictable results unacceptable.[117]

All dermatologists with a specialist interest in hair disorders consulted in the Therapeutic Landscape Delphi panel stated that they were not satisfied with the current treatment options available for patients with severe AA (8/8 clinicians, 100%), with the majority also stating that they were not satisfied at all (6/8 clinicians, 75%), citing poor efficacy, lack of evidence on efficacy and cost-effectiveness, limited long-term maintenance options, and high risk of side effects.[21] One clinician, for example, said:[21]

"It is very frustrating managing patients with this degree of hair loss. There are limited options which are evidence based – some of which are very time consuming/labour intensive e.g., DPCP immunotherapy. Lack of maintenance options that are safe is also frustrating as treatment relapse can be unpredictable."

The limitations of current off-licence treatment options was reflected by PAG representatives, with the majority saying they are not satisfied, mainly due to the perceived lack of efficacy:[22]

[redacted]

It was noted that steroid treatments may work for the duration they were being given but that the benefit would stop once the treatment stopped.[22] Unpleasant side effects associated with steroids and immunotherapy were also cited as reasons for treatment dissatisfaction by PAG representatives.

An epidemiological study in a UK primary care setting (N=2,634,083) has shown that 46% of patients with AA did not receive any prescription medication for their AA which may result from the lack of treatment options currently available.[4]

Patients are limited to non-pharmacological treatment options, but these options are inadequate or unacceptable for many patients.[129] A recent study found that the majority of patients with alopecia (65%) worried about affording new wigs and only 23% reported that wearing a wig improved their confidence/self-esteem in everyday life.[130] As highlighted in Sections B.1.3.2.5 and B.1.3.3.2, the unmet need for patients with AA is so great that patients are frequently purchasing JAK inhibitors from

overseas and purchasing topical minoxidil over the counter despite its lack of efficacy in treating AA, against the advice of their prescribing physician.[17]

The treatment unmet need is great for patients with severe AA.As the likelihood of positive treatment outcomes reduces with severity of hair loss and duration of disease,[7,18,19] there is a clear need for effective treatments to manage severe AA.

As discussed throughout Section B.1.3.3, although there are a number of off-licence pharmacological treatment options available, none of them are considered suitable for use by dermatologists with a specialist interest in hair disorders, nor are they widely available for use in the treatment of patients with severe AA.[17,21] The reasons each of the pharmacological treatment options are not suitable are described in further detail below.

Contact immunotherapy with DPCP may be offered although it is unlicensed, not of pharmaceutical grade (this means its purity has not been established by the British Pharmacopeia) and is not widely available in the UK. Dermatologists in less specialised practices may not be equipped to use it, and patients may not be able to commit to the required weekly clinic visits.[7,131,132] Dermatologists with a specialist interest in hair disorders also commented that it is not offered consistently in the UK.[17]

Long-term use of systemic steroids is limited by side effects such as weight gain, osteoporosis, cataract formation, hypertension, peptic ulceration, metabolic abnormalities, gastrointestinal irritation, and suppression of the hypothalamic-pituitary-adrenal axis along with high relapse rates (22–100%).[116,133,134] Treatment of AA with systemic corticosteroids is not recommended in the BAD guidelines,[7] due to the inevitable systemic toxicities associated with their short-[135] and long-term use.[136]

For patients who are contraindicated or with high-risk factors, both oral corticosteroids and immunosuppressants treatments might be avoided. Use of oral corticosteroids particularly is still controversial due to lack of evidence, safety, short term use and high relapse rate. Moreover, dermatologists with a specialist interest in hair disorders advised that their use is avoided in general practice.[17]

There are also limitations in treating patients with immunosuppressants from a safety perspective. Gastrointestinal symptoms (nausea, epigastric pain, and diarrhoea) have

been reported in 9.3% of patients treated with methotrexate and prednisolone;[137] this regimen is also associated with haematologic (mild to moderate leukopenia), and hepatic (mild, transient increase in transaminases) adverse events (AEs), which have been reported in 9.7% and 6.5% of patients, respectively. Another study has found methotrexate treatment to be associated with transient elevated transaminases (12.1%), persistent nausea (6.1%) and lymphocytopenia (3.0%).[138] Methotrexate plus prednisolone was associated with acne (20%), muscle cramp (20%), anaemia (10%), hypertension (10%), weight gain (10%) and amenorrhea (10%).[139]

Due to the systemic side effects of these older agents, including but not limited to hypertension, malignancies, bone marrow suppression, hepatotoxicity, nephrotoxicity, and pancytopenia, these treatments are not commonly recommended or appropriate for long-term treatment.[7,140] A dermatologist with a specialist interest in hair disorders commented that only clinicians with a particular interest in hair would prescribe an immunosuppressant.[17]

Minoxidil is sometimes used off-licence as adjunctive therapy to oral or topical steroids to stimulate hair growth in AA, although evidence is scarce and histological studies on its effect on perifollicular lymphocytic infiltration in AA are inconsistent.[141–144] Its use is less common in severe disease due to limited efficacy and unsatisfactory results when used as monotherapy.[145]

All dermatologists with a specialist interest in hair disorders consulted in the Therapeutic Landscape Delphi panel agreed that oral minoxidil is not commonly used as a standalone treatment for patients with AA with ≥50% hair loss including AT/AU (8/8 clinicians, 100%).[21]

The current treatment unmet need is especially apparent for adolescent patients with severe AA as clinicians are less likely to prescribe adolescents off-licence systemic therapies due to the fact that the immune system in younger patients is still developing, as further described in Section B.1.3.3.1

This unmet need is greatest of all in patients with AT/AU, where dermatologists with a specialist interest in hair disorders consulted in the Therapeutic Landscape Delphi panel agreed that patients with AT or AU are particularly challenging to treat (8/8

clinicians, 100%).[21] Similarly, PAG representatives noted dissatisfaction to be higher amongst patients seeking treatment for AT and AU as therapies tend to be even less effective.[22]

Adult and adolescent patients with severe AA are least likely to respond to the available treatments and therefore, there is an unmet need for well tolerated treatments that demonstrate satisfactory improvements in hair growth.

B.1.3.3.6 Place of ritlecitinib in treatment pathway

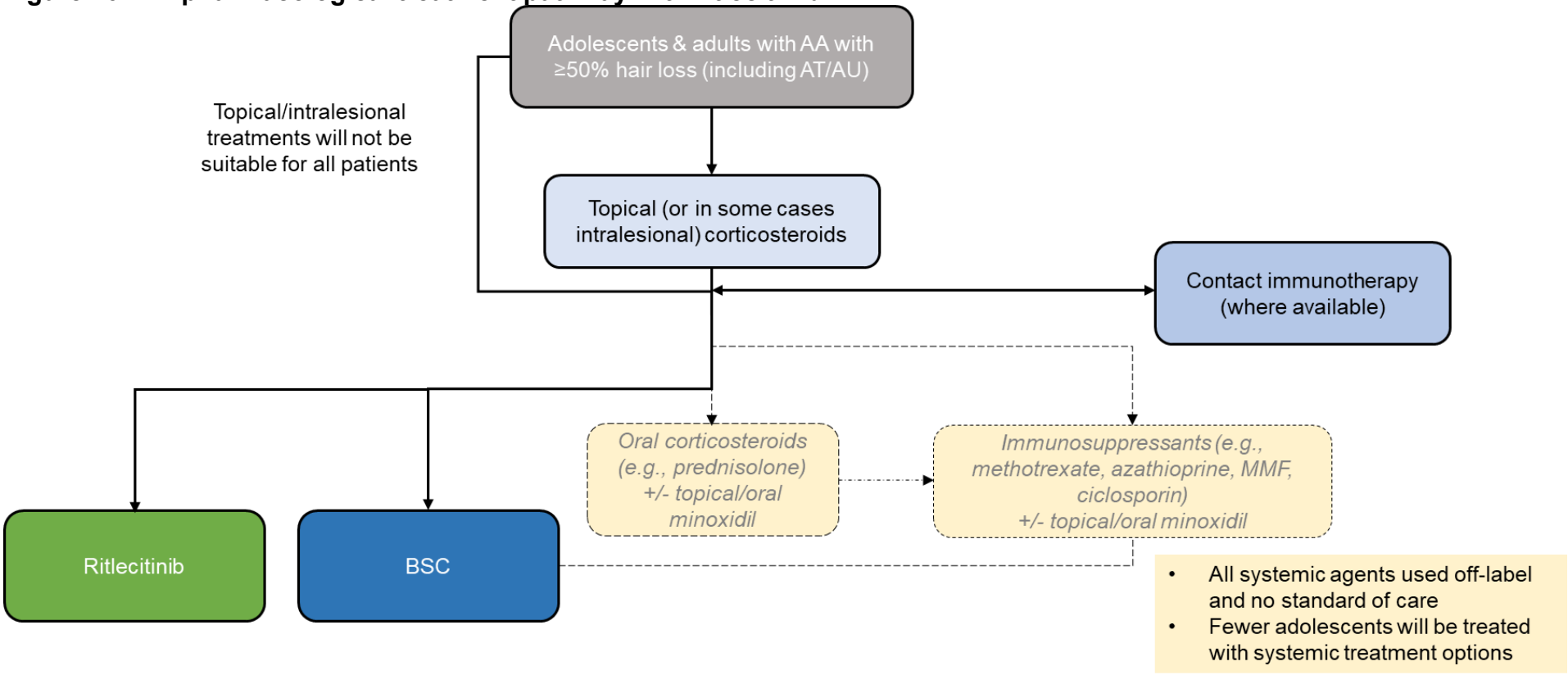
Ritlecitinib is expected to be placed as a systemic treatment option for patients with severe AA, aligned with the technology's full marketing authorisation for this indication.

As discussed throughout Section B.1.3.3.2, the limited options for patients with severe AA are; topical and intralesional steroids, contact immunotherapy and off-licence use of systemics (such as corticosteroids and immunosuppressants). Topical and intralesional steroids are rarely used for patients with severe AA, being impractical for patients with extensive hair loss, with intralesional steroids further associated with risks of atrophy. Contact immunotherapy is time-intensive and dermatologists in less specialist practices may not be equipped in its application; therefore, it is not widely available in the UK. Finally, systemic treatment is off-licence and limited to provision from dermatology specialists; there is limited evidence for their efficacy and broad concerns on their safety. These limited treatment options can be uncomfortable, time consuming and have unacceptable levels of toxicity for the patients. Furthermore, due to the reluctance of clinicians to prescribe off-licence systemic options to adolescents, there is an opportunity to provide treatment options in this patient group.

Because of these reasons, there is an unmet need for well tolerated treatments that demonstrate satisfactory improvements in hair growth. During validation interviews with dermatologists with a specialist interest in hair disorders, they all agreed that the most relevant comparator for ritlecitinib is best supportive care with non-pharmacological treatment (3/3 clinicians, 100%);[17]there are no pharmacological treatment options which can be considered as a comparator to ritlecitinib. The only comparator of interest is therefore best supportive care defined as non-pharmacological therapy (see Figure 10). Based on payor feedback, BSC defined as

non-pharmacological therapy aligns with the placebo group in the ALLEGRO 2b/3 pivotal trial in terms of baseline characteristics and use of prior treatments. Patients on placebo were permitted to use camouflage items, such as wigs, which makes the placebo arm also generalisable to the UK population.

Figure 10: AA pharmacological treatment pathway with ritlecitinib



Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; BSC, best supportive care; MMF, mycophenolate mofetil

B.1.4. Equality considerations

There are currently no licensed systemic treatment options for patients with severe hair loss in the UK. Contact immunotherapy, which may be used prior to systemic treatment options, requires multiple clinic visits over several months and it is not widely available in the UK, thus resulting in inequality of access.

Further, there is a disparity in wig provision across NHSE with local NHS organisations setting limits on the number of wigs available to patients. Many NHS organisations do not have a wig policy in place;[128] furthermore, six NHS trusts reported that alopecia is a "cosmetic issue" and for this reason there is no funding available for people living with alopecia. This raises concerns as the emotional and psychological consequences of alopecia have been widely acknowledged in research, as described in Section B.1.3.2.2, and reporting that alopecia is a "cosmetic issue" suggests the psychological needs of patients are not being taken into account.[128] Furthermore, based on interviews with dermatologists with a specialist interest in hair disorders, patients would only be eligible to receive camouflage treatments such as wigs if they lived within specific catchment areas because of NHS budget constraints.[17] Patients with AA receiving care from NHS trusts which consider alopecia a "cosmetic issue" or receiving care outside of their catchment area are therefore at a disadvantage and will not have access to wig provision through the NHS.[128] This further exacerbates the inequity in access to treatment. In addition to the disparity in wig provision, not all patients are able to wear wigs comfortably – an example is patients with alopecia who also need to wear hearing aids; these patients can find it difficult to get wigs to fit properly and also experience a rustling sound when the wig is next to the ears.

The frequency of AA is also higher for people with skin of colour, particularly for those of Asian backgrounds where it is more than three times more common than in Caucasians.[4] A UK population-based cohort study found that people of non-white ethnicity were more likely to present with AA, especially those of an Asian ethnicity (incident rate ratio (IRR) 3.32). A higher AA incidence was associated with social deprivation (IRR most vs. least deprived quintile 1.47). People from more deprived backgrounds living in urban areas are also more likely to have AA and are less likely to be able to pay for higher quality wigs, where these are not accessible via the NHS.[4]

In addition, people of higher social deprivation were less likely to be referred for specialist dermatology review.[4]

Out of pocket expenses also contribute to the inequities of access to treatments across the UK. Based on interviews with dermatologists with a specialist interest in hair disorders, patients would only be eligible to receive camouflage treatments such as wigs if available subject to local NHS budget constraint.[17] As a result, where there is no budget for wigs on the NHS, patients from more economically deprived communities are disproportionately affected, resulting in considerable inequity.

B.2 Clinical effectiveness

The efficacy and safety of ritlecitinib as a therapy for patients with severe alopecia areata (AA) has been conclusively demonstrated in the ALLEGRO 2b/3 study, a large, international, placebo-controlled trial.

The primary endpoint from the ALLEGRO 2b/3 study was response defined as absolute SALT score ≤20 at week 24; this primary outcome was met and superiority of ritlecitinib versus placebo was demonstrated (Section B.2.6). This demonstrates that over a 24-week period, patients treated with 50 mg of ritlecitinib (once daily) were significantly more likely to experience clinically meaningful hair regrowth than patients who were treated with placebo.

- Statistically significant improvements for ritlecitinib 50 mg versus placebo were observed across further critical endpoints related to hair regrowth, including response based on an absolute SALT score ≤10 at Week 24, which represents near-to-complete hair regrowth, and change from baseline in SALT score by Week 24.
- A significantly greater improvement from baseline in patient global impression of change (PGI-C) was observed for patients treated with ritlecitinib 50 mg versus placebo, demonstrating that significantly more patients treated with ritlecitinib than placebo perceived a positive change in AA as a result of treatment.
- Treatment with ritlecitinib demonstrated improvements in hair regrowth of the eyebrows and eyelashes. The proportion of participants in the 50 mg group with an eyelash assessment (ELA) and eyebrow assessment (EBA) response (≥2-grade improvement or normal ELA/ EBA score) at Week 24 was clinically meaningful and higher than the proportion of participants in the placebo group and increased over time, indicating that participants' hair regrowth was not limited to the scalp.

The ALLEGRO-LT study reports a similar trajectory to ALLEGRO 2b/3 in the attainment of SALT scores ≤20 and ≤10 over prolonged treatment, supporting the findings of the ALLEGRO 2b/3.

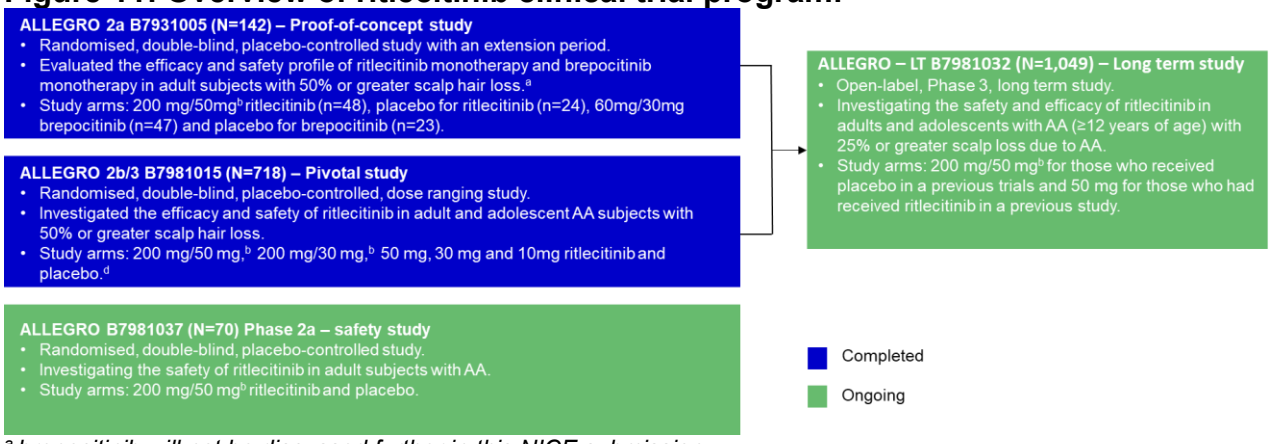
B.2.1. Identification and selection of relevant clinical studies

A systematic literature review (SLR) was conducted to identify relevant literature regarding the efficacy and safety of ritlecitinib compared to other therapies for AA. This SLR was not restricted to the treatment of severe AA. In total, the SLR identified 163 publications reporting on 156 unique studies; of these, two studies reported data that is relevant to the decision problem. Evidence specially addressing the NICE scope and relevant to the UK for ritlecitinib was included. Full details of the methodology and results of the SLR are detailed in Appendix D.

B.2.2. List of relevant clinical effectiveness evidence

The clinical effectiveness of ritlecitinib in the treatment of AA is being examined in a robust clinical trial programme. ALLEGRO Phase 2a proof-of-concept study NCT02974868 (ALLEGRO 2a) and ALLEGRO Phase 2b/3 pivotal study NCT03732807 (ALLEGRO 2b/3) have been completed (coloured blue in Figure 11). ALLEGRO Phase 2a mechanistic study NCT04517864 and a longer term ALLEGRO study NCT04006457 (ALLEGRO-LT) are still ongoing (coloured green in Figure 11).

Figure 11: Overview of ritlecitinib clinical trial program.[146–149]



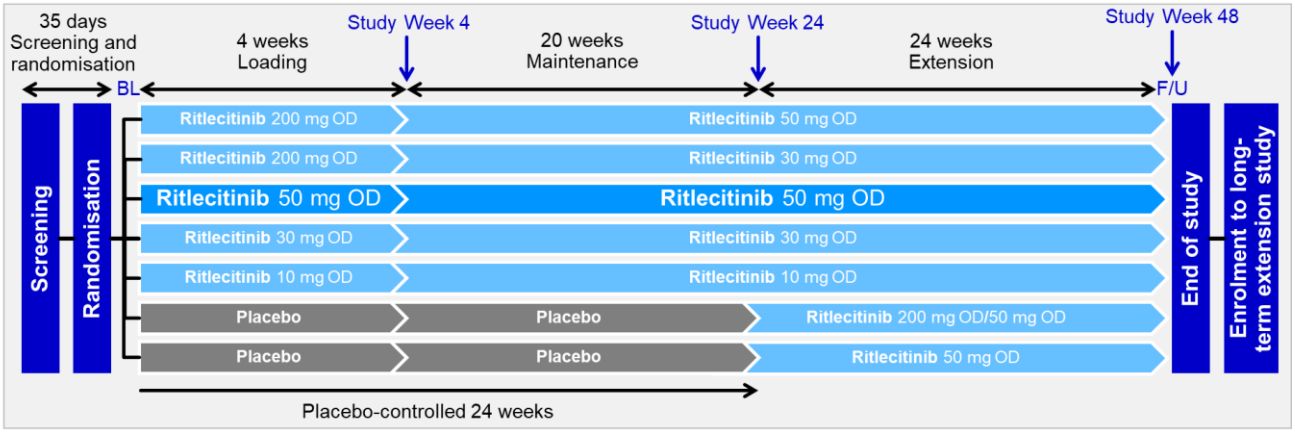
[a] *brepocitinib will not be discussed further in this NICE submission*
[b] *participants received 200 mg ritlecitinib OD for four weeks followed by 50 mg/30 mg OD for the remainder of the trial*
[c] *participants received 60 mg brepocitinib OD for four weeks followed by 30 mg OD for the remainder of the trial*

B.2.2.1. ALLEGRO 2b/3: pivotal study demonstrating the efficacy of ritlecitinib

ALLEGRO 2b/3 is a completed Phase 2b/3 randomised, double-blind, placebo-controlled, dose-ranging pivotal study in adults and adolescents (12 years of age and older) with AA.[147] The study had a maximum duration of 57 weeks including a 5-week

screening period, a 48-week treatment period, and a 4-week follow-up period as shown in Figure 12. The 48-week treatment period consisted of a 24-week placebo-controlled period (four-week loading phase, a 20-week maintenance phase) and a 24-week extension phase where placebo-treated patients switched to active treatment with ritlecitinib in a pre-specified, blinded manner, while other arms continued on the same maintenance dose.

Figure 12: Study design of ALLEGRO 2b/3[2]



Abbreviations: BL, baseline; F/U, follow-up; OD, once daily

Two treatment groups were treated (once daily) with a loading dose of 200 mg ritlecitinib for four weeks before switching to 50 mg or 30 mg ritlecitinib to Week 48 (regimen referred to as 200/50 mg and 200/30 mg, respectively). In the third group, patients were treated with ritlecitinib 50 mg (once daily) throughout the whole study period (48 weeks) with no loading dose administered. In two further treatment groups, patients were treated with 30 mg and 10 mg ritlecitinib (once daily) throughout the whole study period with no loading dose administered. Ritlecitinib 10 mg (once daily) was included in the study exclusively to support the estimation of the exposure-response. The final two groups were treated (once daily) with a placebo for the first 24 weeks before switching to receive either: a loading dose of 200 mg ritlecitinib for four weeks before being treated with 50 mg for 20 weeks, or 50 mg ritlecitinib for the full extension period.

Investigators, subjects, and the sponsor study team were blinded as to the treatment group throughout the duration of the study. Following their last dose of ritlecitinib, discontinued and completed subjects entered a 4-week follow-up period for safety monitoring.

Patients included in the study had severe AA (>50% scalp hair loss), including AT and AU, had no evidence of hair regrowth within the previous 6 months, and their current episode of hair loss was ≤10 years.

The study completion date was 24 June 2021. The study enrolled a total of 718 participants at 118 sites globally, including 6 sites in the UK. The results of this study inform the clinical- and cost-effectiveness of ritlecitinib in this submission. An overview of this pivotal study is provided in Table 7.

Table 7: Overview of the pivotal clinical study

	ALLEGRO 2b/3[147]
Study design	Randomised, double-blind, placebo-controlled, dose-ranging pivotal study.
Population	Patients with 50% hair loss of the scalp (SALT score ≥50% [severe AA]) at both screening and baseline visits, without evidence of terminal hair regrowth within the previous 6 months and with the current episode of hair loss ≤10 years.
Intervention(s)	Once daily: Ritlecitinib 200/50: 200 mg (loading dose; 4 weeks), followed by 50 mg maintenance dose (20 weeks) (n=132) Ritlecitinib 200/30: ritlecitinib 200 mg (loading dose; 4 weeks), followed by 30 mg (maintenance dose; 20 weeks) (n=130) Ritlecitinib 50 mg (licensed dose) (n=130) Ritlecitinib 30 mg (n=132) Ritlecitinib 10 mg (n=63)
Comparator(s)	Placebo-ritlecitinib 200 mg/50 mg (n=65) Placebo-ritlecitinib 50 mg (n=66) At week 24, patients who were randomised to receive placebo were re-randomised to receive either 200 mg/50 mg or 50 mg ritlecitinib (Figure 12)
Indicate if the study supports the application for marketing authorisation	Yes
Indicate if the study is used in the economic model	Yes – ritlecitinib 50 mg (once daily)
Rationale if study not used in the model	N/A
Reported outcomes specified in the decision problem	Disease severity and improvement in hair loss SALT Score PGI-C EBA ELA HRQoL HADS AAPPO EQ-5D SF-36 P-Sat Adverse events SAE Non-serious AE Other

	ALLEGRO 2b/3[147]
	AARU WPAI
All other reported outcomes	N/A

Abbreviations: AA, alopecia areata; AARU, AA resource utilisation; AE, adverse events; CGI, Clinician Global Impression; EBA, eyebrow assessment; ELA, eyelash assessment; HADS, hospital anxiety and depression scale; HRQoL, health-related quality-of-life; N/A, not applicable; SALT, Severity of Alopecia Tool; PGI-C, patients global impression of change; P-Sat, patient satisfaction; SAE, serious adverse events; SF-36, Short Form 36 Health Survey Questionnaire; WPAI, Work Productivity and Activity Impairment

B.2.2.2. Supporting studies

B.2.2.2.1 ALLEGRO 2a proof of concept study[146]

The ALLEGRO 2a study has been completed. It was a Phase 2a, randomised, double-blind, parallel-group, multi-centre study evaluating the safety, tolerability, pharmacokinetics (PK) and efficacy of ritlecitinib and brepocitinib in subjects with AA. Adults at least 18 years of age qualified for inclusion if they had severe AA (≥50% scalp hair loss; including AT and AU), no hair regrowth within 6 months of the screening and baseline visits, and a current episode of fixed hair loss of 7 years or less in duration.

The primary endpoint was change from baseline in SALT score at Week 24. The key secondary endpoint was the proportion of participants achieving SALT 30 at week 24.

The study was designed with the primary objective to meet the single Phase 3 pivotal safety and efficacy trial criteria, both in the context of the guidelines for single pivotal trials and in the context of the CHMP advice regarding the selection of clinically relevant clinical endpoints and application of the stricter statistical significance threshold. All patients had a 4-week loading dose of either 200 mg ritlecitinib (once daily) or 60 mg brepocitinib (once daily) before being randomised to 50 mg ritlecitinib, 30 mg brepocitinib, or placebo for 20 weeks. The safety and efficacy outcomes were consistent with what is seen in ALLEGRO 2b/3.

The study completion date was May 2019; data from this study is published in King *et al.,* 2021.[37] ALLEGRO 2a is not presented further in this submission because subjects received a daily loading dose for 4 weeks prior to treatment with the randomised ritlecitinib dose which does not align with the proposed licensed use. Brepocitinib, also assessed in this study, is not a focus of this submission.

The methodologies for two additional active studies are described below.

B.2.2.2.2 ALLEGRO-2a safety study[150]

ALLEGRO-2a safety study is a global Phase 2a randomised, double-blind, placebo-controlled study to evaluate the safety and tolerability of ritlecitinib in adults aged 18 to ≤50 years of age with ≥25% scalp hair loss due to AA. The study has enrolled 71 participants that will be randomised to receive ritlecitinib 200 mg (once daily) for 4 weeks then ritlecitinib 50 mg OD through to month 60. The control arm of the study received placebo followed by active therapy extension with ritlecitinib. Patients received placebo once daily (4 capsules for 4 weeks then 1 tablet for 8 months) then 200 mg ritlecitinib (four 50 mg capsules) for 4 weeks then 50 mg ritlecitinib capsules through month 24. After month 24, participants continue on 50 mg capsules (once daily), up to month 60. At month 9, participants assigned to this treatment arm will also receive 3 capsules of placebo for 4 weeks to maintain the blind with the other arm.

The primary outcome of this safety study is to measure functional auditory testing via the BAEP at a stimulus intensity of 80 decibels (dB) at Month 9 (time frame: baseline, month 9; see Appendix D for more trial details).

The primary completion date is January 4th, 2022, and the study completion date is estimated to be January 8th 2026. The ALLEGRO 2a safety study is not presented further in this submission as its objective was to provide safety data only. Details on safety across all study of ritlecitinib can be found in Appendix F.

B.2.2.2.3 ALLEGRO – Long Term (LT) study[149]

ALLEGRO-LT is an ongoing long-term Phase 3 open-label, multi-centre study which is investigating the safety and efficacy of ritlecitinib in adults and adolescents (12 years of age and older). This study is recruiting subjects who have previously participated in either the ALLEGRO 2a or 2b/3 study as well as approximately 350 additional '*de novo*' adult and adolescent subjects not previously enrolled in either study (without evidence of hair regrowth within the previous 6 months and with a current episode of hair loss of ≤10 years). *De novo* participants and those with a gap of 30 days between phase 2a or phase 2b/3 are required to have ≥25% scalp hair loss, including AT and AU, to be eligible to enrol in the study. Moreover, participants from the Phase 2a &

2b/3 studies were eligible to enrol within >30 days between their last dose in their prior study and first visit in ALLEGRO-LT, regardless of their SALT score.

Participants who did not receive ritlecitinib in either the Phase 2a or 2b/3 study will receive 200 mg ritlecitinib (once daily) for four weeks, followed by once daily 50 mg for 35 months. Participants previously treated with ritlecitinib in either study will receive ritlecitinib 50 mg (once daily) for 36 months. As a primarily safety-focused study, the primary outcomes focus on key safety signals including the occurrence or adverse events. The primary completion date and study completion dates are estimated to be July 2024 and January 2026, respectively.

B.2.2.3. Patient population

As described in Section B.1.1, the proposed positioning for ritlecitinib is for adult and adolescent patients with AA with severe hair loss. This is aligned with the anticipated licensed population in which ritlecitinib was studied in ALLEGRO 2b/3.

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

B.2.3.1. ALLEGRO 2b/3 study

The methodology for the ALLEGRO 2b/3 study is summarised below in Table 8 and discussed further in the following subsections. The methodology for ALLEGRO 2b/3 has been taken from the study protocol, clinical study report (CSR) and additional sources added as needed.[151–154]

Table 8: Summary of ALLEGRO 2b/3 study methodology

ALLEGRO 2b/3 (B7981015; NCT03732807)[154]	
Study objective	To evaluate the efficacy and safety of ritlecitinib in adult and adolescent AA subjects with 50% or greater scalp hair loss (measured by an absolute SALT score ≤20) at Week 24.
Study location	118 sites globally (including 6 sites in the UK)
Method of randomisation	A stratified randomisation was used to achieve a target global composition for AT/AU and adolescent subjects in the enrolled population. The targets for enrolment were approximately 40% AT/AU and approximately 15% adolescents. • In regions enrolling both adolescents and adults, there were four strata: - >18 years of age and AT/AU, - <18 years of age and not AT/AU, - ≥18 years of age and AT/AU and, - ≥18 years of age and not AT/AU

ALLEGRO 2b/3 (B7981015; NCT03732807)[154]	
	• Within each of these strata, subjects were randomised in a 2:2:2:2:1:1:1 manner to blinded ritlecitinib and matching placebo for a total of 7 treatment sequences. • In regions enrolling only adults, there were two strata: - ≥18 years of age and AT/AU and, - ≥18 years of age and not AT/AU • In these regions, subjects were randomised using the same ratio as described for regions enrolling both adolescents and adults. • Randomisation was performed via an interactive web response system.
Method of blinding (care provider, patient, and outcome assessor)	• Investigators, subjects, and the sponsor study team were blinded as to treatment group throughout the duration of the study. • At the end of the maintenance period, placebo-treated subjects were advanced in a prespecified, blinded manner to one of two active treatment sequences for the remainder of the study (through Week 48). • In order to achieve the proper dosage and maintain the blind throughout the study, capsules were dispensed in a blinded fashion to ensure that all subjects, regardless of the assigned treatment sequence, took the same number of capsules/day.
Key eligibility criteria for participants (full criteria reported in Appendix D)	Inclusion criteria: • Male or female aged ≥12 years • Have a clinical diagnosis of AA with no other etiology of hair loss (e.g., telogen effluvium, androgenetic alopecia). • ≥50% hair loss of the scalp (measured by SALT), including AT and AU, without evidence of terminal hair regrowth within 6 months at both the screening and baseline visits. • Current episode of hair loss ≤10 years. • If receiving permitted concomitant medications for any reason other than AA, subjects should have been on a stable regimen, which was defined as not starting a new drug or changing dosage within 7 days or 5 half-lives (whichever is longer) prior to Day 1. Subject must have been willing to stay on a stable regimen during the duration of the study. Exclusion criteria: • Participation in other studies involving investigational drug(s) within 8 weeks or within 5 half-lives (if known), whichever is longer, prior to study entry and/or during study participation. • Other types of alopecia • Other scalp disease that may impact AA assessment • Active systemic diseases that may cause hair loss • Any psychiatric condition including recent or active suicidal ideation or behaviour that meets any of the following criteria:

ALLEGRO 2b/3 (B7981015; NCT03732807)[154]	
	○ Suicidal ideation associated with actual intent and a method or plan in the past year: "Yes" answers on items 4 or 5 of the C-SSRS ○ For subjects who had previous history of suicidal behaviours in the past >1 year to <5 years: "Yes" answer (for events that occurred in the past 5 years) to any of the suicidal behaviour items of the C-SSRS or any lifetime history of serious or recurrent suicidal behaviour, a risk assessment must be performed, and documented, by a qualified mental health professional to assess whether it is safe for the subject to participate in the trial ○ Clinically significant depression as indicated by the PHQ-8 total score ≥15 ○ The presence of any current major psychiatric disorder that is not explicitly permitted in the inclusion/exclusion criteria NOTE: For any subject who had significant depression or any suicidal behaviour, the subject was not randomised and referred for appropriate evaluation and treatment.
Duration of study	A maximum duration of 57 weeks, including a five-week screening period, a 48-week treatment period, and a four-week follow-up period.
Trial drugs	In total, 718 participants were randomised to receive once daily treatment as follows: • Ritlecitinib 200/50 mg: n=132 • Ritlecitinib 200/30 mg: n=130 • Ritlecitinib 50 mg: n=130 • Ritlecitinib 30 mg: n=132 • Ritlecitinib 10 mg: n=63 • Placebo-ritlecitinib 200/50 mg: n=65 • Placebo-ritlecitinib 50 mg: n=66
Permitted and disallowed concomitant medications (full list presented in Appendix D)	Permitted: • CYP3A inhibitors (given there is no interaction with ritlecitinib) • Acetaminophen may be used intermittently (not to exceed 3 g/day) Disallowed: • Medications and treatments that could affect AA • Medications with potential drug-drug interactions for potential safety concerns.
Primary outcomes	In this submission, we describe the hierarchy of endpoints according to the study protocol. • The primary endpoint for the study was response based on an absolute SALT ≤20 at Week 24. The primary endpoint for regulatory agencies varies as follows • EMA: the proportion of participants achieving an absolute SALT score ≤10 (response) at Week 24. • FDA: response based on SALT ≤20 at Week 24.
Secondary/exploratory outcomes	Key secondary endpoints: • Study: response based on absolute SALT ≤10 at Week 24 • EMA: PGI-C response (a measure of patient reported treatment benefit) defined as a score of "moderately improved" or "greatly improved" at Week 24. • FDA: N/A

ALLEGRO 2b/3 (B7981015; NCT03732807)[154]	
	Additional secondary endpoints: • Exposure-response characterised by response based on an absolute SALT score ≤20 at Week 24 • Response based on an absolute SALT score ≤20 and ≤10 up to Week 48 • Response based on a 75% improvement in SALT score from baseline (SALT 75) up to Week 48 • Change from baseline in SALT scores at up to Week 48 • Response based on at least a 2-grade improvement from baseline or a score of 3 in Eyebrow Assessment (EBA) score and Eyelash Assessment (ELA) score up to Week 48 ○ EBA and ELA are an NRS developed to characterise eyebrow/eyelash hair loss. The numeric rating scale ranges from 0 (none) to 3 (normal). • PGI-C response defined as a score of "moderately improved" or "greatly improved" up to Week 48 • Change from baseline in AAPPO up to Week 48 Exploratory outcomes: • Improvement on P-Sat items was included as an exploratory endpoint. P-Sat measures patient-reported satisfaction with hair growth across three items, amount of hair grown, overall hair grown back and quality of new hair. Improvement, defined as "slightly", "moderately", or "very satisfied", was measured up to Week 48.
Patient-reported assessment	• AAPPO was measured on baseline, Week 4, Week 8, Week 12, Week 18, Week 24, Week 34, Week 40 and Week 48. • PGI-C and P-SAT were both measured on Week 4, Week 8, Week 12, Week 18, Week 24, Week 34, Week 40 and Week 48. • HADS was measured on baseline, Week 4, Week 8, Week 12, Week 24 and Week 48. • EQ-5D-5L and EQ-5D-Y were both measured on baseline, Week 4, Week 12, Week 24 and Week 48. • PHQ-8 was measured during the screening period. • SF36v2 Acute measured on baseline, Week 4, Week 8, Week 12, Week 24 and Week 48. • AARU was measured on baseline, Week 12, Week 24, Week 34 and Week 48. • WPAI was measured on baseline, Week 12, Week 24, Week 34 and Week 48.
Safety assessments performed	• SAE • Non-serious AE
Pre-planned subgroups	• Age, • BMI, • Weight, • Gender, • Race, • Region, • AA severity, • Duration since AA diagnosis, • Duration of current AA episode,

ALLEGRO 2b/3 (B7981015; NCT03732807)[154]	
	• Prior pharmacological treatment from AA.

Abbreviations: AA, alopecia areata; AAPPO, Alopecia Areata Patient Priority Outcomes; AARU, Alopecia Areata Resource Utilisation; AE, adverse event; AT, alopecia totalis; AU, alopecia universalis; CYP3A, Cytochrome P450, famil3, subfamily A; EBA, Eyebrow assessment; ELA, Eyelash assessment; EU, European Union; FDA, Food and Drug Administration; HADS, Hospital Anxiety and Depression Scale; MCS, mental component score; PGI-C-Patient's Global Impression of Change; PHQ, Patient Health Questionnaire – 8 Items; P-Sat, Patient Satisfaction with Hair Growth; SAE, serious adverse event; SALT, Severity of Alopecia Tool; SF36v2 Acute, 36-Item Short Form Health Survey Version 2 Acute; UK, United Kingdom; WPAI: AA, Work Productivity and Activity Impairment: Alopecia Areata

B.2.3.1.1 Trial outcome definitions

ALLEGRO 2b/3 pivotal study protocol (see Section B.2.2.1)

Multiple endpoints were assessed in ALLEGRO 2b/3; the including study protocol trial endpoints and the FDA and EMA driven endpoints. The key study protocol trial endpoints are presented in this submission (Table 9). Further endpoints are discussed in B.2.7

Table 9: ALLEGRO 2b/3 protocol primary and key secondary objectives and endpoints

Objectives	Endpoints	
Primary		
To evaluate the efficacy of ritlecitinib compared to placebo in adult and adolescent AA participants with ≥50% scalp hair loss on regrowth of lost hair	**Overall Study and FDA** Response based on an absolute SALT score ≤20 at Week 24	**EMA** Response based on an absolute SALT score ≤10 at Week 24
Key Secondary		
Overall Study To evaluate the efficacy of ritlecitinib compared to placebo in adult and adolescent AA participants with ≥50% scalp hair loss on regrowth of lost hair	**Overall Study** Response based on absolute SALT score ≤10 at Week 24	
EMA To evaluate the effect of ritlecitinib on patient-centered outcomes at Week 24	**EMA** PGI-C response defined as a score of "moderately improved" or "greatly improved" at Week 24	

Abbreviations: AA, alopecia areata; EMA, European Medicines Agency; FDA, United States Food and Drug Administration; PGI-C, Patient's Global Impression of Change; SALT, Severity of Alopecia Tool
Source: Pfizer B7981015 Top Line Report (2021)[155]

The primary endpoint for the ALLEGRO 2b/3 study was response defined as an absolute SALT ≤20 at Week 24. SALT is a quantitative assessment of AA severity based on scalp hair loss, as described in Section B.1.3.1. Change from baseline is

defined as the baseline value minus the value at a specific visit. A positive change from baseline signifies an improvement.

The study was designed with the primary objective to meet the 'single Phase 3 pivotal safety and efficacy trial' criteria, both in the context of the guidelines for single pivotal trials and in the context of the CHMP advice regarding selection of clinically relevant clinical endpoints and application of the stricter statistical significance threshold.

The trial outcome definitions were agreed in a scientific advice meeting with the EMA Committee for Medicinal Products for Human Use (CHMP) in 2018. Ref: European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) advice letter (EMEA/H/SA/3875/1/2018/HTA/III). This advice was given before the Brexit withdrawal agreement was implemented. Post Brexit implementation the MHRA adopted the EMA scientific advice and the same dossier was submitted to the MHRA and the EMA.

Secondary outcome definitions are reported in Appendix D.

B.2.3.1.2 Baseline demographics and disease characteristics

The full trial population baseline characteristics for participants enrolled in the ALLEGRO 2b/3 study are presented in Table 10.[154]

The majority of participants [redacted] were adults; a total of [redacted] adolescents were enrolled. In the ritlecitinib 50 mg arm, age, ethnicity and primary diagnosis and duration were balanced compared to the placebo arms, assuming a minimally important difference of no more than 5%. There was a higher percentage of males in the ritlecitinib 50 mg arm compared to both placebo arms (ritlecitinib 50 mg = [redacted]; placebo-ritlecitinib 200/50 mg = [redacted]; placebo-ritlecitinib 50 mg = [redacted]). A higher proportion of patients in the ritlecitinib 50 mg arm were Asian (ritlecitinib 50 mg = [redacted]; placebo-ritlecitinib 200/50 mg = [redacted]; placebo-ritlecitinib 50 mg = [redacted]) and fewer were white compared to both placebo arms (ritlecitinib 50 mg = 60.8%; placebo-ritlecitinib 200/50 mg = [redacted]; placebo-ritlecitinib 50 mg = [redacted]).

AA history was similar across treatment groups.[154] Fewer people in the ritlecitinib 50 mg arm had AU compared to placebo (ritlecitinib 50 mg = [redacted]; placebo-ritlecitinib 200/50 mg = [redacted]; placebo-ritlecitinib 50 mg = [redacted]). The mean (SD) duration since

AA diagnosis across all groups was [redacted] years and was similar across all treatment groups. The mean (SD) duration of the current AA episode was [redacted] years. The distribution of participants by AA severity was balanced across treatment groups, despite fewer people with AU. The mean (SD) baseline SALT score was similar across treatment groups, ranging from [redacted] to [redacted]. A total of [redacted] participants were classified as AT/AU, based on a baseline SALT score of 100%. The distribution of participants with either AT or AU compared to neither AT nor AU was similar across treatment groups. Interviews with UK clinicians confirmed that the trial population was generalisable to patients with AA in the UK likely to receive treatment, with the placebo arm specifically generalisable to BSC defined as non-pharmacological therapy, as discussed in Section B.1.3.3.6.

Table 10: Baseline demographic and disease characteristics of patients enrolled in the ALLEGRO 2b/3 study (B7981015)

Characteristics	Ritlecitinib 200/50 mg (n=132)	Ritlecitinib 200/30 mg (n=130)	Ritlecitinib 50 mg (n=130)	Ritlecitinib 30 mg (n=132)	Ritlecitinib 10 mg (n=63)	Placebo-ritlecitinib 200/50 (n=65)	Placebo-ritlecitinib 50 mg (n=66)	Placebo pooled (n=131)
Age (years), n (%)								
12-17								
≥18								
≥65								
Median (range)								
Sex, n (%)								
Male								
Race, n(%)								
White								
Black								
Asian								
Other								
Multiracial								
Not reported								
Ethnicity, n (%)								
Hispanic/Latino								
Unknown								
Primary diagnosis and durations								
Mean years since diagnosis (SD)								
Mean years since onset of current episode (SD)								
Extent of disease								
Mean SALT score (SD) at baseline								
AT, n (%)								
AU, n (%)								

AT/AU not specified, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-AT/AU, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior pharmacological non-systemic treatment for AA n(%)								
Intralesional corticosteroid injection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other immunotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other topical anti-inflammatory	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SC immunotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Topical JAK	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Topical anthralin/dithranol	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Topical corticosteroid	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Topical immunotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Topical vasodilators	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Unknown steroid	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous exposure to systemic therapies for AA, n (%)[d]								
Biologics	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral anti-inflammatory	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Oral immunosuppressant	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral vasodilator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral/IV/IM steroids	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other non-oral systemic immunosuppressant	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Unknown methotrexate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] *treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants in the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg ritlecitinib. For analysis of endpoints up to Week 24, groups Pbo→200/50 mg and Pbo→50 mg were generalized together and labelled as 'Placebo'.*

[b] *for analysis of endpoints up to Week 24, groups Pbo→200/50 mg and Pbo→50 mg were summarised together and labelled as 'Placebo'.*

[c] *includes all prior non-systemic pharmacological treatments for AA in participants of Study B7981015.*

[d] *includes prior systemic therapies that are reflective of those used in UK clinical practice. The cohort of participants with previous exposure to systemic therapies included in the post-hoc analysis includes participants who received at least one of these systemic treatments.*

Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg ritlecitinib. For analysis of endpoints up to Week 24, groups Pbo→200/50 mg and Pbo→50 mg were generalize together and labelled as 'Placebo'.

Abbreviations: FAS, full analysis set; mg, milligram; n, number of participants; Pbo, placebo

B.2.3.2. ALLEGRO-LT

The methodology for the ALLEGRO-LT study is summarised below in Table 11 and discussed further in the following subsections. The methodology for ALLEGRO-LT has been taken from the study protocol, clinical study report (CSR) and additional sources added as needed.[151–154]

Table 11: Summary of ALLEGRO-LT study methodology

ALLEGRO-LT (B79810132; NCT04006457)	
Study objective	The primary objective of ALLEGRO-LT is to evaluate the long-term safety and tolerability of ritlecitinib in adults and adolescents with AA up to Month 36. The secondary objective is to assess the long-term efficacy of ritlecitinib in adults and adolescents with AA for 36 months.
Study location	148 study locations (including 4 sites in the UK)
Method of randomisation	The study is not randomised.
Method of blinding (care provider, patient, and outcome assessor)	ALLEGRO-LT is open-label.
Key eligibility criteria for participants (full criteria reported in Appendix D)	Inclusion criteria: • Patients ≥ 12 years • Diagnosis of AA with ≥ 25% scalp hair loss due to AA, including AT or AU • No evidence of terminal hair regrowth within 6 months at both screening and baseline visits (this applies to *de novo* patients only) • Maximum duration of current episode of hair loss ≤ 10 years (this applies to *de novo* patients only)
Duration of study	36 months
Trial drugs	• Roll-over participants from ALLEGRO 2a/ALLEGRO 2b/3 were treated with Ritlecitinib 50 mg • *De novo* participated were treated with ritlecitinib 200/50mg
Permitted and disallowed concomitant medications	• None specified in the SAP
Primary outcomes	• Incidence of AEs • SAEs • AEs leading to discontinuation.
Secondary/exploratory outcomes	• Full list presented in Appendix D
Safety assessments performed	Safety assessments were the primary endpoints for the study
Pre-planned subgroups	Not known

Abbreviations: AA, alopecia areata; AAPPO, Alopecia Areata Patient Priority Outcomes; AARU, Alopecia Areata Resource Utilisation; AE, adverse event; AT, alopecia totalis; AU, alopecia universalis; CYP3A, Cytochrome P450, famil3, subfamily A; EBA, Eyebrow assessment; ELA, Eyelash assessment; EU, European Union; FDA, Food and Drug Administration; HADS, Hospital Anxiety and Depression Scale; MCS, mental component score; PGI-C-Patient's Global Impression of Change; PHQ, Patient Health Questionnaire – 8 Items; P-Sat, Patient Satisfaction with Hair Growth; SAE, serious adverse event; SALT, Severity of Alopecia Tool; SF36v2 Acute, 36-Item Short Form Health Survey Version 2 Acute; UK, United Kingdom; WPAI: AA, Work Productivity and Activity Impairment: Alopecia Areata

B.2.3.2.1 Trial outcome definitions

For the primary study objective, safety endpoints included:[156]

- Incidence of AEs
- SAEs
- AEs leading to discontinuation.

For the secondary objective, efficacy endpoints included:

- SALT response based on an absolute SALT score ≤ 20 (i.e., ≤ 20% of the scalp without hair)
- SALT response based on an absolute SALT score ≤ 10 (i.e., ≤ 10% of the scalp without hair)
- PGI-C response, defined as a score of "moderately improved" or "greatly improved"
- The definition of these outcomes were as in the ALLEGRO 2b/3 study.

B.2.3.2.2 Baseline demographics and disease characteristics

At the interim analysis for the *de novo* cohort, the mean age was [redacted] years and approximately [redacted] of patients were adolescent, 12 and 17 years of age. About a [redacted] of the *de novo* patient cohort had AT and/or AU (Table 12). The mean SALT score at baseline in the *de novo* cohort was [redacted].

Table 12: ALLEGRO-LT: Baseline Patient Demographics and Disease Characteristics (*De Novo* Cohort, Interim Analysis)

Characteristic	Ritlecitinib 200/50 mg QD (FAS; N = 449)
Age	
Mean (SD), years	[redacted]
12–17 years, n (%)	[redacted]
≥ 18 years, n (%)	[redacted]
Female, n (%)	[redacted]
White, n (%)	[redacted]
Type of AA, n (%)	
AT/AU*	[redacted]
SALT score among all patients, mean (SD)	[redacted]
Duration of AA since diagnosis, mean (SD), years	[redacted]
Duration of current AA episode, mean (SD), years	[redacted]

Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; FAS, full analysis set; SALT, Severity of Alopecia Tool; SD, standard deviation.
*Notes: *Participants in the AT/AU category had a SALT score of 100% at baseline.*
Source: Sinclair R, et al. (EADV 2022).[156]

The baseline of patients who rolled over from the ALLEGRO 2a or ALLEGRO 2b/3 studies was taken from their enrolment to their originator trials.

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

B.2.4.1. ALLEGRO 2b/3

B.2.4.1.1 Analysis sets

The primary analysis population for efficacy was the full analysis set (FAS), defined as all participants who were randomised regardless of whether they received study intervention

All patients who received the investigational product (ritlecitinib) were included in safety analyses and thus considered in the safety population (Table 13).

Table 13: Summary of analysis sets presented for ALLEGRO 2b/3

Analysis set	Description
FAS	Defined as all participants who were randomised regardless of whether they received study intervention. Participants were analysed in treatment groups as randomised. Analysis set applies to efficacy.
SAS	Defined as all participants who received at least one dose of study intervention, classified according to actual study intervention received for most of the time during the study. Analysis set applies to treatment administration/compliance and safety.

Abbreviations: FAS, full analysis set; SAS, safety analysis set

Table 14 summarises the statistical analysis plan for the ALLEGRO 2b/3 study.

Table 14: Summary of the statistical methodology for ALLEGRO 2b/3

ALLEGRO 2b/3[154]	
Hypothesis objective	There were four key hypotheses: To demonstrate whether 200 mg/50 mg (once daily) dose regimen is superior to placebo for the primary endpoint. To demonstrate whether the 200 mg/30 mg (once daily) dose regimen is superior to placebo for the primary endpoint. To demonstrate whether the 50 mg/50 mg (once daily) dose regimen is superior to placebo for the primary endpoint. To demonstrate whether the 30 mg/30 mg (once daily) dose regimen is superior to placebo for the primary endpoint.
Population definitions	Efficacy analyses were performed using the FAS population
Sample size, power calculation	Total of 660 patients, with 120 per group randomised to either 200 mg/50 mg (once daily), 200 mg/30 mg (once daily), 50 mg/50 mg (once daily), or 30 mg/30 mg (once daily). This would provide more than 90% power to demonstrate that at least the 200 mg/50 mg group is superior to placebo by a difference of 24% in the proportion of subjects achieving the primary endpoint (SALT ≤20 at Week 24), assuming a placebo response rate of no more than 5%, at alpha = 0.05 (2-sided significance level).

Statistical analysis of primary endpoints	The primary endpoint of the study was analysed by Miettinen and Nurminen (MN) method for the difference in the proportion of responders between each active treatment group and placebo.[157] Note the EMA/MHRA, primary endpoints were also analysed by the MN method. In the comparison to placebo, the data from the 2 placebo groups up to Week 24 will be pooled to form one placebo group (see Figure 12)
Statistical analysis of secondary and other endpoints	For secondary endpoints, all binary endpoints were analysed in the same way as the primary endpoint. The study was tested at an overall significance level (α) of 0.05. While this significance level was deemed sufficient for a declaration of effect, more stringent levels have been advised by regulatory agencies for submission of this study to support marketing authorization, namely α=0.01 for the EMA and MHRA, and α=0.00125 for the FDA
Data management, patient withdrawals	Data management was completed by the sponsor. The study used an external data monitoring committee, responsible for ongoing monitoring of the efficacy, safety of patients in the study. Patients were permitted to withdraw from the study at any time at their request, or they were withdrawn at any time at the discretion of the investigator or sponsor for safety or behavioural reasons or the inability of the subject to comply with the protocol-required schedule of study visits or procedures at a given study site.
Missing data	In this submission, missing data is analysed according to the overall study protocol. Patients missing due to COVID-19 were excluded from the analysis. If patients were missing due to other reasons, they were classified as non-responders. For reference for the EMA and MHRA, for primary and key secondary endpoints patients missing due to COVID-19 were classified as missing at random. If patients were missing due to other reasons, they were classified as non-responders

Abbrevations: EMA, Eurpoean Medicines Agency; FAS, full set analsis; FDA, Food and Drug Administration; MN, Miettinen and Nurminen; PDMA, Prescription Drug Marketing Act; VHP, The Voluntary Harmonisation Procedure

Table 15 summarises the different approaches by study or region. In this submission, results are reported from the overall study.

Table 15: Summary of approaches by study or region

Study or region	Endpoint	Analysis designation	Analysis #	Statistical method	Missing due to COVID-19	Missing due to other reasons
Overall study	SALT ≤20 at Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	SALT ≤10 at Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FDA/PMDA	SALT ≤20 at Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EMA and competent authorities in VHP countries	SALT ≤10 at Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	PGI-C response at Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: FDA, Food and Drug Administration; PGI-C, patient global impression of change; PMDA, Pharmaceuticals and Medical Devices Agency; SALT, Severity of Alopecia Tool; VHP, Voluntary Harmonisation Procedure

B.2.4.1.2 Exposure-response analysis

The statistical analysis used to characterise the exposure-response of ritlecitinib on the regrowth of scalp hair according to SALT ≤10 is detailed in Appendix D. The analysis demonstrated that there is a dose-response relationship when considering regrowth of scalp hair. It also demonstrated that a loading dose increased the response to treatment, though similar response is attainable without a loading dose over a longer timeframe. This analysis confirmed the selection of ritlecitinib 50 mg as the dose applied for in the Marketing Authorisation Application. Results for the exposure-response analysis based on SALT ≤20 response were performed post-hoc and were consistent with those described for SALT ≤10.

B.2.4.2. ALLEGRO-LT

The statistical analysis used to evaluate the long-term safety and tolerability of ritlecitinib over time is detailed in Appendix D.

B.2.5. Critical appraisal of the relevant clinical effectiveness evidence

A quality assessment of ALLEGRO 2b/3 and ALLEGRO-LT are presented in Appendix D and have both been determined as high quality overall. As no other studies are included, it is not relevant to do a quality assessment of any other studies.

B.2.6. Clinical effectiveness results of the relevant studies

This section presents data from the full trial populations of ritlecitinib studies for outcomes that are specified in the scope for this appraisal.

In the ALLEGRO 2b/3 trial, 50 mg and 30 mg ritlecitinib doses/regimens, with and without a loading dose of 200 mg for the first four weeks of treatment, demonstrated efficacy in hair regrowth measured by SALT score and PGI-C response at Week 24 compared to placebo. The 48-week treatment period consisted of a 24-week placebo-controlled period and a 24-week extension phase where placebo-treated patients switched to active treatment with ritlecitinib in a pre-specified, blinded manner, while other arms continued on the same maintenance dose. As such, comparative efficacy data is only available up to week 24. In addition, given there is no difference between the 50mg dose and the 200/50mg dose by Week 48, we can conclude that induction dose did not have any impact on the long-term efficacy.

The 50 mg dose has been selected as the dose for registrations with regulatory authorities and is the focus of the results in this NICE submission.

Summary of key efficacy data

- ALLEGRO 2b/3 met its primary endpoint: response based on SALT score ≤20 was statistically significantly higher in the ritlecitinib 50 mg treatment group [redacted] than placebo [redacted] at Week 24 ($P < 0.001$).
- Through to Week 48, the response rates based on SALT ≤20 in the 50 mg treatment group [redacted] continued to improve compared to Week 24
- The secondary endpoint of response rate based on SALT score ≤10 was statistically significantly higher in the ritlecitinib 50 mg treatment group [redacted] than placebo [redacted] at Week 24 ($P < 0.001$).

- The secondary endpoint of PGI-C response was statistically significantly higher in the ritlecitinib 50 mg treatment group [redacted] than placebo [redacted] at Week 24 ($P < 0.001$).
- ALLEGRO-LT demonstrated a similar trajectory to the ALLEGRO 2b/3 study in the proportion of patients achieving SALT scores of ≤20 and ≤10 over time, though in the absence of placebo-control the significance of this finding cannot be determined.

B.2.6.1. ALLEGRO 2b/3

B.2.6.1.1 Efficacy endpoints

Response based on absolute SALT through Week 48

Ritlecitinib administered at a dose of 50 mg (once daily) met the primary study endpoint of significantly more patients achieving SALT ≤20 compared to placebo. This regimen resulted in both clinically meaningful and statistically significant scalp hair regrowth at Week 24 [redacted] compared with placebo ([redacted], $P < 0.001$, Table 16).[17,154] Through to Week 48, the response rates based on SALT ≤20 in the 50 mg treatment group [redacted] continued to improve compared to Week 24, though there is no placebo-control against which to determine the significance of this result.

Table 16: Summary of response based on SALT score ≤20 and ≤10 at Week 24 and Week 48 (FAS)

Cohort[a]	Ritlecitinib 200/50 mg (N=132)	Ritlecitinib 200/30 mg (N=130)	Ritlecitinib 50 mg (N=130)	Ritlecitinib 30 mg (N=132)	Ritlecitinib 10 mg (N=63)	Placebo (N=131)
SALT score ≤20 up to Week 24[b]						
Participants with SALT ≤20 Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Estimated response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference from placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value	<0.001 [c]	<0.001 [c]	<0.001 [c]	<0.001 [c]	0.963 [c]	-
SALT score ≤20 up to Week 48 (Please note patients on placebo switched to active treatment at week 24)						
Estimated response rate, n/N (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SALT score ≤10 up to Week 24[b]						
Participants with SALT ≤10 Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Estimated response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference from placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value	[redacted]	[redacted]	[redacted]	[redacted]	-	-
SALT score ≤10 up to Week 48 (Please note patients on placebo switched to active treatment at week 24)						
Estimated response rate, n/N (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose.

[b] a generalised linear mixed effect model without imputation using observed data up to Week 24 was used as the imputation model, assuming MAR under a Bayesian framework with non-informative/weakly informative prior densities and Markov chain Monte Carlo. For a participant with missing SALT score at Week 24 due to any reason, imputation was done based on predictive Bernoulli distribution with a probability equal to the probability under MAR calculated using the sampled parameters. MN method was used for the calculation of 95% Cis and p-values for testing the difference in the proportion of response between each treatment group and placebo.

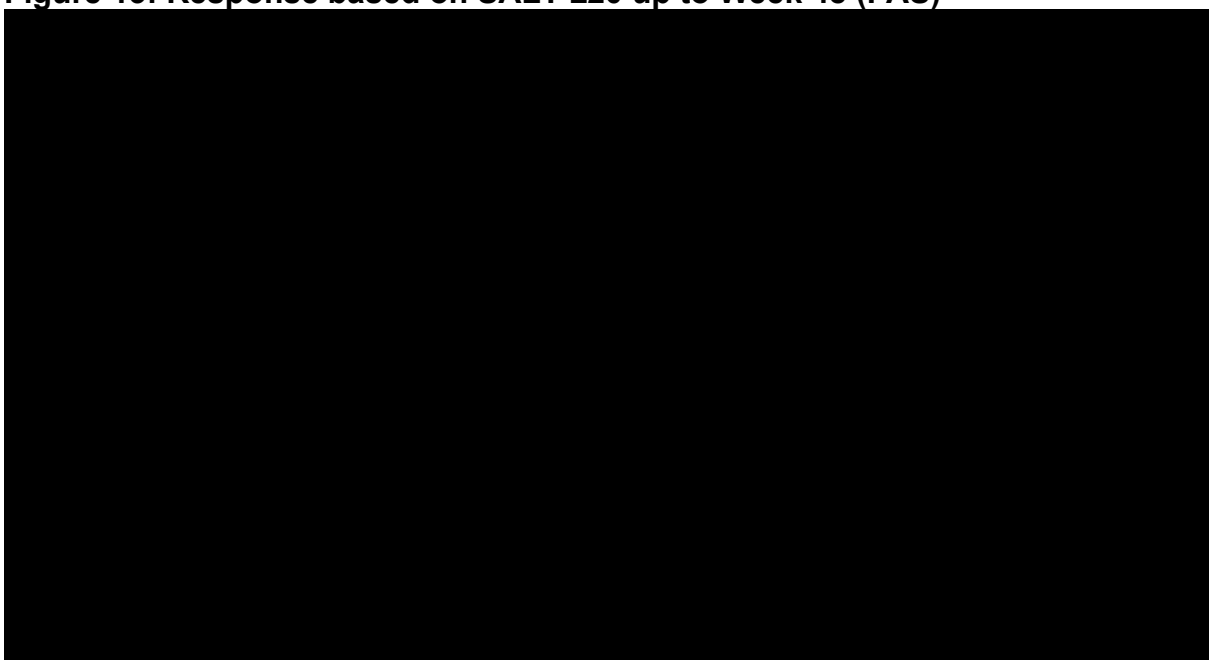
[c] confidence interval and p-value are calculated using MN method. Missing data due to COVID-19 was excluded from this analysis, whereas participants with missing data due to other reasons were considered as non-responders.

Abbreviations: CI, confidence interval; FAS, full analysis set; MAR, missing at random; mg, milligrams; MN, Miettinen and Nurminen; n, number of participants with SALT ≤20 at Week 24/48 per group; SALT, Severity of Alopecia Tool

The difference in the response rate for SALT ≤20 showed a statistically significant difference between the ritlecitinib and placebo groups as early as Week 12 for the 200/50 mg group ($P = 0.001$) and Week 18 for the 200/50 mg and 50 mg groups ($P < 0.001$ and $P < 0.001$, respectively).

Across treatment groups, [redacted] participants [redacted] missed doses (range: [redacted]) up to Week 24 due to COVID-19, for a mean of [redacted] days. These participants were excluded from the response analysis based on SALT ≤20 at Week 24 (primary endpoint for the overall study; Table 8). Patients with missing data due to other reasons were counted as non-responders. Tipping point analysis assessed the impact of missing SALT scores related to COVID-19 on the conclusion of the primary analysis, which used multiple imputation for missing data due to COVID-19 and considered missing data due to other reasons as non-responders. The results of all explored scenarios supported the conclusion of the primary analysis in all the ritlecitinib and placebo groups and are presented in Appendix D.

Figure 13: Response based on SALT ≤20 up to Week 48 (FAS)



Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants in the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg.
Abbreviations: FAS, full analysis set; mg, milligram; Pbo, placebo; SALT, Severity of Alopecia Tool

For patients who were switched from placebo to ritlecitinib at Week 24, SALT ≤20 response rates at Week 48 were similar to that of active treatment participants on the same regimen at Week 24 (Pbo→200/50 mg at Week 48 [redacted] vs 200/50 mg at Week 24 [redacted]; Pbo→50 mg at Week 48 [redacted] vs 50 mg at Week 24 [redacted]).

Ritlecitinib 50 mg met the key secondary endpoint of a response based on SALT ≤10 at Week 24, i.e., the more patients achieved a SALT score of 10 or lower at Week 24 (redacted) compared with placebo (redacted). Through Week 48, response rates based on SALT ≤20 followed a similar trend to that of SALT ≤10 although the absolute response rates were lower for SALT ≤10 than SALT ≤20, which was expected given that the SALT ≤10 response reflects greater hair regrowth, which is a more stringent endpoint compared to the SALT ≤20 response. Up to week 24, the estimated response rate for patients on ritlecitinib 50mg was redacted and at week 48, the response rate increased to redacted.

As with the primary endpoint, (response based on SALT ≤20) tipping point analysis assessed the impact of missing SALT scores related to COVID-19. The results of all explored scenarios supported the conclusion of the primary analysis in all the ritlecitinib and placebo groups. The results of this analysis are also presented in Appendix D.

Through to Week 48, response rates based on SALT ≤10 followed a similar trend to that of SALT ≤20. In the 50 mg group, redacted out of 125 participants (redacted; 95% CI redacted, redacted) had achieved SALT ≤10 up to Week 48 (Table 16 and Figure 14).

Figure 14: Response based on SALT ≤10 up to Week 48 (FAS)

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization.
Abbreviations: FAS, full analysis set; mg, milligram; Pbo, placebo; SALT, Severity of Alopecia Tool

Change in SALT scores up to Week 48 from baseline

The least-square mean (LSM) change from baseline in SALT score improved (became more negative) from Week 4 to Week 24 in all of the ritlecitinib treatment groups, achieving statistical significance compared to placebo by Week 8 in the 50 mg group ($P = 0.012$), as well as in the groups that received a loading dose (200/50 mg, $P < 0.001$; 200/30 mg, $P < 0.001$) (Figure 15).

Figure 15: Least squared means of absolute change from baseline in SALT score for initial active groups up to Week 48 (FAS)

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization.
Abbreviations: FAS, full analysis set; mg, milligram; Pbo, placebo; SALT, Severity of Alopecia Tool

Up to Week 24, the estimated LSM change from baseline in SALT score was numerically greater in participants who had received a 200 mg loading dose for four weeks than in participants treated for 24 weeks with the same maintenance dose but without a loading dose (200/50 mg ( ) vs 50 mg (.)]; 200/30 mg ( ) vs 30 mg ( )) (Figure 15). The LSM change from baseline was significantly greater ($P < 0.001$) in these groups compared to placebo ( ). By Week 48, the estimated LSM change from baseline in SALT scores was similar between the 200/50 mg and 50 mg groups (vs , respectively), and the 200/30 mg and 30 mg groups (- vs , respectively).

Patient's global impression of change response

PGI-C, a key secondary outcome, is a single-item measure that evaluates whether there has been a global improvement or worsening in AA compared to the start of the study. PGI-C response is defined as a score of 'moderately improved' or 'greatly improved' compared to baseline. Based on a pre-specified analysis, the ritlecitinib 50 mg group ( ) demonstrated a significantly higher proportion of participants with

a PGI-C response at Week 24 compared to the placebo cohort ([redacted]) (Table 17). The PGI-C response rates increased across all treatment groups through Week 48 and were highest in the 50 mg dose groups with and without loading dose (Table 17 and Figure 16).

Table 17: PGI-C response at Week 24 and 48 (FAS)

Treatment group[a]	200/50 mg (n=132)	200/30 mg (n=130)	50 mg (n=130)	30 mg (n=132)	10 mg (n=63)	Pbo→ 200/ 50 mg (n=65)	Pbo→ 50 mg (n=66)
Week 24							
Participants with PGI-C Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Difference from placebo (95% CI)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
p-value	<0.001	<0.001	<0.001	<0.001	-	-	
Week 48							
n/N (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI[c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization. For analysis of endpoints up to Week 24, groups Pbo→200/50 mg and Pbo→50 mg were generalize together and labelled as 'Placebo'.

[b] a generalized linear mixed effect model without imputation using observed data up to Week 24 was used as the imputation model. Estimation of model parameters was performed assuming MAR under a Bayesian framework with non-informative/weakly informative prior densities and Markov chain Monte Carlo. For a participant with missing SALT score at Week 24 due to COVID-19 related reasons, imputation was done based on predictive Bernoulli distribution with a probability equal to the probability under MAR calculated using the sampled parameters. Participants with missing SALT score at Week 24 due to reasons other than COVID-19 were considered non-responders. A single complete imputed data set for Week 24 was then analysed using the Miettinen method as the analysis model.

Abbreviations: CI, confidence interval; FAS, full analysis set; MAR, missing at random; mg, milligrams; MN, Miettinen and Nurminen; n, number of participants with PGI-C response at Week 24 per group; Pbo, placebo; PGI-C, Patient's Global Impression of Change; SALT, Severity of Alopecia ToolConfidence

[c] Interval is calculated based on normal approximation

In addition to the results presented in Table 17 and Figure 16, post-hoc sub-analyses of PGI-C results are presented in B.2.6.1.2.

Figure 16: PGI-C response to Week 48 (FAS)

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization.
Abbreviations: FAS, full analysis set; mg, milligram; Pbo, placebo; PGI-C, Patient's Global Impression of Change

Eyebrow and eyelash assessment scores up to Week 48

Participants without normal eyebrow assessment (EBA) or eyelash assessment (ELA) scores at baseline were included in an analysis of response based on EBA and ELA. The proportion of participants with EBA and ELA response (defined as at least a 2-grade improvement from baseline or a normal EBA and ELA score, respectively) increased over time (Figure 17 and Figure 18, respectively).[154] A similar pattern was observed in the proportion of ELA response and EBA response from Week 4 to Week 48.

At Week 24, the proportions of participants with an EBA response were clinically meaningful and higher than the proportion of participants in the placebo group placebo group ([ ] in the 50 mg group compared to [ ] in the placebo group), P < 0.001 [Table 18]).

Table 18: Participants with an EBA and ELA response at Week 24 and 48 (FAS)

Treatment group[a]	200/50 mg (n=132)	200/30 mg (n=130)	50 mg (n=130)	30 mg (n=132)	10 mg (n=63)	Pbo→ 200/ 50 mg (n=65)	Pbo→ 50 mg (n=66)
EBA Week 24							
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference from placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value	<0.001	<0.001	<0.01	0.005	0.368	-	-
EBA Week 48							
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ELA Week 24							
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference from placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
p-value	<0.001	0.001	<0.001	<0.01	0.946	-	-
ELA Week 48							
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization. For analysis of endpoints up to Week 24, groups Pbo→200/50 mg and Pbo→50 mg were generalize together and labelled as 'Placebo'.

[b] a generalized linear mixed effect model without imputation using observed data up to Week 24 was used as the imputation model. Estimation of model parameters was performed assuming MAR under a Bayesian framework with non-informative/weakly informative prior densities and Markov chain Monte Carlo. For a participant with missing SALT score at Week 24 due to COVID-19 related reasons, imputation was done based on predictive Bernoulli distribution with a probability equal to the probability under MAR calculated using the sampled parameters. Participants with missing SALT score at Week 24 due to reasons other than COVID-19 were considered non-responders. A single complete imputed data set for Week 24 was then analysed using the Miettinen method as the analysis model.

Abbreviations: CI, confidence interval; EBA, eyebrow assessment; FAS, full analysis set; MAR, missing at random; mg, milligrams; MN, Miettinen and Nurminen; N/A, not applicable; Pbo, placebo; SALT, Severity of Alopecia Tool

Figure 17: Response based on at least a 2-grade improvement from baseline or a normal EBA score, up to Week 48 (participants without normal EBA at baseline in FAS)

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization. Abbreviations: EBA, eyebrow assessment; FAS, full analysis set; mg, milligram; S/E, standard error

The proportions of participants with an ELA response at Week 24 were also clinically meaningful and higher than the proportion of participants in the placebo group ([ ]) in the 50 mg group ([ ], P < 0.001).

Figure 18: Response based on at least a 2-grade improvement from baseline or a normal ELA score, up to Week 48 (participants without normal ELA at baseline in FAS)

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization.
Abbreviations: ELA, eyelash assessment; FAS, full analysis set; mg, milligram; OD, once daily, SALT, Severity of Alopecia Tool; S/E, standard error

Exploratory endpoint: P-sat items

Up to Week 24, an improvement was reported versus placebo in the 50 mg group for satisfaction with amount of hair grown ( ), satisfaction with overall hair grown back (ritlecitinib 50 mg:) and satisfaction with quality of new hair (ritlecitinib 50 mg:). The proportion of participants showing improvement in all three aspects of P-Sat was numerically higher in all active treatment groups (Table 19).[154] In groups that received ritlecitinib from the start of the trial, the levels of satisfaction in each P-Sat item were improved from Week 24 to Week 48. In participants who had first received placebo (Pbo→200/50 mg; Pbo→50 mg), the largest increase in each aspect of the P-Sat was seen in the first weeks after switching to treatment with little additional change thereafter up to Week 48.

Table 19: Response based on improvement in P-Sat (FAS)

Treatment group[a]	200/50 mg (n=132)	200/30 mg (n=130)	50 mg (n=130)	30 mg (n=132)	10 mg (n=63)	Pbo→ 200/50 mg (n=65)	Pbo→ 50 mg (n=66)
Satisfaction with amount of hair grown							
Week 24 n/N1 (% [95% CI])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 48 n/N1 (% [95% CI])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Satisfaction with overall hair grown back							
Week 24 n/N1 (% [95% CI])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 48 n/N1 (% [95% CI])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Satisfaction with quality of new hair							
Week 24 n/N1 (% [95% CI])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 48 n/N1 (% [95% CI])	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose. Pbo groups received Pbo until Week 24, after which they were re-randomised to receive either a 200 mg ritlecitinib OD for four weeks, followed by 50 mg ritlecitinib OD for the remaining 20 weeks, or 50 mg ritlecitinib OD for 24 weeks

Abbreviations: CI, confidence interval; FAS, full analysis set; mg, milligrams; n, number of participants with patient satisfaction with hair growth response as 'slightly', 'moderately', or 'very satisfied'; N1, number of participants with valid data at Week 24 per group (excluding participants who missed doses of ritlecitinib due to COVID-19); Pbo, placebo; P-Sat, Patient Satisfaction with Hair Growth

B.2.6.1.2 Correlation between SALT score and patients' perception of hair regrowth

In a pre-specified analysis of data collected from the ALLEGRO 2a proof of concept study, absolute SALT and PGI-C scores were strongly correlated at Weeks 24 and 48. In a similar, post-hoc analysis of patient-reported satisfaction with hair regrowth (at Week 24 and 48), strong correlations were observed for satisfaction with quality and amount of hair regrowth, in addition to overall satisfaction. Taken together, these results indicate that an improvement in clinician-assessed efficacy measures from ritlecitinib treatment is associated with an increase in patients' perception of improvement and satisfaction with hair regrowth.

Full detail of the analyses of the correlation of SALT response with PGI-C and P-Sat are presented below.

SALT score by PGI-C score

The mean SALT scores by PGI-C score at Week 24 are presented in Table 20. Patients whose PGI-C was greatly improved had a mean SALT score of [redacted] whereas patients with a PGI-C of slightly improved, not changed or worsened had a mean SALT score of [redacted].

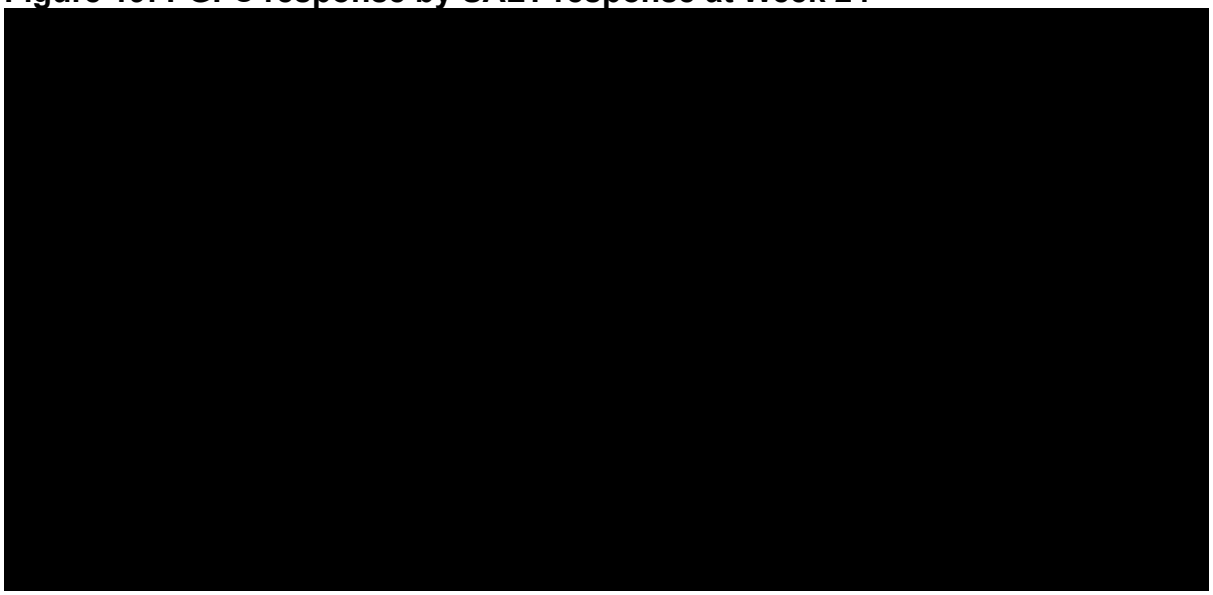
Table 20: ALLEGRO 2b/3 mean SALT Scores by PGI-C score at Week 24

	Greatly improved	Moderately improved	Slightly improved	Not changed	Slightly worsened	Moderately worsened	Greatly worsened
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean SALT score (SD)							

Abbreviations: PGI-C, Patient's Global Impression of Change; SALT, Severity of Alopecia Tool; SD, standard deviation

Of all patients who achieved SALT ≤20 in the ALLEGRO 2b/3 study by Week 24, [redacted] were PGI-C responders (determined by a response of "greatly improved" or "moderately improved"). Conversely, just [redacted] of patients who did not achieve a score of SALT 20 were PGI-C responders (Figure 19). The correlation between PGI-C response and SALT 20 response at Week 24 score demonstrated a [redacted] correlation (Spearman's correlation coefficient r=[redacted]). This demonstrates that patients with lower SALT scores felt the greatest improvements in their condition.

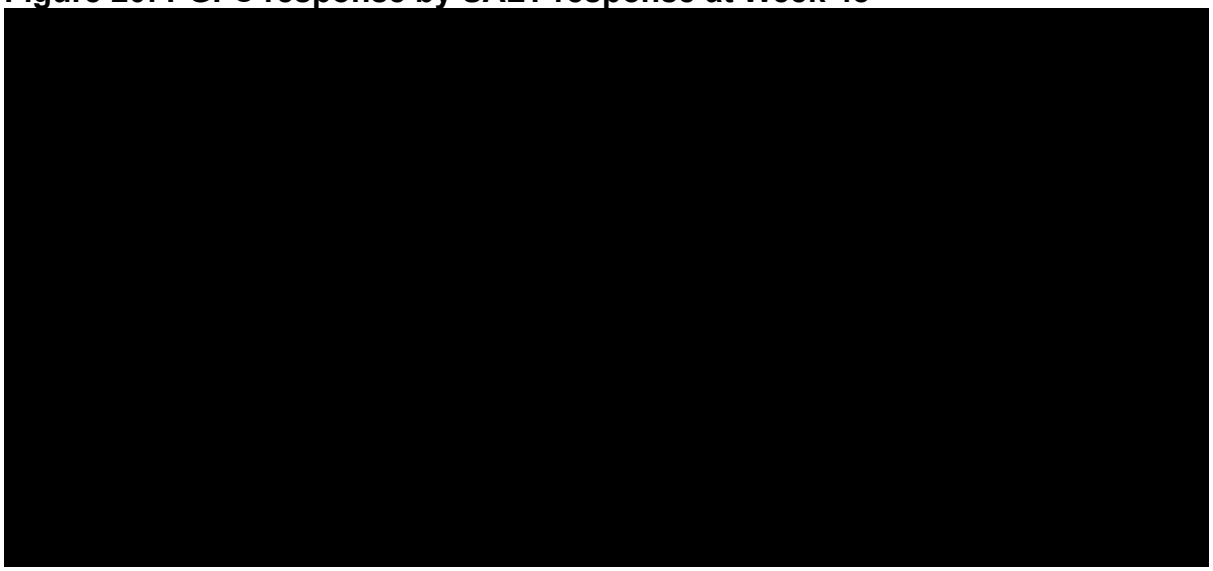
Figure 19: PGI-C response by SALT response at Week 24



Abbreviations: PGI-C, Patient's Global Impression of change; SALT, Severity of Alopecia Tool.
Source: Law E. et al (EADV 2022).[158]

Similarly, of all patients who achieved SALT ≤20 by Week 48, [redacted] were PGI-C responders (determined by a response of "greatly improved" or "moderately improved"). Conversely, just [redacted] of patients who did not achieve a score of SALT 20 were PGI-C responders (Figure 20). The correlation between PGI-C response and SALT 20 response at Week 48 score demonstrated a [redacted] correlation (Spearman's correlation coefficient r=[redacted]). This demonstrates that patients with lower SALT scores felt the greatest improvements in their condition.

Figure 20: PGI-C response by SALT response at Week 48



Abbreviations: PGI-C, Patient's Global Impression of change; SALT, Severity of Alopecia Tool.
Source: Law E. et al (EADV 2022).[158]

SALT score by P-Sat score

The mean SALT score by P-Sat score for overall satisfaction with hair at Week 24 are presented in Table 21. Patients who reported being very satisfied had a mean SALT score of [redacted] whereas patients who reported being slightly satisfied, neither satisfied or dissatisfied, or dissatisfied had a mean SALT score ≥70.

Table 21: ALLEGRO 2b/3 Mean SALT Scores by P-Sat (overall satisfaction with hair) at Week 24

	Very satisfied	Moderately satisfied	Slightly satisfied	Neither satisfied or dissatisfied	Slightly dissatisfied	Moderately dissatisfied	Very dissatisfied
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean SALT score (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: P-Sat, Patient Satisfaction with Hair Growth; SALT, Severity of Alopecia Tool; SD, standard deviation

Of all patients who achieved SALT ≤20 (responders) in the ALLEGRO 2b/3 study by Week 24, a [redacted] of patient reported satisfaction (i.e., “very satisfied”, “moderately satisfied”, or “slightly satisfied”) versus non-responders (Figure 21) for amount of hair regrowth ([redacted] versus [redacted], respectively), quality of hair regrowth ([redacted] versus [redacted], respectively) and overall satisfaction of hair regrowth ([redacted] versus [redacted], respectively).

Figure 21: P-Sat response by SALT response at Week 24

Abbreviations: PGI-C, Patient's Global Impression of change; SALT, Severity of Alopecia Tool.
Source: Law E. et al (EADV 2022).[158]

A [redacted] proportion of patients who achieved SALT ≤20 (responders) in the ALLEGRO 2b/3 study by Week 48 reported satisfaction (i.e., "very satisfied", "moderately satisfied", or "slightly satisfied") versus non-responders (Figure 22) for amount of hair regrowth ([redacted] versus [redacted], respectively), quality of hair regrowth([redacted] versus [redacted], respectively), and overall satisfaction of hair regrowth([redacted] versus [redacted], respectively).

Figure 22: P-Sat response by SALT response at Week 48

Abbreviations: PGI-C, Patient's Global Impression of change; SALT, Severity of Alopecia Tool.
Source: Law E. et al (EADV 2022).[158]

All P-Sat scores were [redacted] correlated with SALT score change from baseline at Week 24 (correlation coefficient range, [redacted] to [redacted], $P < 0.05$) and Week 48 (correlation coefficient range, [redacted] to [redacted], $P < 0.05$).

Table 22. ALLEGRO 2b/3: Correlation Between P-Sat Domains and Change From Baseline in SALT Scores

Measure	P-Sat Domains, Correlation Coefficient (95% CI) Satisfaction With:			P Value
	Overall Hair Regrowth	Amount of Hair Regrowth	Quality of Hair Regrowth	
Change from Baseline SALT At Week 24 (n = 647)	[redacted]	[redacted]	[redacted]	< 0.05
Change from Baseline SALT At Week 48 (n = 621)	[redacted]	[redacted]	[redacted]	< 0.05

Abbreviations: CI, confidence interval; P-Sat, Patient Satisfaction With Hair Growth; SALT, Severity of Alopecia Tool.
Source: Sinclair R, et al (EADV 2022).[156]

B.2.6.2. ALLEGRO-LT

The dose for the *de novo* cohort was specified before the ALLEGRO 2b/3 study was complete, and as such prior to the determination of the dose of ritlecitinib that will be used in clinical practice. Despite this, the efficacy results from the *de novo* cohort are presented below, demonstrating a similar trajectory in proportions of patients treated with ritlecitinib attaining SALT ≤20 over time as to the ALLEGRO 2b/3 study. The results demonstrate that the vast majority of patients who achieve SALT ≤20 do so by Week 48 (Month 12), with small numbers of patients going on to do so with further treatment.

SALT response

For the interim analysis, response rates based on SALT ≤20 (Figure 23) and SALT ≤10 (Figure 24) revealed sustained, long-term efficacy with a continued increase over time to Month 24 in both SALT score measures.

Figure 23: ALLEGRO-LT: Response Based on SALT ≤ 20 up to Month 24 (Interim Analysis *De Novo* Cohort)

Abbreviations: CI, confidence interval; QD, once daily; SALT, Severity of Alopecia Tool.
Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤20. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.
Source: Sinclair R, et al. (EADV 2022).[156]

Figure 24: ALLEGRO-LT: Response Based on SALT ≤ 10 up to Month 24 (Interim Analysis *De Novo* Cohort)

Abbreviations: CI, confidence interval; QD, once daily; SALT, Severity of Alopecia Tool.
Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤10. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.
Source: Sinclair R, et al. (EADV 2022).[156]

PGI-C response

For the interim analysis, patient-reported rates of improvement in AA as measured by the PGI-C increased over time to Month 24 (Figure 25).[291]

Figure 25: ALLEGRO-LT: PGI-C Response up to Month 24 (Interim Analysis *De Novo* Cohort)

Abbreviations: CI, confidence interval; PGI-C, Patient Global Impression of Change; QD, once daily.
Note: PGI-C response is defined as "moderately improved" or "greatly improved". n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving a PGI-C response.
The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.
Source: Sinclair R, et al. (EADV 2022).[156]

B.2.7. Subgroup analysis

Forest plots for all the subgroups that were pre-specified in the trial are presented in Appendix E. The differences between each ritlecitinib group and placebo in the proportion of response based on SALT ≤20 at Week 24 [redacted] across most pre-specified subgroups (age, BMI, weight, gender, race, region, severity of disease, duration since diagnosis, duration of current episode, prior pharmacological treatment for AA) for all doses.[154]

There were some exceptions to this:

- Race: For analysis based on race, the numbers of participants who identified as 'Black' or 'Other' were [redacted]
- AA severity: the proportion of patients with AT/AU who achieved SALT ≤10, SALT ≤20 and PGI-C was [redacted] than in those with non-AT/AU across all ritlecitinib treatment groups.

No subgroup analyses were considered in the economic evaluation [redacted]

B.2.7.1. Age subgroups

When considering subgroups by age, the difference in the proportion of responders at Week 24 between ritlecitinib and placebo was [redacted] in adolescents and adults for SALT ≤20 and ≤10 and PGI-C for each treatment group, as seen in Table 23

Table 23 Response based on SALT ≤20 and ≤10 by age group (FAS)

Treatment group[a]	200/50 mg (n=60)	200/30 mg (n=60)	50 mg (n=60)	30 mg (n=61)	10 mg (n=29)	Pbo (n=131)
SALT ≤10						
12-17						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
≥18						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
18-44						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
45-64						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥65						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SALT ≤20						
12-17						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥18						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18-44						

n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
45-64						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥65						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose
Abbreviations: CI, confidence interval; FAS, full analysis set; mg, milligrams; NR, not reported; n, number of participants at Week 24 per group; N1, number of participants with valid data at Week 24 per group (excluding participants who missed doses of ritlecitinib due to COVID-19); PGI-C, Patient's Global Impression of Change; SALT, Severity of Alopecia Tool

B.2.7.2. Participants with AT/AU

For participants with AT/AU, a significant difference was observed in the proportion of responders in the 50 mg group, as well as the 200/50 mg, 200/30 mg and 30 mg groups compared to placebo based on SALT ≤20 at Week 24 ($P < 0.05$) (Table 24).[154] At Week 48 the SALT ≤20 response rate in the 50 mg group ([redacted]) was higher than that in the 200/50 mg ([redacted]), while the response rate in the 30 mg group ([redacted]) was lower than that in the 200/30 mg group ([redacted]). For response based on SALT ≤10 at Week 24, a significant difference was observed in the AT/AU cohort between the ritlecitinib and placebo groups for the groups that received the loading dose i.e., the 200/50 mg and 200/30 mg groups ($P < 0.05$) (Table 24).[154]

The proportion of participants with PGI-C response at Week 24 between ritlecitinib and placebo was [redacted] in participants with AT/AU and the whole cohort (and was statistically significant versus placebo for the 200/50 mg, 200/30 mg, 50 mg and 30 mg groups.[154]

Table 24: Response of AT/AU patients based on SALT ≤10, based on SALT ≤20 and PGI-C response at Week 24 (FAS)

Treatment group[a]	200/50 mg (n=60)	200/30 mg (n=60)	50 mg (n=60)	30 mg (n=61)	10 mg (n=29)	Pbo→ 200/50 mg (n=32)	Pbo→ 50 mg (n=28)

Response based on SALT ≤10						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response based on SALT ≤20						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PGI-C response						
n/N1 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference vs placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose
Abbreviations: CI, confidence interval; FAS, full analysis set; mg, milligrams; NR, not reported; n, number of participants at Week 24 per group; N1, number of participants with valid data at Week 24 per group (excluding participants who missed doses of ritlecitinib due to COVID-19); PGI-C, Patient's Global Impression of Change; SALT, Severity of Alopecia Tool

The SALT ≤20 response rates with ritlecitinib treatment were consistently [redacted] in participants with AT/AU than in those without AT/AU. For four of the ritlecitinib doses (200/50 mg, 200/30 mg, 50 mg and 30 mg), the 95% CI for the differences in proportions of responders based on SALT ≤20 at Week 24 between ritlecitinib and placebo excluded zero in both subgroups, indicating a treatment effect in both AT/AU and non-AT/AU participants at a nominal significance level of α=0.05.

B.2.8. Meta-analysis

A meta-analysis was not conducted as the ALLEGRO 2a study had an initial four week induction dose of 200 mg ritlecitinib (once daily) for all ritlecitinib treated patients, meaning the data from the study does not represent patients treated with the licensed dose of ritlecitinib and it is not suitable for a meta-analysis with the ALLEGRO 2b/3 study.[154,159]

B.2.9. Indirect and mixed treatment comparisons

The Therapeutic Landscape Delphi panel supports the position of ritlecitinib as a comparator to BSC defined as non-pharmacological therapy. Since it was confirmed by expert opinion that the placebo arm of ALLEGRO 2b/3 is generalisable to BSC defined as non-pharmacological therapy, meaning the most relevant efficacy and

safety data to this appraisal are provided from the ALLEGRO 2b/3 study. Therefore, an indirect treatment comparison was not conducted.

Adverse reactions

The safety and tolerability of ritlecitinib in the treatment of participants with severe AA has been evaluated in two studies; the ALLEGRO 2b/3 and 2a study and support the following findings:[32,154]

- Short-and long-term use of ritlecitinib was well tolerated.
- Most adverse events were mild, self-limited, and seldom required interruption or permanent discontinuation of therapy. The most common TEAEs (reported in ≥5%) were nasopharyngitis, headache, acne, diarrhoea, upper respiratory tract infection, nausea, folliculitis, and AD.
- The incidence of serious adverse events (SAE) was low. In total, 14 participants experienced serious AEs, which were generally balanced across treatment groups.
- The adverse event and laboratory profiles suggest that there are no risks unique to the adolescent population.

B.2.10. Adverse reactions

B.2.10.1. ALLEGRO 2b/3

In this section we present the safety data from the pivotal Phase 2b/3 trial. This data is consistent with data from the Phase 2a trial.

B.2.10.1.1 Adverse event definitions

Treatment emergent adverse events (TEAE) are defined as any untoward medical occurrence which emerged or worsened during the treatment period but these were not necessarily causally related to treatment unless classified as treatment-related.

A serious adverse event (SAE) is any untoward medical occurrence at any dose that results in death, is life-threatening (immediate risk of death), requires inpatient hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability/incapacity (substantial disruption of the ability to conduct normal

life functions), results in congenital anomaly/birth defect, and is an important medical event based on investigator's judgment.

The investigators assessment of AE intensity is classified as mild, moderate or severe. A mild AE does not interfere with patient's usual function whereas a moderate AE interferes to some extent. A severe AE is one that interferes significantly with a patient's usual function although it is not necessarily an SAE. For example, a headache may be severe (interferes significantly with the patient's usual function) but would not be classified as serious unless it met one of the criteria for SAEs.

B.2.10.1.2 Duration of follow-up

A total of 1,097 participants were screened. There were 379 screens failures, and 718 participants were enrolled at 118 sites in 18 countries. A safety analysis set (SAS) consisting of all participants with severe AA who received at least one dose of study treatment were evaluated (n=715).

Of the participants who were randomised, [redacted] ([redacted]) discontinued treatment and a total of [redacted] ([redacted]) completed the study. The proportion of participants who completed the study was similar between treatment groups, except for the ritlecitinib 30 mg arm, where [redacted] of participants completed the study compared to [redacted] in the remaining ritlecitinib arms. In the placebo arm, [redacted] of participants completed the study.

The most common reasons for discontinuation across all groups were "Withdrawal by participant", "Adverse event", "Lost to follow-up", "Lack of efficacy", and "Physician decision". There were [redacted] ([redacted]) discontinuations due to COVID-19. Discontinuation due to COVID-19 was similar across treatment groups and the most common reasons for discontinuation due to COVID-19-related reasons were "Withdrawal by participant" and "Physician decision". No impact was observed for COVID-19-related illness or other pandemic-related reason on the overall intended population being treated.

B.2.10.1.3 Safety overview

The safety and tolerability of ritlecitinib in the treatment of participants with severe AA has been evaluated in two studies; the ALLEGRO 2b/3 and 2a study.[32,154] In the studies, the incidence of AEs was reported including TEAEs and TRAEs. Ritlecitinib

at all doses studied, including once daily 50 mg, was well tolerated up to Week 48 in the patient population included in the ALLEGRO 2b/3 study and the safety profile of ritlecitinib in the study was consistent with that observed in previous studies in healthy volunteers and in patients with AA.[154] The incidence of all AEs was also similar between the 50 mg treatment group and the placebo groups at Week 24.[154]

As summarised in Table 25 in ALLEGRO 2b/3, the proportion of participants who experienced TEAEs was [redacted] across treatment groups through the placebo-controlled period (up to Week 24) and up to Week 48 .[19] Up to Week 24, the percentage of participants with TEAEs across treatment groups ranged from [redacted] in the 10 mg group to [redacted] in the 50 mg group and were similar to the placebo group ([redacted]). In the entire study (up to Week 48 and including the follow up period) the percentage of participants with TEAEs was [redacted] in the 50 mg group and ranged from [redacted] in the 10 mg group to [redacted] in the placebo to 50 mg (Pbo→50 mg) group. Most TEAEs ([redacted]) throughout the 48-week study period were mild to moderate in severity, severe adverse events occurred in between [redacted] (50 mg) and [redacted] (200/30 mg) of participants across all groups. The incidence of TEAEs did not appear to be dose dependent across treatment groups.[19] .

Table 25: Overall summary of TEAEs (SAS)

Treatment group[a]	200/ 50 mg (n=131)	200/ 30 mg (n=129)	50 mg (n=130)	30 mg (n=132)	10 mg (n=62)	Pbo→ 200 /50 mg (n=65)	Pbo→ 50 mg (n=66)
TEAEs up to Week 24, n (%)							
Number of AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Participants with AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Participants with serious AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Participants with severe AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
TEAE up to Week 48, n (%)							
Number of AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Participants with AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Participants with serious AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Participants with severe AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] *treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants*

in the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg ritlecitinib.
Abbreviations: AE, adverse event; mg, milligram; n, number of participants; Pbo, placebo; SAS, safety analysis set; TEAE, treatment-emergent adverse event

The incidence rates of serious TEAEs were also similar before and during the COVID-19 pandemic ([redacted] and [redacted] per 1,000 patient-years, respectively). However, the incidence rate of all TEAEs was higher before compared to during the COVID-19 pandemic when factoring in exposure in patient-years ([redacted] vs [redacted] events per 1,000 patient-years of exposure, respectively).

B.2.10.1.4 Commonly reported treatment-emergent adverse events

The TEAEs that occurred in ≥5% of participants in any group up to Week 24 and Week 48, and their incidence rates, are reported in Table 26.[154] In the 50 mg group to Week 24, the most commonly reported TEAEs were nasopharyngitis, headache, and diarrhoea and to Week 48 nasopharyngitis, headache, diarrhoea, and acne. Across all treatment groups the most frequently reported TEAEs at both Week 24 and Week 48 included nasopharyngitis, headache, and upper respiratory tract infection.[154] At Week 24 the incidence of headache, diarrhoea, and acne was numerically higher in participants treated with the 50 mg dose of ritlecitinib (once daily) compared to placebo.

In addition, there were [redacted] participants with COVID-19-related TEAEs. The number of participants with COVID-19 TEAEs ranged from Pbo→50 mg group [redacted] to [redacted] in the 50 mg group.

There were [redacted] severe adverse events reported in the 50 mg group to Week 48; these were [redacted]. The most frequently reported severe TEAEs up to Week 24 were [redacted] in the 200/30 mg group. By Week 48, the most frequently reported severe TEAEs were [redacted] in the 10 mg group (Table 26).[154]

Table 26: TEAEs with an incidence rate of ≥5% patients in any treatment group by Preferred Term (SAS)

Treatment group[a]	200/ 50 mg (n=131)	200/ 30 mg (n=129)	50 mg (n=130)	30 mg (n=132)	10 mg (n=62)	Pbo→ 200 /50 mg (n=65)	Pbo→ 50 mg (n=66)
TEAE with incidence rate of ≥5% up to Week 24[b], n (%)							
URTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Folliculitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Urticaria	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Dizziness	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
UTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Acne	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Myalgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
TEAE with incidence rate of ≥5% up to Week 48[b], n (%)							
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
URTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Folliculitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Urticaria	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dizziness	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Influenza	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acne	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myalgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rash	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oropharyngeal pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Insomnia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abdominal pain upper	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nasal congestion	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Constipation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants in

the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg ritlecitinib. [b] in at least one group, participants were only counted once per group per event
Abbreviations: mg, milligram; n, number of participants; Pbo, placebo; SAS, safety analysis set; TEAE, treatment-emergent adverse event; URTI, upper respiratory tract infection; UTI, urinary tract infection

B.2.10.1.5 Treatment-related adverse events

The proportion of participants with TRAEs was similar across treatment groups.[154] In the 50 mg group, TRAEs were reported in [redacted] ([redacted]) participants to Week 24, and in [redacted] ([redacted]) participants to Week 48 (Table 27).[154]

Severe TRAEs were reported in [redacted] participants, including [redacted] participant in the 50 mg group, up to Week 24 with no further events reported from Week 24 to Week 48. [redacted] participants experienced serious TRAEs up to Week 24; these were in the 200/50 mg[redacted]10 mg [redacted] groups. By Week 48, there were a total of [redacted] participants who experienced serious TRAEs; [redacted] who experienced a serious TRAE [redacted] was in the 50 mg group (see Section B.2.10.1.7).[154]

Table 27: Overall summary of TRAEs (SAS)

Treatment group[a]	200/ 50 mg (n=131)	200/ 30 mg (n=129)	50 mg (n=130)	30 mg (n=132)	10 mg (n=62)	Pbo→ 200 /50 mg (n=65)	Pbo→ 50 mg (n=66)
TRAE up to Week 24, n (%)							
Number of AEs	94	82	83	97	38	86	
Participants with AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Participants with serious AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Participants with severe AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
TRAE up to Week 48, n (%)							
Number of AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Participants with AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Participants with serious AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Participants with severe AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants in the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg ritlecitinib.
Abbreviations: AE, adverse event; mg, milligram; n, number of participants; Pbo, placebo; SAS, safety analysis set; TRAE, treatment-related adverse event

B.2.10.1.6 Common treatment-related adverse events

In the 50 mg group there were no TRAEs with a 5% higher frequency up to Week 24, while up to Week 48, acne [redacted] and headache [redacted] were frequently reported. The most frequently reported TRAEs across all groups up to Week 24 were headache, upper respiratory tract infection, nausea and folliculitis (Table 28). Up to Week 48, the reported events were similar to Week 24, the most frequently reported TRAEs were headache, acne, upper respiratory tract infection, folliculitis, nausea, nasopharyngitis and diarrhoea.

Table 28: Common treatment-related adverse events

Treatment group[a]	200/ 50 mg (n=131)	200/ 30 mg (n=129)	50 mg (n=130)	30 mg (n=132)	10 mg (n=62)	Pbo→ 200 /50 mg (n=65)	Pbo→ 50 mg (n=66)
TRAE with incidence rate of ≥5% up to Week 24[b], n (%)							
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Folliculitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
URTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Myalgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
TRAE with incidence rate of ≥5% up to Week 48[b], n (%)							
Folliculitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
URTI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acne	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myalgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants in the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg ritlecitinib.
Participants were only counted once per group per event

B.2.10.1.7 Serious adverse events

In total, [redacted] participants experienced [redacted] SAEs up to Week 48; of these, [redacted] participants experienced [redacted] SAEs up to Week 24. In the 50 mg arm, there were [redacted] reported serious adverse events. In [redacted] participants, the serious AEs were considered related to treatment with ritlecitinib by the investigator.[154]

There were [redacted] participants with [redacted] events of serious infections and no opportunistic infections. All TEAEs of herpes zoster were mild to moderate and occurred in participants treated with ritlecitinib. By Week 48, [redacted] participants experienced serious TRAEs; these were in the 200/50 mg ([redacted] and [redacted] in [redacted] participant), 50 mg ([redacted]) and 10 mg ([redacted]) groups.

Up to Week 24, [redacted] participants experienced [redacted] serious TEAEs of serious infection; by Week 48, [redacted] participants ([redacted]) had experienced [redacted] serious infections. These included [redacted] and [redacted]. The outcome for these events was recovered/resolved. There were no serious TEAEs adjudicated by an external Opportunistic Infection Review Committee as opportunistic infections up to Week 48.

Up to Week 48, [redacted] participant in the 50 mg group experienced a serious TEAE of [redacted] of moderate severity on Day 169.[154] The sponsor deemed the event to be related to the study intervention, but not related to concomitant drugs or clinical trial procedure. Potential risk factors for the pulmonary embolism included COVID-19 infection, ongoing medical history of monoclonal gammopathy of undetermined significance, morbid obesity, sleep apnoea, and cardiovascular disease (hypertension, hyperlipidaemia).

Up to Week 48, [redacted] participants each experienced a serious AE of malignancy).[154] One female participant in the 200/50 mg group was diagnosed with [redacted] on Day 68, which was assessed by the investigator as unrelated to treatment with ritlecitinib. A second female in the 50 mg group was diagnosed with [redacted] on Day 195, assessed by the investigator as being related to study treatment. Both participants permanently discontinued from the study. As of database release, the outcome of the former TRAE was not recovered/not resolved whereas the status of the latter TRAE was recovering/resolving.[154] There were no AEs adjudicated as nonmelanoma skin cancer up to Week 48.[154]

B.2.10.1.8 Deaths

There were no deaths reported during the study up to Week 48.[154]

B.2.10.1.9 Safety events of interest

Dermatological events

By Week 48, [redacted] participants experienced dermatological TEAEs ([redacted] in the 10 mg group to [redacted] in the 200/30 mg group).The dermatological TEAEs reported most frequently (≥5% of participants) across treatment groups were acne ([redacted]), folliculitis ([redacted]), and urticaria ([redacted]). Rash was experienced numerically more frequently in the 50 mg group ([redacted] of participants) compared to other groups, including placebo (range [redacted]). Folliculitis, urticaria, and pruritus were experienced more frequently in the 200/50 mg ([redacted] and [redacted], respectively) and 200/30 mg groups ([redacted] and [redacted] respectively) compared to other groups, including placebo. There were [redacted] participants ([redacted] in the 200/30 mg group and [redacted] in the 50 mg group) that discontinued from the study or study drug due to TEAEs of urticaria.[154]

Neurological and audiological events

In the non-clinical studies there was a species-specific (dogs only) finding of axonal dystrophy (swelling). At high systemic exposures (33 times the unbound AUC at the chronic human dose of 50 mg), the finding was reversible but considered adverse because it was associated with abnormal functional auditory testing via brainstem auditory evoked potential (BAEP). Due to this finding potential auditory changes were monitored in the clinical studies. In the Phase 2b/3 study there were [redacted] participants overall with TEAEs confirmed to meet the criteria for neuro-safety events of interest; in [redacted] participants these events of interest were neurological and in [redacted] participants they met the criteria for sensorineural hearing loss. None of the TEAEs were consistent with central hearing disorder and no participants discontinued the study due to these events. (Full detail about the assessment of neurological and audiological events is presented in appendix F).

Serious infections

An increased incidence of serious infections has been observed during treatment with other JAK inhibitors.[160]

In the ritlecitinib studies, the data do not suggest a meaningful increase in the incidence of serious infections overall compared to placebo nor exhibit a dose response. The most frequent serious infections in ritlecitinib-treated patients were herpes zoster and herpes simplex.

[redacted] serious infections occurred in [redacted] participants ([redacted]) treated with ritlecitinib in the ALLEGRO 2b/3 study, as detailed in Appendix F. [redacted] of the [redacted] serious infections were deemed not related to treatment. None of the serious infections occurred in the ritlecitinib 50 mg treatment arm, nor did any serious infections occur in patients treated with placebo.

There were no opportunistic events of tuberculosis, fungal infections (including invasive fungal infections), mycobacterial infections, or infections with other intracellular bacteria.

Herpes zoster

Up to Week 24, [redacted] participants reported herpes zoster TEAEs; by Week 48, a total of [redacted] participants experienced herpes zoster TEAEs across the 200/50 mg [redacted] 200/30 mg ([redacted]) and 50 mg ([redacted]) groups. Participants who reported herpes zoster ranged in age from 28 to 65 years and most were female ([redacted]). No occurrences were disseminated or multi-dermatomal and all were mild to moderate in severity, all participants recovered, and no participants discontinued from the study due to these TEAEs.[154]

Herpes simplex

Up to Week 24, [redacted] participants experienced herpes simplex TEAEs, increasing to a total of [redacted] participants at Week 48. Events were reported across all groups, including [redacted] in the 50 mg group, and were not dose dependent. All herpes simplex TEAEs were mild to moderate in severity, all participants recovered, and no participant was discontinued from the study due to a herpes simplex TEAE. [154]

Cardiovascular and thromboembolic events

[redacted] participant experienced a SAE of pulmonary embolism of moderate severity in the 50 mg group. The [redacted] was discontinued from the study and the event was recovered. No events of MACE were reported.[154]

Laboratory abnormalities

Treatment with ritlecitinib was associated with changes in haematological parameters, some of which were dose dependent.[151] Neutrophil count decreased: [redacted] participants ([redacted] in the 200/50 mg, 30 mg and Placebo/50 mg groups). CD4 lymphocytes decreased: [redacted] participants ([redacted] in groups 200/30 mg, 50 mg, 30 mg, Placebo/200/50 mg). Lymphocyte count decreased: [redacted] participants ([redacted] in the 200/50 mg group, [redacted] each in groups 50 mg and 30 mg)

Ritlecitinib had a small, transient, dose dependent effect on total cholesterol, HDL cholesterol, and LDL cholesterol (relative to placebo).[154] Up to Week 24, there were dose dependent increases in total cholesterol (200/50 mg: median change from baseline: [redacted] mg/dL at Week 4; 50 mg: median change from baseline: [redacted] mg/dL at Week 4) and LDL cholesterol (200/50 mg: median change from baseline: [redacted] mg/dL at Week 4; 50 mg: median change from baseline: [redacted] mg/dL at Week 4) compared to placebo (median change from baseline: [redacted] and [redacted] mg/dL at Week 4, respectively). Levels of total cholesterol and LDL cholesterol remained stable up to Week 48. [154] Up to Week 24, there was a small, transient increase in HDL cholesterol in participants who received a 200 mg loading dose (200/50 mg, median change from baseline: [redacted] mg/dL at Week 4), whereas for placebo there was no change compared to baseline at Week 4. HDL cholesterol then remained stable up to Week 48. [154] Up to Week 24, there were small, variable changes in triglycerides; there were no long-term changes up to Week 48.[154]

There were [redacted] clinically relevant effects of ritlecitinib on ALT, AST, bilirubin, or alkaline phosphatase. The incidence of elevation in hepatic enzymes was [redacted]. Up to Week 48, there were [redacted] potential Hy's law cases. [154]

There were [redacted] clinically meaningful changes in CK across treatment groups.[154]

The median immunoglobulin values for [redacted] participants were within normal range at baseline. There were [redacted] clinically meaningful changes in IgG, IgM or IgA across treatment groups Up to Week 48.[154]

B.2.10.2. ALLEGRO-LT

For this interim results analysis, safety data were included up to the data cut-off of February 2022 and efficacy data were included through Month 24.[156]

In the *de novo* ALLEGRO-LT interim analysis, the safety profile for ritlecitinib was consistent with the primary ALLEGRO 2b/3 study results. Among [redacted] patients (safety population), there were a total of [redacted] AEs and the [redacted] of these AEs ([redacted]) were mild to moderate in severity (Table 29).[156] Among patients who discontinued the study due to an AE, the most common discontinuation AEs were pregnancy ([redacted]) and dizziness ([redacted]).[156]

Table 29: ALLEGRO-LT: Interim Safety Summary (*De Novo* Cohort)

***De Novo* Cohort**	**Ritlecitinib 200/50 mg (N = 447)***
Patients with AEs, n (%)	[redacted]
Patients with severe AEs, n (%)	[redacted]
Patients with SAEs, n (%)	[redacted]
Patients who discontinued from study due to AEs, n (%)	[redacted]
Patients with a temporary study drug discontinuation due to AEs, n (%)	[redacted]

Abbreviations: AE, adverse event; SAE, serious adverse event.
*Note: *Safety analysis set.*
Source: Source: Sinclair R, et al. (EADV 2022).[156]

At the interim analysis, the most common AEs were headache [redacted] SARS-CoV-2 test positive [redacted], acne [redacted], nasopharyngitis [redacted] and urticaria [redacted] (Table 30).

Table 30: ALLEGRO-LT: Most Commonly Occurring AEs (≥ 5% of Patients) at Interim Analysis

***De Novo* Cohort**	**Ritlecitinib 200/50 mg (N = 447)***
Headache	[redacted]
SARS-CoV-2 test positive	[redacted]
Acne	[redacted]
Nasopharyngitis	[redacted]
Urticaria	[redacted]
Pyrexia	[redacted]

Fatigue	[redacted]
Cough	[redacted]
Upper respiratory tract infection	[redacted]

Abbreviations: AE, adverse event.
*Note: *Safety analysis set.*
Source: Sinclair R, et al. (EADV 2022).[156]

At the interim analysis for AEs of special interest, [redacted] opportunistic infections were reported.[291] All events of herpes zoster were mild or moderate and none were disseminated. The [redacted] death reported was due to breast cancer and considered by the investigator to be unrelated to the study drug.[156]

For laboratory measures at the interim analysis, [redacted] clinically relevant median changes from baseline in hematologic parameters were noted.[291] [redacted] study patients discontinued or met criteria for discontinuation for neutrophils, lymphocytes, or platelet counts. [redacted] study patients met criteria for discontinuation based on AST, ALT, or CK values. For lipid levels, [redacted] clinically relevant changes were observed over time.[156]

B.2.10.3. Pooled safety analysis

A safety analysis pooling data across the clinical trial programme for ritlecitinib is available in Appendix F.

B.2.11. Ongoing supportive studies

There are currently two active ongoing supportive studies as discussed in section B.2.2.2:

- ALLEGRO-2a mechanistic study:[161] A global Phase 2a randomised, double-blind, placebo-controlled study to evaluate the safety and tolerability of ritlecitinib in adults aged 18 to ≤50 years of age with ≥25% scalp hair loss due to AA.
- ALLEGRO – LT:[162] An ongoing Phase 3 open-label long-term (2 year) study to investigate the safety and efficacy of ritlecitinib in adults and adolescents (12 years of age and older).

B.2.12. Innovation

- **Ritlecitinib offers a novel mode of action, and the selective inhibition of JAK3 spares the undesirable sides effects of JAK2 inhibition**

Ritlecitinib is a novel treatment for AA which inhibits the JAK3 and TEC signalling pathways involved in the pathogenesis of the disease . It is orally-administered and targets several cytokine pathways implicated in AA. As it is a small molecule there is no anticipated immunogenicity and so it is unlikely to generate antidrug antibodies which may potentially result in loss of efficacy over time.

Ritlecitinib has been demonstrated to achieve both clinically meaningful and statistically significant scalp hair regrowth. It blocks JAK3 and is less potent against other JAK isoforms. Inhibition of JAK3 is a desirable target to modulate a broad range of cytokines involved in the pathogenesis of AA while reducing the risk for undesirable effects of JAK2 inhibition, such as neutropenia and anaemia.

- **The oral route of administration for ritlecitinib is preferable for some patients**

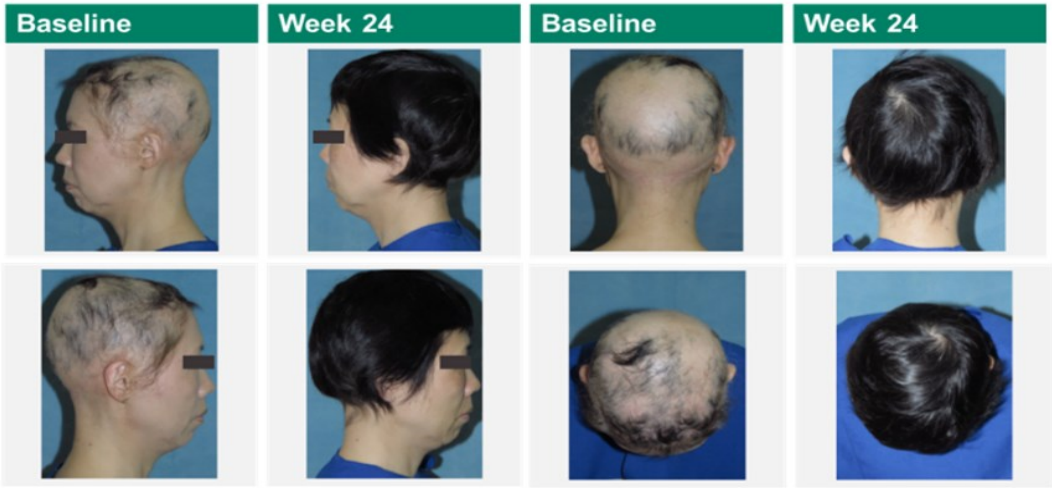
There remains a clear unmet need for oral treatments that are effective with an acceptable safety profile for adult and adolescent patients with AA. The oral route of administration for ritlecitinib is preferable for some patients, might potentially reduce unnecessary visits to hospital or specialist settings, along with once daily dosing which potentially aids compliance.

B.2.13. Interpretation of clinical effectiveness and safety evidence

B.2.13.1. Principal findings from the clinical evidence highlighting the clinical benefits and harms of the technology

Ritlecitinib is a new treatment option for patients aged 12 years and older with severe AA. The efficacy and safety of ritlecitinib as therapy for patients with severe AA has been conclusively demonstrated in the ALLEGRO 2b/3 trial (Section B.2.2.1). Figure 26 illustrates the impact of ritlecitinib on a patient with AA after 24 weeks who was treated within the clinical trial programme.

Figure 26: Hair appearance before and during treatment with ritlecitinib



Shared with patient consent. This patient was randomised to 50mg ritlecitinib in the ALLEGRO 2b/3 study, Abbreviations: SALT, Severity of Alopecia Tool

In the ALLEGRO 2b/3 study, ritlecitinib 50 mg, the dose to be considered for this submission and to be used in clinical practice, demonstrated a statistically significantly improved response based on both SALT ≤10 and ≤20 at Week 24 compared to placebo at an overall significance level (α) of 0.01.[154] This demonstrates that over a 24 week period, patients treated with 50 mg of ritlecitinib (once daily) are significantly more likely to experience clinically meaningful hair regrowth than patients who are treated with placebo. The trajectory seen in ALLEGRO-LT, showing the proportion of patients achieving a SALT score of ≤20 or ≤10 over time was consistent with the ALLEGRO-2b/3 study, however, as it was not placebo-controlled the significance of this cannot be determined.

For participants receiving ritlecitinib 50 mg, the difference in response based on PGI-C at Week 24 was statistically significantly different to placebo at an overall significance level (α) of 0.01 showing that the patients perceive a positive change in AA as a result of treatment with ritlecitinib.[154] As for SALT scores, the trajectory of the proportion of patients achieving a PGI-C response over time was consistent in the ALLEGRO-LT study, however, as it was not placebo-controlled the significance of this cannot be determined. The findings on PGI-C in both the ALLEGRO 2b/3 and ALLEGO-LT studies were further supported by P-Sat scores at Week 24 in the ritlecitinib 50 mg group which were over double the P-sat score recorded in patients treated with placebo, though this was not tested for statistical significance. Since the

burden of AA to patients is predominantly psychosocial (Section B.1.3.2), patient satisfaction with treatment is integral to improving the quality of life in patients with AA.

In addition to the benefits on hair regrowth on the scalp, treatment with ritlecitinib demonstrated improvements in hair regrowth in the eyebrows and eyelashes. The proportion of participants in the 50 mg group with an ELA and EBA response at Week 24 were also statistically significantly different compared to the placebo group, indicating that participants experienced regrowth of hair both on their scalp and their eyebrows and eyelashes.[154]

Subgroup analyses in the ALLEGRO 2b/3 study looked at age and participants with AT/AU. Ritlecitinib is similarly effective in adolescent and adult patients. Response rates based on SALT ≤10 and SALT ≤20 at Week 24 were similar in adolescents and adults. Ritlecitinib has also been shown to be effective in patients with AT/AU providing evidence that ritlecitinib is a treatment option for patients with AT/AU. A significant difference was observed at Week 24 in the proportion of responders based on SALT ≤20 and response based on PGI-C in the 50 mg group. Patients with non-AT/AU AA had a higher response based on SALT ≤20 than those with AT/AU, indicating that while ritlecitinib is effective in patients with AT/AU, treating a patient with ritlecitinib sooner can lead to a better response. This result is not unexpected as patients with AT/AU have more severe hair loss at baseline than those with non-AT/AU, and as such may require extended treatment to reach SALT ≤20 thresholds.

Overall, the efficacy data from the phase 2b/3 trial demonstrates ritlecitinib to be an effective treatment for regrowing hair in both adult and adolescent patients with severe AA. Statistically significant response rates in PGI-C, ELA, EBA and P-Sat in the 50 mg group compared to placebo demonstrate the positive effects that hair regrowth, measured by SALT score, has on the treatment satisfaction of patients with severe AA as a result of response to treatment with ritlecitinib. The favourable benefit-risk assessment of ritlecitinib represents an important advance for the treatment of adult and adolescent subjects with severe AA, including AT and AU.[6]

B.2.13.2. Safety

Ritlecitinib was well tolerated at all doses studied, including once daily 50 mg, up to Week 48 in the patient population included in the ALLEGRO 2b/3 study and the safety profile of ritlecitinib in the study was consistent with that observed in previous studies in healthy volunteers and in patients with AA.[154] The incidence of all AEs was also similar between the 50 mg treatment group and the placebo groups at Week 24.[154]

The most frequently reported TEAEs up to Week 48 are consistent with the reported TEAEs in the previous Phase 2a trial.[32,159] Groups treated with a 200 mg loading dose had a higher incidence of some TEAEs and numerically greater changes in some haematological parameters compared to the 50 mg group, further justifying the choice of treating patients with AA with once daily 50 mg of ritlecitinib without a loading dose in clinical practice.

Most TEAEs were mild or moderate in severity. In total, 14 participants experienced serious AEs, which were generally balanced across treatment groups and are summarised in Section 0.

B.2.13.3. Strengths and limitations of the clinical evidence base for the technology

B.2.13.3.1 Strengths

- The ALLEGRO 2b/3 study (B7981015) is a robust placebo-controlled, multi-centre, multinational clinical trial programme which enrolled over 700 patients with severe AA. Baseline characteristics were comparable between treatment arms and were broadly comparable between trials. Similarly, the trial included 10 sites in the UK and it is therefore expected that the benefits reported for this trial are likely to be reflected in clinical practice in England and Wales (See Section B.1.3.1.3 on epidemiology of AA).
- The full trial populations for the trial included patients who had not been previously exposed to systemic therapies, which aligned with the licensed indication. Therefore, data are presented for the anticipated population, providing justification for using the full trial population in the model.

- As patients with AA have no treatment options (as discussed in Section B.1.3.3.5), placebo is an appropriate comparator meaning the ALLEGRO 2b/3 study (B7981015) provides head-to-head evidence with best current standard of care in clinical practice in England and Wales.
- The ALLEGRO 2b/3 study was a comprehensive trial of ritlecitinib which demonstrated a positive response based on SALT score ≤20 and ≤10 which was statistically significantly higher in the ritlecitinib 50 mg treatment group than placebo at Week 24 ($P < 0.0001$ and $p<0.001$ respectively).
- The efficacy results of the ALLEGRO 2b/3 study have been supported by similar trends in the ALLEGRO-LT study which studied open-label treatment of ritlecitinib.
- The incidence of AEs in all treatment placebo groups were broadly similar and the safety profile of ritlecitinib in the study was consistent with that observed in previous studies in healthy volunteers and in patients with AA.

B.2.13.3.2 Limitations

- Data from the phase 2b/3 clinical trial provide evidence of the efficacy of ritlecitinib against placebo until 24 weeks only. However, this timeframe was still sufficient to demonstrate a statistically significant improvement versus placebo with continually increasing response trends. Placebo arm patient response trends based on SALT <10 and SALT <20 remained constant from Week 12 through to Week 24 at [redacted]. However, in ALLEGRO 2b/3, there was a 24-week extension phase where placebo-treated patients switched to active treatment with ritlecitinib in a pre-specified, blinded manner, while other arms continued on the same maintenance dose. We would not expect any additional change from placebo beyond 24 weeks. This data therefore provides additional evidence of clinical response from ritlecitinib cross over group. Moreover, continued follow-up in the open-label extension for a maximum of 36 months also provides evidence of the efficacy of ritlecitinib in the long term.
- The exclusion of patients with comorbid psychiatric conditions including recent or active suicidal ideation or behaviour is a requirement for clinical trial design

and is standard across clinical trials.[163] Investigation is needed to determine how comorbid psychiatric conditions may impact the efficacy of treatment with ritlecitinib.

- Whilst the ALLEGRO-LT study rolled over patients from ALLEGRO 2b/3 and ALLEGRO 2a on a dose of ritlecitinib 50 mg, the study enrolled *de novo* patients onto ritlecitinib 50 mg with a 200 mg loading dose for four weeks, so results from *de novo* patients are not directly comparable to the ALLEGRO 2b/3 study. Additionally, the ALLEGRO-LT study is not placebo-controlled, so the statistical significance of the outcomes observed cannot be determined.

B.3 Cost-effectiveness

B.3.1. Published cost-effectiveness studies

An economic SLR was performed to identify published economic evidence for the treatment of AA up to 15th October 2021. This SLR sought to identify both cost-effectiveness studies and cost and resource use studies. Please see Appendix G for the methods used to identify relevant studies, and the description and quality assessment of any identified studies.

There were no relevant cost-effectiveness analyses in patients with AA identified in the SLR. All studies identified in the SLR were disease burden analysis and evaluated global, regional, and national disability-adjusted life years (DALYs) from the Global Burden of Disease Study.

B.3.2. Economic analysis

A *de novo* economic model is included in the submission, comparing ritlecitinib 50 mg with BSC for adult and adolescent patients with severe AA.

A *de novo* economic model was developed because there are no published cost-effectiveness analyses of treatments of AA, or other relevant economic models, to inform a model adaptation.

B.3.2.1. Patient population

The population entering the cost effectiveness model (CEM) includes adult and adolescent patients with severe AA, in line with the population considered in the decision problem (Section B.1.1).

B.3.2.2. Comparator

Established clinical management is specified as the comparator for ritlecitinib. Although there are a large number of potential off-licence treatments, none are considered as suitable comparators for the treatment of severe AA by dermatologists with a specialist interest in hair disorders and, moreover, all of the dermatologists agreed the most relevant comparator for ritlecitinib is BSC defined as no

pharmacological treatment (3/3 clinicians, 100%),[17] as further described in Section B.1.3.3.6.

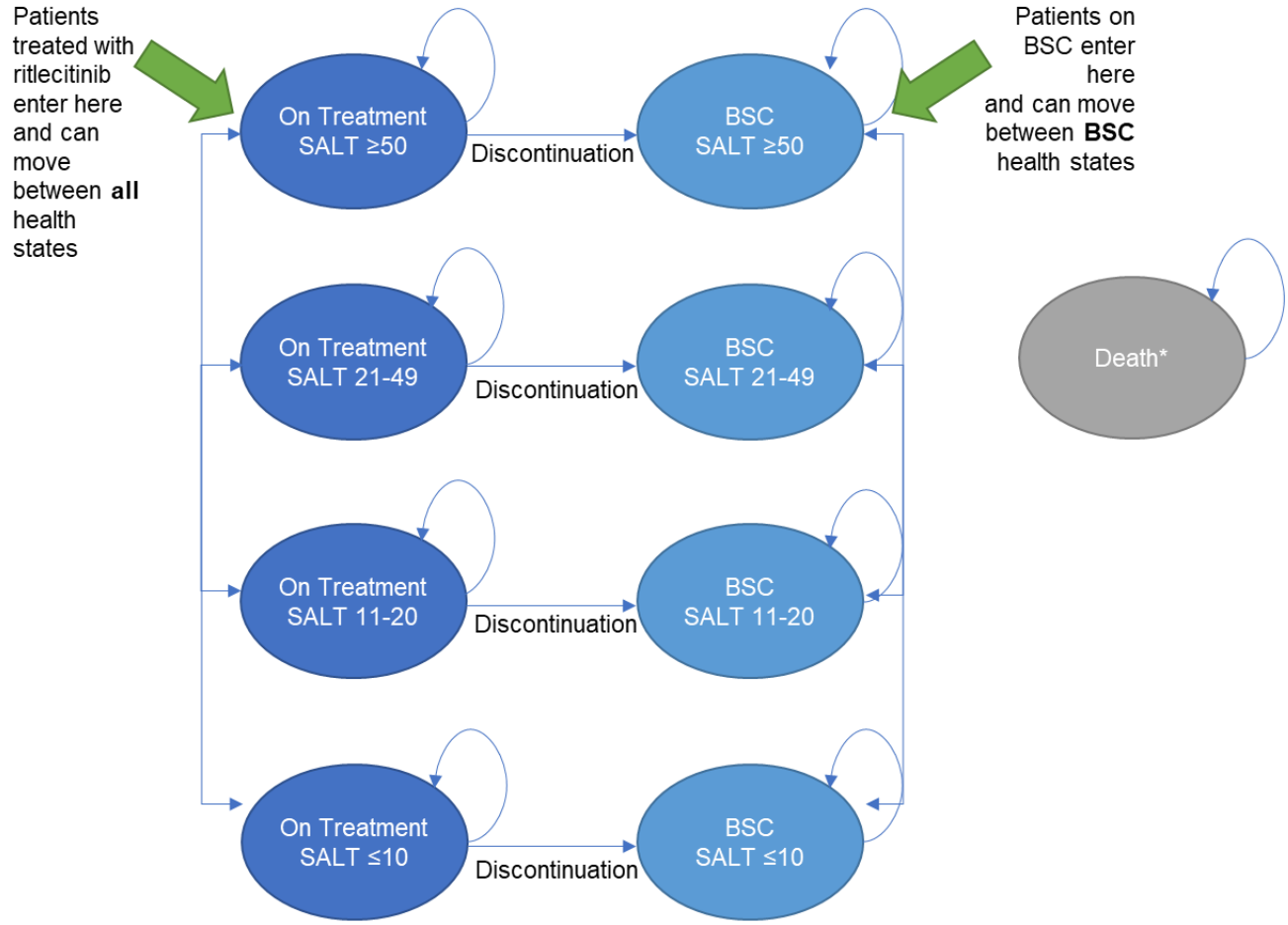
B.3.2.3. Model structure

A *de novo* semi-Markov model was adopted to capture the long-term, chronic nature of AA. The model structure was influenced by discussions with NICE Scientific Advice regarding potential economic modelling approaches to AA. As part of these discussions, Pfizer submitted an early CEM using pre-clinical trial data for ritlecitinib when used in patients with AA, based on which NICE made recommendations which were used to inform the final model structure and choice of health states. The choice of a semi-Markov structure, health states and use of the SALT score to define health states was also validated by UK dermatologists with a specialist interest in hair disorders.[17]

The model structure, as described in Figure 27, simulates the movement of patients between health states based on the absolute SALT score of patients when treated with either ritlecitinib or BSC. Patients can move into the death state at any time in the model.

Patients enter the model in the SALT ≥50 health state because the patient population being considered are those with severe AA, which is deemed as interchangeable with having ≥50% scalp hair loss (SALT ≥50), as fully described in Section B.1.1. Patients treated with ritlecitinib begin on-treatment whereas those on BSC do not. Whilst on treatment with ritlecitinib or on BSC, patients move between the health states; patients 'on treatment' can move to the BSC health states but not vice versa. The health states are linked to patients' absolute SALT scores, detailed below (Table 31), aligned with the primary outcome of the ALLEGRO 2b/3 study.

Figure 27: Graphical representation of the model structure



**All patients may transition to death from any health state. A user option exists to disable death in the first 48 weeks so that the model results match the observed values in the clinical trial.*

Table 31: Health states and definitions

Health state	Definition
On Treatment; SALT ≥50	All patients receiving ritlecitinib treatment (i.e., not BSC) enter the model in this health state and remain on active treatment.
On Treatment; SALT 21-49	Patients in this health state remain on active treatment are considered to have a partial response to treatment as their SALT score has improved from baseline (SALT 50-100).
On Treatment; SALT 11-20	Patients in this health state are considered to have a response to ritlecitinib with SALT 11-20 and remain on active treatment.
On Treatment; SALT ≤10	Patients in this health state are considered to have a response to ritlecitinib with SALT ≤10 and remain on active treatment.
BSC; SALT ≥50	BSC patients enter the model in this health state. Patients who discontinue active treatment will accumulate in this health state if they do not experience spontaneous remission (as described in B.3.2.3.3).
BSC; SALT 21-49	Patients in this health state are not on active treatment and have a SALT score of 21-49.

BSC; SALT 11-20	Patients in this health state are not on active treatment and have a SALT score of 11-20.
BSC; SALT ≤10	Patients in this health state are not on active treatment and are assumed to have spontaneous remission (as described in B.3.2.3.3).
Death	Death has occurred due to any cause. Patients can transition to the Death heath state from any health state.

Abbreviations: BSC, best standard of care, SALT, Severity of Alopecia Tool

B.3.2.3.1 Short-term state membership (to week 48)

During the first 48 weeks (four cycles), patients treated with ritlecitinib are partitioned in SALT-based health states based on the ALLEGRO 2b/3 clinical trial, hence the "semi" portion of the model description. The distribution of patients treated with ritlecitinib across different SALT scores during the first 48 weeks is therefore defined and not linked to a previous health state.

Patients on BSC are partitioned in SALT-based health states based on the ALLEGRO 2b/3 clinical trial until Week 24 (two cycles). The distribution of patients on BSC across different SALT scores during the first 24 weeks is therefore defined and not linked to previous health state.

B.3.2.3.2 Stopping rule

An interim and a final stopping rule is applied at Week 24 and Week 48 for patients treated with ritlecitinib and informs transitions to BSC or 'off-treatment' health states. The interim stopping rule is that patients will discontinue treatment with ritlecitinib if they show an initial worsening at Week 24 compared to baseline (i.e., a worse SALT score). The final stopping rule is that patients who do not achieve a SALT score ≤20 at Week 48 will discontinue treatment.

Based on outputs from the Therapeutic Landscape Delphi panel, all dermatologists with a specialist interest in hair disorders referred to the SALT score when asked what quantitative measures they use to define severity (7/7 clinicians, 100%) and the majority of dermatologists with a specialist interest in hair disorders agreed that in the absence of DLQI, the SALT score alone is appropriate for measuring response in clinical practice for patients with severe AA including AT/AU (6/8 clinicians, 75%).[21] Moreover, when asked what might be an appropriate measure of clinical response for patients with severe AA when treated with JAK inhibitors, dermatologists with a

specialist interest in hair disorders agreed the SALT score alone is a good objective measure (8/8 clinicians, 100%). The use of the SALT score for the stopping rule was further validated by two KOLs, who both stated it is an appropriate measure and one KOL added:[17]

"The degree of hair loss is in line with patient feelings, SALT score is basis of everything in terms of treatment."

The majority of dermatologists with a specialist interest in hair disorders agreed that achieving a clinically meaningful response was associated with a meaningful improvement in HRQoL (7/8 clinicians, 88%) and also agreed that there is a correlation for individual patients between the severity of hair loss and HRQoL, meaning patients with more hair loss would have a lower HRQoL (8/8 clinicians, 100%).[21] The ALLEGRO 2b/3 study showed strong correlations of PGI-C responders and P-Sat responders with SALT ≤20 responders, which demonstrates the value of achieving SALT ≤20 to patients with AA (Figure 20 and Figure 22).

A two-phase stopping rule is proposed because hair growth is not immediate and patients continue to reach higher thresholds of response beyond Week 24 and patients who show no clinically meaningful response at Week 24 can go on to respond by Week 48, as shown in the ALLEGRO 2b/3 clinical trial (see Section B.2.6 and Figure 28) and supported by the opinion of consultant dermatologists.[21] An interim stopping rule at Week 24 therefore prevents patients discontinuing who are slower to respond to treatment with ritlecitinib while alleviating the need to treat all eligible patients for at least 48 weeks as, based on clinical opinion, 24 weeks allows sufficient time to determine whether a patient would worsen with treatment. The final stopping rule is that patients who do not achieve a SALT score ≤20 at Week 48 will discontinue treatment. The final stopping rule at Week 48, was chosen in line with the anticipated label and because it allows sufficient time to show long-term response to ritlecitinib based on clinical data and opinion, and achieving a SALT score ≤20 is a clinically meaningful threshold for patients with AA.[17,21] One of the three dermatologists with a specialist interest in hair disorders added that they felt using SALT score ≤20 rather than SALT score ≤10 was more appropriate as a SALT score ≤10 was too stringent; if

they told a patient they had to discontinue treatment who had a SALT score of 20 the patient would be “very upset”.[17]

Figure 28: Week 48 SALT score given relative improvement in SALT score at Week 24 compared to baseline

Abbreviations: SALT, Severity of Alopecia Tool
Graph shows that patients who show even minimal improvement at 24 weeks respond at 48 weeks i.e., 1 patient achieves SALT 11-20 at week 48 whom at 24 weeks showed 0 to <1% improvement. Only those that worsen at 24 weeks were not shown to go onto respond at 48 weeks.

A scenario analysis was conducted using only the final stopping rule (i.e., only patients who do not achieve a SALT score ≤20 at Week 48 will discontinue treatment due to lack of response) to evaluate the impact of the interim stopping rule on the incremental cost-effectiveness ratio (ICER) of ritlecitinib compared to BSC.

B.3.2.3.3 Discontinuation of ritlecitinib

Following discontinuation of ritlecitinib to BSC at any time point in the model, it is assumed that patients gradually lose any prior improvement in SALT score. This assumption was validated by dermatologists with a specialist interest in hair disorders, stating that most patients will progressively regress after discontinuing treatment.[17] Patients initially enter a BSC health state with the same SALT score range they were in while on treatment with ritlecitinib. Patients remain in that state for one cycle before

transitioning sequentially through the health states with a greater SALT score each cycle until reaching ‘BSC SALT ≥50’. For example, a patient would transition from ‘On Treatment SALT 11-20’ to ‘BSC SALT 11-20’ where they would remain for one cycle before transitioning to ‘BSC SALT 21-49’, followed by ‘BSC SALT ≥50’. The exception to this is patients who are assumed to be in spontaneous remission after discontinuing treatment and transition to ‘BSC SALT ≥ 10’, as fully described in Section B.3.2.3.4.

B.3.2.3.4 Longer term transitions (week 24+ for BSC and week 48+ for ritlecitinib)

The movement of the cohort through health states according to SALT score, mortality and treatment discontinuation following the final discontinuation rule at Week 48 are handled through Markov processes.

Patients who are treated with ritlecitinib can move between the on-treatment health states according to SALT score. Given there is no waning effect seen in the ALLEGRO-LT study, it is assumed that patients remain within the same health state after Week 96 unless they discontinue ritlecitinib treatment.

After Week 48, any ritlecitinib-treated patients who move to a health state where their SALT score is >20 are assumed to discontinue treatment due to loss of response. Patients may discontinue ritlecitinib treatment for any reason from health states with a SALT score ≤20 which is derived by extrapolating time on treatment amongst patients with a SALT score ≤20. Following discontinuation of ritlecitinib treatment, it is assumed that patients gradually lose any existing improvement in SALT score.

Patients initially enter a BSC health state with the same SALT score range they were in while on treatment with ritlecitinib for one cycle, before transitioning sequentially through the health states with a greater SALT score each cycle until reaching ‘BSC SALT ≥50’.

Patients on BSC (whether they had previously discontinued treatment with ritlecitinib or not) are able to experience spontaneous remission by moving to the ‘BSC SALT ≤10’ health state. As it is not known whether spontaneous remission is durable, it was assumed some patients lose spontaneous remission over time and an equal number of patients gain spontaneous remission over time. Thus, the percentage of

BSC patients with spontaneous remission remains constant amongst those alive (see Section B.3.3.3). Therefore, patients treated with BSC (i.e., those who begin on BSC or who have discontinued ritlecitinib to BSC and returned to their baseline SALT score) do not move between health states in the model after Week 24 unless they die.

B.3.2.3.5 Other model features

A lifetime horizon that allowed patients to live to 100 years (weighted average age at baseline was 37 in adults, 14 in adolescents, and 34 in adults and adolescents, giving an average time horizon of 63 years, 86 years, and 66 years, respectively) was adopted as per the NICE reference case for a non-fatal chronic disease.[164] Alternative time horizons have been considered in the scenario analyses.

A cycle length of 12 weeks was used in the model and a half-cycle correction was applied.

Costs and outcomes are discounted at 3.5% per annum in line with the NICE reference case.[164]

The model adopts a UK NHS and personal social services (PSS) perspective on costs in line with the NICE reference case, however PSS costs do not have a significant impact on results and are therefore not considered. The perspective on outcomes considers all direct health effects for patients and their caregivers, in line with the NICE reference case.[164]

Societal costs are included as a scenario analysis for adults to broaden the perspective of the model. Societal costs are incorporated into the model as productivity losses, i.e., the relative reduction in both full-time and part-time work in patients in the model compared to the UK general population per treatment cycle, based on both absenteeism (work time missed) and presenteeism (percent impairment while working). This can be used to inform the committee's deliberations as a non-reference case analysis.[164]

The model is flexible to test alternative assumptions, which are described further in Section B.3.10.3.

Table 32: Features of the *de novo* economic analysis

	Current evaluation	
Factor	**Chosen values**	**Justification**
Time horizon	Lifetime	The time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect all important differences in costs or outcomes between technologies being compared.[164] Therefore, a lifetime horizon was chosen since patients accumulate differential costs and QALYs until death
Model structure	Semi-Markov model with 12-week cycles	Able to capture the short-term treatment decisions in clinical practice and the long-term extrapolation of response following the final discontinuation rule
Response criteria	Steady or improved SALT score at Week 24 and SALT ≤20 at Week 48	Clinician input found SALT ≤20 to be the most clinically relevant measure of response for adults and adolescents[21]
Discontinuation rate	Parametric extrapolation of patient level data is applied for patients who did not stop treatment after the final discontinuation rule at Week 48. Following discontinuation of ritlecitinib to BSC, it is assumed that patients gradually lose any existing improvement in SALT score.	Based on ALLEGRO-LT trial data
Perspective	UK NHS & PSS	NICE reference case[164]
Treatment waning effect?	Included from ALLEGRO-LT data	Given that response was demonstrated to plateau in the ALLEGRO-LT trial (Section B.2.6.2), it is reasonable to assumed that patients remain in state.
Source of utilities	Vignette study data	Suitable utilities for patients with AA are not available in the literature and the EQ-5D and SF-36 data from the trial did not adequately capture change in utility for patients with AA, as discussed in Section B.1.3.2.3
Source of costs	Sources of cost data included the British National Formulary for drug costs, and national cost databases (NHS Reference Costs)	NICE reference case[164]
Discount of 3.5% for utilities and costs	Yes	NICE reference case[164]

Were health effects measured in QALYs; if not, what was used?	Yes	NICE reference case[164]

Abbreviations: AA, alopecia areata; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PSS, Personal Social Services; QALY, quality-adjusted life years; SALT, Severity of Alopecia Tool; SF-36, Short Form 36 Health Survey Questionnaire; UK, United Kingdom

B.3.2.4. Intervention technology and comparators

BSC defined as non-pharmacological therapy is considered the appropriate comparator for ritlecitinib, in line with the comparators defined in the NICE scope and decision problem (see Section B.1.1 and Table 2) and as further discussed in Section B.1.3.3.6.

B.3.3. Clinical parameters and variables

As discussed in Section B.2.2, the clinical evidence base for ritlecitinib to inform the economic model comes from the ALLEGRO 2b/3 and ALLEGRO-LT clinical trials. The FAS population of the studies are utilised for data inputs.

The clinical inputs which were included in the economic model were:

- Baseline characteristics
- Short-term response
- Long-term response
- Spontaneous remission
- Adverse events
- Discontinuation
- Mortality

For patients treated with ritlecitinib, there are multiple considerations to ensure that a patient may who remains on treatment is a responder. These are summarised in Table 33.

Table 33: Mechanisms for on-treatment patients to discontinue

Mechanism	Approach
Initial stopping rule	Patients who show a worsening in their SALT score between baseline and Week 24 will discontinue.
Final stopping rule	Patients who do not achieve a SALT ≤20 at Week 48 will discontinue treatment.

Loss of response	If, after Week 48, patients transition to a health state with a SALT >20, they will immediately discontinue treatment due to loss of response.
Discontinuation for other reasons	After Week 48, patients may discontinue treatment with a SALT score of 0-20 for reasons other than loss of response.

B.3.3.1. Baseline characteristics

The mean age of patients and the percentage of males and females entering the model were aligned with the baseline characteristics of the ALLEGRO 2b/3 trial.[147] The mean age and proportion of females entering the model are used to inform the mortality rate per cycle.

Table 34: Baseline characteristics of patients entering the model

	Value	Reference
Mean age - ≥12 to <18 years	[redacted]	ALLEGRO 2b/3 trial[147]
Mean age - ≥18 years		
Mean population age - ≥12 years		
Female - ≥12 to <18 years (%)	[redacted]	ALLEGRO 2b/3 trial[147]
Female - ≥18 years (%)		
Mean percentage female (%) - ≥12 years		

Abbreviations: AA, Alopecia Areata

B.3.3.2. Short-term response (to week 48)

The response rate to treatment with ritlecitinib was parameterised in the model according to absolute SALT score. Distributions of patients between health states up to Week 48 were derived from only the ALLEGRO 2b/3 clinical trial; the SALT score of patients was determined via review of photographs of the scalp by independent consultants.

Patients on BSC are partitioned between health states until Week 24 based on the placebo arm of the ALLEGRO 2b/3 clinical trial; SALT score was measured in the same way for placebo and ritlecitinib patients.

B.3.3.3. Spontaneous remission on BSC

Given that a [redacted] number ([redacted]) of adult and adolescent patients treated with placebo reached SALT score ≤10 in the ALLEGRO 2b/3 trial at Week 24, patients on BSC were allowed to reach full response status in the model.

In the base case, the percentage of patients on BSC assumed to achieve spontaneous remission from Week 24 was [redacted], equal to the percentage of adult and adolescent patients in the ALLEGRO 2b/3 trial who had a SALT score ≤10 at Week 24. When considering adults only in a scenario analysis, the percentage of patients on BSC assumed to achieve spontaneous remission from Week 24 is [redacted], equal to the percentage of adult patients in the ALLEGRO 2b/3 trial who had a SALT score ≤10 at Week 24. Further scenario analysis explores varying the BSC spontaneous remission assumption further to understand the impact this has on the cost effectiveness, to an upper limit of 10% (as described in Section B.1.3.1.2).

As explained in Section B.3.2.3, the percentage of patients in spontaneous remission is constant as it is assumed that an equal number of patients lose and gain spontaneous remission over time, respectively.

B.3.3.4. Long-term response (week 48+)

Following Week 48, transition matrices for patients treated with ritlecitinib were derived from the ALLEGRO-LT clinical trial to calculate the transitions between health states from Week 48 until Week 96.

To calculate the transitions from Week 48 until Week 96, data for patients with SALT ≤20 after 48 weeks of exposure to ritlecitinib were considered, ensuring transition matrices are parametrised only by patients who would have passed the final stopping rule. The following patients treated with ritlecitinib in the ALLEGRO 2b/3 and ALLEGRO-LT trials were considered:

- Patients who were treated with a 50 mg dose in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial

- Patients who were treated with a 200 mg loading dose followed by a 50 mg dose in the ALLEGRO 2b/3 trial, followed by a 50 mg dose in the ALLEGRO-LT trial
- Patients who were treated with a 30 mg dose in ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial
- Patients who were treated with a 200 mg loading dose followed by a 30 mg dose in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial
- Patients who began on placebo and transitioned to a 50 mg dose in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial
- Patients who began on placebo and transitioned to a 200 mg loading dose followed by a 50 mg dose in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial
- Patients entering the ALLEGRO-LT trial *de novo* who were treated with a 200 mg loading dose followed by a 50 mg dose of ritlecitinib

These data were included as the loading dose/initial treatment with placebo were not anticipated to have a significant impact on the patients' health states in the long term, and pooling the data reduced the uncertainty in the transition matrices calculations.

From Week 96 onwards, it was assumed that patients remained in state given that response was demonstrated to plateau in the ALLEGRO-LT trial (Section B.2.6.2).

B.3.3.5. Adverse events

The adverse events considered in the analysis were TEAEs occurring in greater than 5% of patients in the ritlecitinib 50 mg arm in the FAS population in the ALLEGRO 2b/3 trial, as described in Table 35. The rate of adverse events at Week 48 were used for patients treated with ritlecitinib. Serious infections were considered but not modelled as the occurrence of them was very rare ([redacted] of participants), the majority of the cases were not related to treatment, and they were not observed in the ritlecitinib 50 mg arm of the ALLEGRO 2b/3 study, as detailed in Section B.2.10.1.7.

Table 35: Rate of adverse events with ritlecitinib at Week 48

Adverse event	Ritlecitinib 50 mg
Acne	[redacted]
Diarrhoea	[redacted]
Folliculitis	[redacted]
Headache	[redacted]
Nasopharyngitis	[redacted]
Rash	[redacted]
Upper respiratory tract infection	[redacted]
Urticaria	[redacted]

Abbreviations: BSC, best supportive care

Adverse event probabilities from Week 48 for ritlecitinib were adjusted to calculate annual rates, which in turn were used to calculate probabilities per cycle to be used in the model for patients treated with ritlecitinib, using methods described by Briggs *et al.* (2006).[165]

$$p_B = 1 - \exp\left(-\left[-\frac{\ln(1-p_A)}{n}\right]\right)$$

Where p_A is the probability of the event from the study, n is the number of cycles occurring over the period that the event was observed, and p_B is the probability of the event over a cycle.

As placebo was only given until Week 24 in ALLEGRO 2b/3, the rate of adverse events for patients on BSC was derived from applying the risk ratio of adverse events with placebo relative to ritlecitinib 50 mg at Week 24 to the probability of adverse events for ritlecitinib each cycle. The risk ratio of adverse events with placebo relative to ritlecitinib at Week 24 was equal to [redacted] calculated from the number of participants with AEs by Week 24 (Table 36).

Table 36: Rate of adverse events with ritlecitinib and placebo at Week 24

	Ritlecitinib 50 mg (N=130)	Placebo (N=131)
Participants with adverse events, N (%)	[redacted]	[redacted]

For both ritlecitinib and BSC, the risk of adverse events is assumed to be constant over the modelled time horizon, which is a simplifying assumption given the lack of

longer-term data. The probability per cycle of each adverse event occurring is described in Table 37.

Table 37: Probability of adverse events per cycle

Adverse event	Ritlecitinib 50 mg	BSC
Acne	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Folliculitis	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]
Urticaria	[redacted]	[redacted]

Abbreviations: BSC, best supportive care

B.3.3.6. Discontinuation for reasons other than response

Parametric distributions were used to extrapolate the time on treatment following 48 weeks of ritlecitinib treatment from the ALLEGRO-LT study.

The sample of patients from the ALLEGRO-LT trial described in Section B.3.3.4 was used to inform time on treatment. Patients who were treated in China during the ALLEGRO 2b/3 and ALLEGRO-LT trials (n=81) were not included in the discontinuation estimates because the patient-level data was not available due to confidentiality requirements.

As the discontinuation was applied from Week 48, only patients who had a SALT score of 20 or less at Week 48 were included in the analysis. Patients whose SALT score increased above SALT 20 at any time after Week 48 we excluded as these patients were assumed to automatically discontinue treatment; their removal ensured discontinuation was not double-counted.

Table 38 shows the resulting Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics for each of the parametric distributions. The Generalised Gamma does not fit the data well, indicating a lack of convergence. As the exponential distribution is the next best statistically fitting to the data (according to AIC and BIC), the exponential distribution is used in the base case analysis. As shown in Figure 29, according to the exponential curve, approximately [redacted] of patients remained on treatment with ritlecitinib four years after achieving SALT score ≤20

following 48 weeks of ritlecitinib treatment. The other curves are tested in scenario analyses.

Table 38: AIC and BIC statistics for parametric distributions fit to ALLEGRO-LT discontinuation

Distribution	AIC	BIC
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]
Lognormal	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion

Figure 29: Parametric distributions fit to ALLEGRO-LT discontinuation

B.3.3.7. Mortality

All-cause mortality was estimated using UK national life tables for the years 2018-2020 with no adjustment made for AA-specific mortality because there is no conclusive evidence of any difference in the survival of patients with AA compared to the general population.[166] Although patients with severe AA report higher levels of suicide ideation and depression, there is limited evidence of a direct link between AA and increased mortality.[13,91,92] It is yet to be determined whether better management of AA through approved treatments results in lower rates of depression and therefore requires further

research. A conservative approach was taken to apply no adjustment to mortality in the cost-effectiveness analysis.

As no patients died during the ALLEGRO 2b/3 trial, a user option exists to disable death in the first 48 weeks so that the model results match the observed values in the clinical trial.

B.3.4. Measurement and valuation of health effects

AA is a complex heterogenous condition impacting across several key HRQoL domains including, but not limited to, physical, emotional and social functioning. This makes the significant burden of severe AA multi-faceted, with unique consequences on the HRQoL of individual patients lived experience. See section B.1.3.2.3.

- Thus, when considering and interpreting the relevant literature and trial data, it is prudent to acknowledge the extent of this heterogeneity, and the potential impact it may have on measuring and aligning individual lived experience with clinical results and findings.
- The ALLEGRO 2b/3 trial shows ritlecitinib has a meaningful impact on hair regrowth in adults and adolescents with severe AA, and an acceptable safety profile (Section B.2). However, despite these clinical and physical symptom improvements, differences between arms were not detected from baseline to Week 24 in EQ-5D-5L or SF-36. This is inconsistent with the burden, and HRQoL described in the literature (Sections B.1.3.2 and B.3.4.4.2)

The lack of sensitivity between treatment arms of the ALLEGRO 2b/3 study may be, in part, because generic measures of HRQoL capture some, but not all AA patient-relevant domains, such as social functioning, personal relationships, and appearance as described in Section B.1.3.2.

- Pfizer engaged with PAGs and clinicians in the UK to understand the disconnect between generic measures of HRQoL and burden of disease. They fed back that; While there is an association between the extent of hair loss and HRQoL, the relationship may not always be linear. They also stated that generic measures of HRQOL such as ED-5D lack the specificity, and relevant domains, to capture the impacts faced by patients with AA in a meaningful way.[21,22] (Section B.3.4.4.2).

- Other potential reasons for the lack of sensitivity in the EQ-5D and SF-36 between treatment arms include: High ceiling effects making it difficult to demonstrate any meaningful improvement with treatment. ALLEGRO 2b/3 eligibility criteria excluded patients with "major psychiatric conditions" which may also contribute to the high mean scores at baseline throughout the trial. Patients enrolled in the trial may have developed coping mechanisms given the mean number of years since first diagnosis of AA ([redacted]).[11,167–171] (Section B.1.3.2)
- Responder analysis for AAPPO from patients enrolled in the ALLEGRO 2b/3 trial shows a correlation between clinical response and HRQoL, indicating that increasing the number of patient-relevant domains increases sensitivity of capturing the HRQoL of patients with AA. However, AAPPO cannot be used to describe utilities as it is not a preference-based measure. (Section B.3.4.2).

In summary, generic preference-based measures that quantify the HRQoL of patients with AA are not able to overcome the complexity of correctly estimating the burden of AA. They are not specific or sensitive enough to comprehensively capture the full patient experience as reported in the literature alongside PAG and Clinician feedback.

Therefore, we propose that vignettes developed through direct patient and clinician input, as well as available literature and clinical data, valued by members of the UK general population, are the most valid and equitable source of health state utilities for cost-effectiveness analyses in AA. This is in concordance with the hierarchy of preferred HRQoL methods from NICE guidance.[170] The utility values estimated were within a similar range to those reported in other skin conditions, with similar levels of differentiation by severity / treatment response.[172–174]

B.3.4.1. Health-related quality-of-life data from clinical trials

HRQoL was measured in the ALLEGRO 2b/3 study via the disease-specific AAPPO (as reported in Section B.2.2.1) and generic HRQoL instruments: EQ-5D-5L, EQ VAS, SF-36 and HADS. A summary of the HRQoL outcomes is reported in the following sections.

B.3.4.1.1 EQ-5D-5L

Figure 30 shows the baseline EQ-5D-5L utility scores (for adults) and EQ-5D-Y (for adolescents) for all patients enrolled in the ALLEGRO 2b/3 study with EQ-5D valued according to different country-specific tariffs. The histograms for the UK, along with other countries, highlight a skew in the baseline utility values towards one, the upper bound of utility index scores. The baseline EQ-5D utility weights in patients were similar to that of population norms.[175]

Figure 31 and Figure 32 shows the scores across each of the five dimensions of EQ-5D-5L for ritlecitinib 50 mg and placebo, respectively, across baseline, Week 24, and Week 48. For each timepoint, more than [redacted] of respondents reported 'No problems' in all dimensions, which is the best possible health state. In the mobility, self-care and usual activity dimensions, 'No problems' were reported in more than [redacted] of responses. For the pain/discomfort dimension, 'No problems' were reported in more than [redacted] of responses, whilst 'No problems' were reported for the anxiety and depression dimension in over [redacted] of responses.

Together, these results demonstrate that, according to the EQ-5D-5L questionnaire, perfect health was reported in a large proportion of patients.

Figure 30: Adult Baseline Utility Weight Histogram: a) United Kingdom EQ-5D-5L (DSU/Hernandez-Alvarez) b) United States EQ-5D-5L c) Slovenia EQ-5D-Y d) Germany EQ-5D-Y e) Netherlands EQ-5D-Y

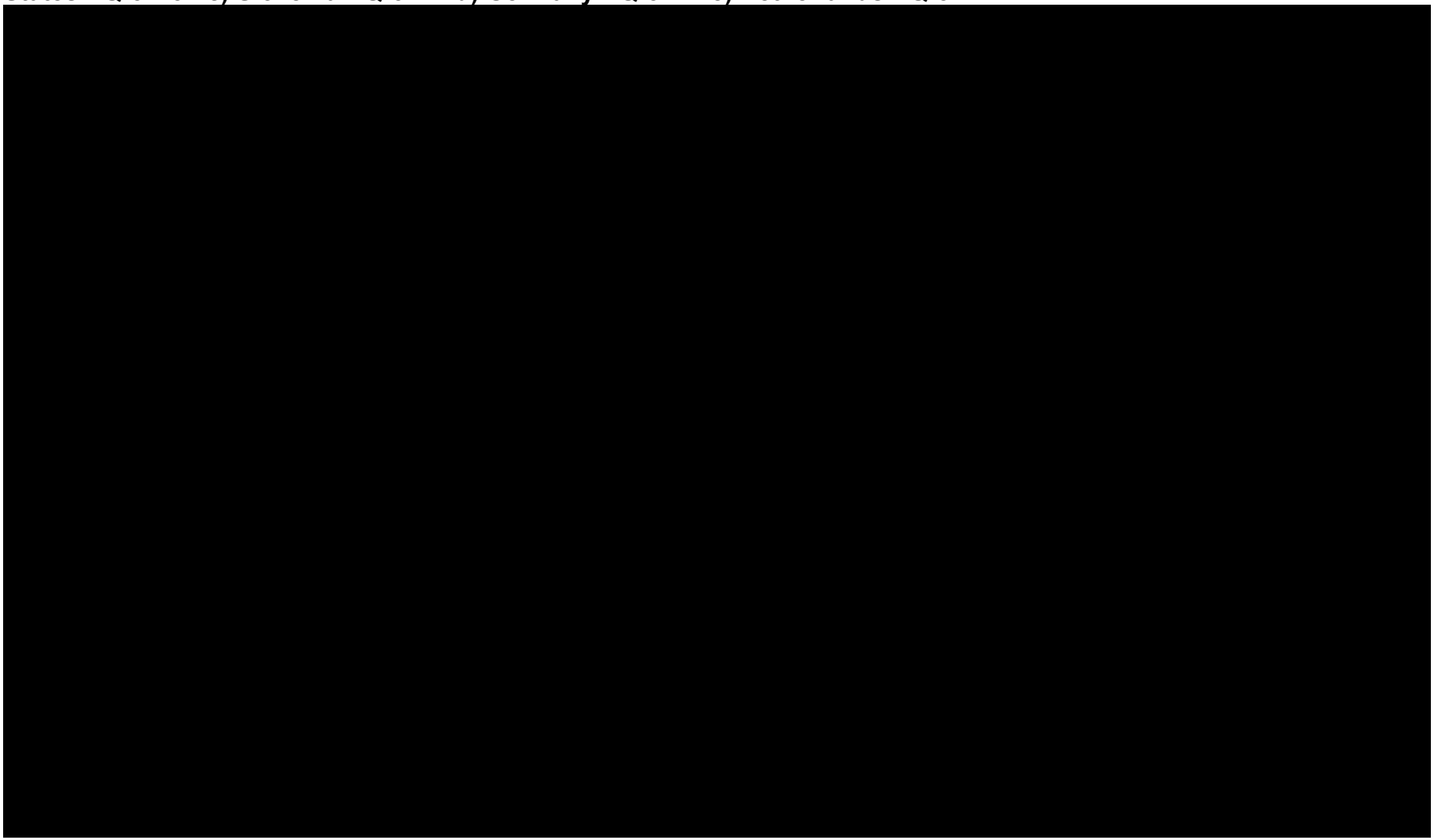


Figure 31: Ritlecitinib 50 mg EQ-5D-5L score by dimension

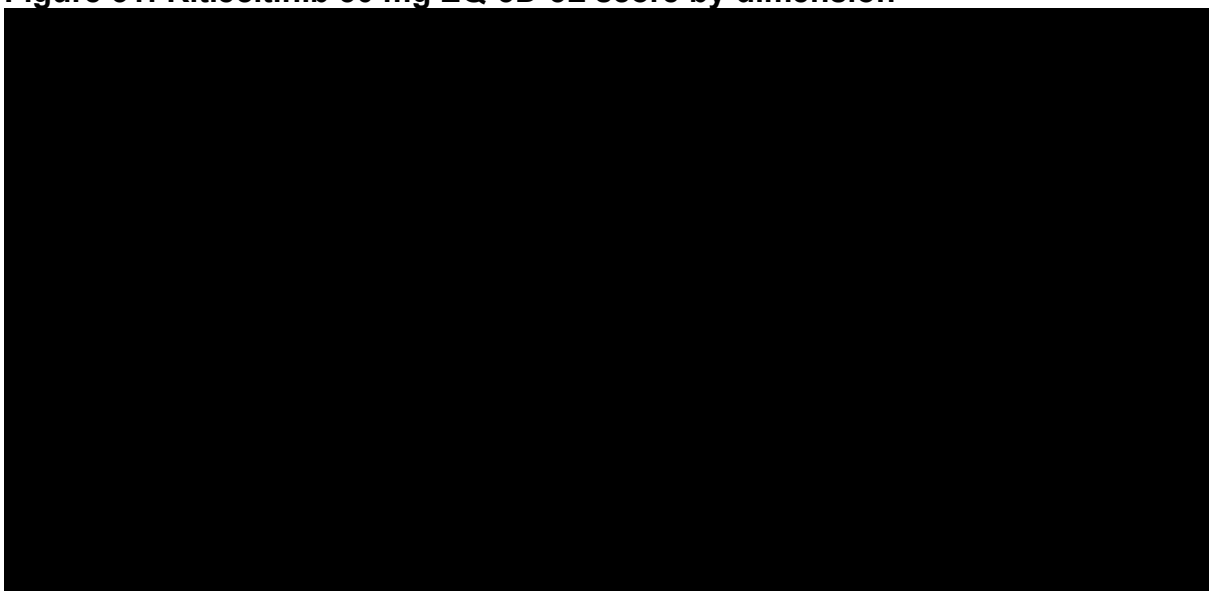


Figure 32: Placebo EQ-5D-5L score by dimension

Table 39 provides a summary of the mean change from baseline in EQ-5D index value by treatment group up to Week 48. Across all treatment groups, there is a marginal improvement in EQ-5D utility scores for patients with a SALT score ≤20 compared to patients with smaller improvements or worsening in patients with a SALT score >20. This demonstrates a pattern of improved outcomes with EQ-5D according to response, but this improvement is obstructed by high EQ-5D scores at baseline.

Table 39: Summary of Change from Baseline in EQ-5D Index Value by Treatment Group, Time Point and SALT Response (≤ 20 and >20) - Full Analysis Set (Overall)

Cohort	Ritlecitinib 200/50 mg (N=132)	Ritlecitinib 200/30 mg (N=130)	Ritlecitinib 50 mg (N=130)	Ritlecitinib 30 mg (N=132)	Ritlecitinib 10 mg (N=63)	Placebo (N=131)
Week 24 SALT score ≤20						
Participants with SALT ≤20 Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (Min, Max)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 24 SALT score >20						
Participants with SALT >20 Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (Min, Max)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 48 SALT score ≤20						
Participants with SALT ≤20 Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Median (Min, Max)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Week 48 SALT score >20						
Participants with SALT ≤20 Response, n	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Median (Min, Max)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	

Abbreviations: CI, confidence interval; Max, maximum; Min, minimum; mg, milligrams; SALT, Severity of Alopecia Tool; SD, standard deviation

When assessing utility values according to different SALT scores, a trend in improved utility scores according to lower SALT scores was observed. The EQ-5D-5L score for SALT 0-10 was [redacted] and for SALT 100 was [redacted]. The degree of difference estimated as an effect size using Cohen's d statistic was interpreted as small to medium effect (Cohen's d=0.24).[176]

The EQ-5D-Y showed an even smaller effect size, and with counter-intuitive increases, but was also based on less data. EQ-5D-Y score for SALT 0-10 was [redacted] and for SALT 100 was [redacted] (d=-0.03).[176]

B.3.4.1.2 EQ VAS

As previously shown with the EQ-5D-5L, the absolute EQ VAS scores (shown in Table 40) in the ALLEGRO 2b/3 trial show little differentiation between baseline, Week 24 and Week 48 for all treatment groups. EQ VAS asks patients to rate their current overall health status of a vertical visual analogue scale from 0 (worst imaginable health) to 100 (best imaginable health). Baseline EQ VAS scores were high, with a mean score >80 in all treatment groups. Median (interquartile range) EQ VAS scores at baseline ranged from [redacted] for ritlecitinib 50 mg to [redacted] for placebo, demonstrating limited range in the EQ VAS scores at baseline. At Week 24 and Week 48, there were nominal changes in the EQ VAS scores of all treatment groups. When considering change in EQ VAS according to treatment outcomes, patients who had achieved a SALT score ≤20 had a statistical significant improvement in their EQ VAS score of [redacted] (SD, [redacted]; 95% CI, [redacted]) at Week 24; at Week 48 there was a numerical improvement of [redacted] points increase in EQ VAS but this was not significant (SD, [redacted]; 95% CI, [redacted]).

Table 40: Absolute EQ VAS scores reported in ALLEGRO 2b/3 study

		Ritlecitinib 200/50 mg (N=112)	Ritlecitinib 200/30 mg (N=111)	Ritlecitinib 50 mg (N=112)	Ritlecitinib 30 mg (N=112)	Ritlecitinib 10 mg (N=54)	Placebo-> Ritlecitinib 200/50 mg (N=55)*	Placebo-> Ritlecitinib 50 mg (N=57)*
Baseline	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SD, standard deviation

B.3.4.1.3 SF-36

The SF-36 is a 36-item survey that assesses four subscales of physical health and four subscales of mental health; after evaluation of the responses to the 36 items, each subscale is allocated a score of 0-100. A summary of the absolute SF-36 scores by domain reported in the ALLEGRO 2b/3 study are presented in Figure 33 and

Figure 34; they show little change in scores by domain between baseline and Week 48 for patients treated with ritlecitinib and placebo.

Figure 33: Ritlecitinib 50 mg SF-36 scores by domain

The SF-36 tool is a 36-item survey that assesses four subscales of physical health (physical functioning, role physical, bodily pain, and general health) and four subscales of mental health (vitality, social functioning, role emotional, and mental health).[177] The results are summed (0-100) with lower scores suggesting worse HRQoL.

Figure 34: Placebo SF-36 scores by domain

The SF-36 tool is a 36-item survey that assesses four subscales of physical health (physical functioning, role physical, bodily pain, and general health) and four subscales of mental health (vitality, social functioning, role emotional, and mental health).[177] The results are summed (0-100) with lower scores suggesting worse HRQoL.

The physical and mental component summary scores of SF-36 (as summarised in Table 41 and Table 42) are similar across all arms at baseline and small differences are seen between baseline, Week 24 and Week 48. However, no change observed is greater than the minimally important change of 3 points. There also does not appear to be a trend in change from baseline in component summary scores according to whether ritlecitinib or placebo was allocated.

As with the EQ-5D results, the SF-36 showed very little differentiation between SALT score groups. For the Physical component (Table 41), summary scores were [redacted] (SALT 0-10) and [redacted] (SALT 100). For the Mental component (Table 42), summary scores were [redacted] (SALT 0-10) and [redacted] (SALT 100), which was a slightly larger effect size, but still classified as small (d=0.19).[176]

Table 41: Summary of absolute and change from baseline SF-36 scores reported in ALLEGRO 2b/3 study – Physical component summary

		Ritlecitinib 200/50mg (N=132)	Ritlecitinib 200/30 mg (N=130)	Ritlecitinib 50 mg (N=130)	Ritlecitinib 30 mg (N=132)	Ritlecitinib 10 mg (N=63)	Placebo-> Ritlecitinib 200/50 mg (N=65)*	Placebo-> Ritlecitinib 50 mg (N=66)*
Baseline	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

**Patients who began treatment on placebo crossed over to treatment with ritlecitinib in Week 24 of the ALLEGRO 2b/3 trial*
N: Number of Participants in FAS Population. n: Number of participants with observed data.
Baseline is defined as the latest non-missing value from the pre-treatment period

Table 42: Summary of absolute and change from baseline SF-36 scores reported in ALLEGRO 2b/3 study – Mental component summary

		Ritlecitinib 200/50mg (N=132)	Ritlecitinib 200/30 mg (N=130)	Ritlecitinib 50 mg (N=130)	Ritlecitinib 30 mg (N=132)	Ritlecitinib 10 mg (N=63)	Placebo-> Ritlecitinib 200/50 mg (N=65)*	Placebo-> Ritlecitinib 50 mg (N=66)*
Baseline	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

**Patients who began treatment on placebo crossed over to treatment with ritlecitinib in Week 24 of the ALLEGRO 2b/3 trial*
N: Number of Participants in FAS Population. n: Number of participants with observed data.
Baseline is defined as the latest non-missing value from the pre-treatment period

A post-hoc assessment was done comparing change in SF-36 scores according to whether patients achieved a SALT score of ≤20 (responder) or not. From this analysis, it was found that SALT ≤ 20 responders showed a greater mean improvement in change from baseline SF-36 Mental Component Summary (MCS) scores vs non responders at Week 24: [redacted] (95% CI, [redacted]) vs [redacted] (95% CI, [redacted]), respectively.[158] No differences were noted in the SF-36 Physical Component Summary (PCS) scores between SALT ≤ 20 responders and non-responders at Week 24.

When considering each of the subscales separately, SALT ≤20 responders demonstrated greater change from baseline scores in general health perception, vitality, social functioning, role emotional, and general mental health domains than the physical functional, role physical and bodily pain domains at Week 24 and Week 48. In the social functioning domain, mean change from baseline in the social functioning domain were close to a three-point difference, which is the minimally important change across SF-36 sub-scores.

Figure 35: ALLEGRO 2b/3: SF-36v2 Domain Score Change (Baseline to Week 24)[158]

Abbreviations: BP = bodily pain; GH = general health; MH = mental health; PF = physical functioning; RE = role emotional; RP = role physical; SALT = Severity of Alopecia Tool; SF = social functioning; SF-36 = Short Form 36; VT = vitality.

Figure 36: ALLEGRO 2b/3: SF-36v2 Domain Score Change (Baseline to Week 48)[158]

Abbreviations: BP = bodily pain; GH = general health; MH = mental health; PF = physical functioning; RE = role emotional; RP = role physical; SALT = Severity of Alopecia Tool; SF = social functioning; SF-36 = Short Form 36; VT = vitality.

B.3.4.1.4 HADS

HADS consists of seven questions relating to the Depression subscale and seven questions relating to Anxiety subscale with each question scored on a scale 0–3 to give a Depression and an Anxiety score out of 21. A HADS anxiety (HADS-A) subscale score of >7 for adults and >8 for adolescents is indicative of anxiety; a HADS depression (HADS-D) subscale score of >7 for adults and >6 for adolescents is indicative of depression.[143]

A summary of the HADS-D and HADS-A scores over time in the ALLEGRO 2b/3 study are provided in Table 43 and Table 44, respectively. Scores are similar across all arms at baseline and small differences are seen between baseline, Week 24 and Week 48. According to the thresholds which are indicative for depression and anxiety, few patients would be classified with either depression or anxiety as the inter-quartile ranges are contained below the upper bounds for each measure.

Table 43: Summary of absolute and change from baseline HADS-D scores reported in ALLEGRO 2b/3 study

		Ritlecitinib 200/50mg (N=132)	Ritlecitinib 200/30 mg (N=130)	Ritlecitinib 50 mg (N=130)	Ritlecitinib 30 mg (N=132)	Ritlecitinib 10 mg (N=63)	Placebo-> Ritlecitinib 200/50 mg (N=65)*	Placebo-> Ritlecitinib 50 mg (N=66)*
Baseline	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

**Patients who began treatment on placebo crossed over to treatment with ritlecitinib in Week 24 of the ALLEGRO 2b/3 trial*
N: Number of Participants in FAS Population. n: Number of participants with observed data.
Baseline is defined as the latest non-missing value from the pre-treatment period

Table 44: Summary of absolute and change from baseline HADS-A scores reported in ALLEGRO 2b/3 study

		Ritlecitinib 200/50mg (N=132)	Ritlecitinib 200/30 mg (N=130)	Ritlecitinib 50 mg (N=130)	Ritlecitinib 30 mg (N=132)	Ritlecitinib 10 mg (N=63)	Placebo-> Ritlecitinib 200/50 mg (N=65)*	Placebo-> Ritlecitinib 50 mg (N=66)*
Baseline	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 24								
Week 24	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Absolute values at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline at Week 48								
Week 48	N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Median (Q1, Q3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Min, Max	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

**Patients who began treatment on placebo crossed over to treatment with ritlecitinib in Week 24 of the ALLEGRO 2b/3 trial*
N: Number of Participants in FAS Population. n: Number of participants with observed data.
Baseline is defined as the latest non-missing value from the pre-treatment period

Of the [redacted] participants with a HADS score indicative of anxiety at baseline and who achieved SALT ≤10 at Week 24, [redacted] had a normal HADS anxiety score at Week 24; [redacted] participants with a HADS score indicative of anxiety at baseline had a SALT score >10 at Week 24, of whom [redacted] had a normal HADS anxiety score at Week 24.[133] Only [redacted] participants with SALT ≤10 at Week 24 had a HADS score indicative of depression at baseline and [redacted] participants did not have depression according to HADS by Week 24; [redacted] participants of [redacted] with SALT >10 at Week 24 who had a HADS score indicative of depression at baseline did not have a score indicative of depression after 24 weeks of treatment.[133] As there were limited participants with existing anxiety or depression at baseline, Week 24 and Week 48, care should be take not to over-interpret this data.

B.3.4.1.5 AAPPO

The AAPPO is a novel PRO tool that was developed by Pfizer which captures the consequences of AA of the highest priority to patients, consisting of three subscales: emotional symptoms, activity limitations and hair loss.[73] The 11-item AAPPO contains four hair loss items in which the patient is asked to describe the current amount of hair loss from (1) the scalp, (2) eyebrows, (3) eyelashes, and (4) body; using a five-point response scale that ranges from 0 (no hair loss) to 4 (complete hair loss). The remaining seven items ask the patient to rate emotional symptoms (self-consciousness, embarrassment, sadness and frustration about hair loss) and activity limitations (limiting outdoor activity, physical activity and interaction with others) of AA over the past week on a five-point scale ranging from “never” to “always”. Cognitive debriefing interviews with patients confirmed the content validity of the measure, which was developed in a manner consistent with regulatory guidance.[73]

In the ALLEGRO 2b/3 study, improvement in AAPPO score was measured for patients with AAPPO score ≥2 at baseline for each of the hair loss items where improvement was defined as achieving a score of 0 (no hair loss) or 1 (little hair loss). At Week 24 the number of patients reporting was statistically significantly higher in the 50 mg group for hair loss on the scalp [redacted], eyebrows [redacted] and eyelashes [redacted] compared to placebo (scalp: [redacted], $P < 0.001$; eyebrows:

[redacted], $P = 0.001$; eyelashes: [redacted], $P < 0.001$). For hair loss on the body, the number of participants reporting improvement was numerically greater in the 50 mg group [redacted] than in the placebo group ([redacted] $P = 0.244$).[154] After transitioning to ritlecitinib at Week 24, participants who had received placebo showed improvement in the four hair loss items to Week 48 (Figure 37); this was generally similar to the improvements observed in participants treated with the same regimen at Week 24.

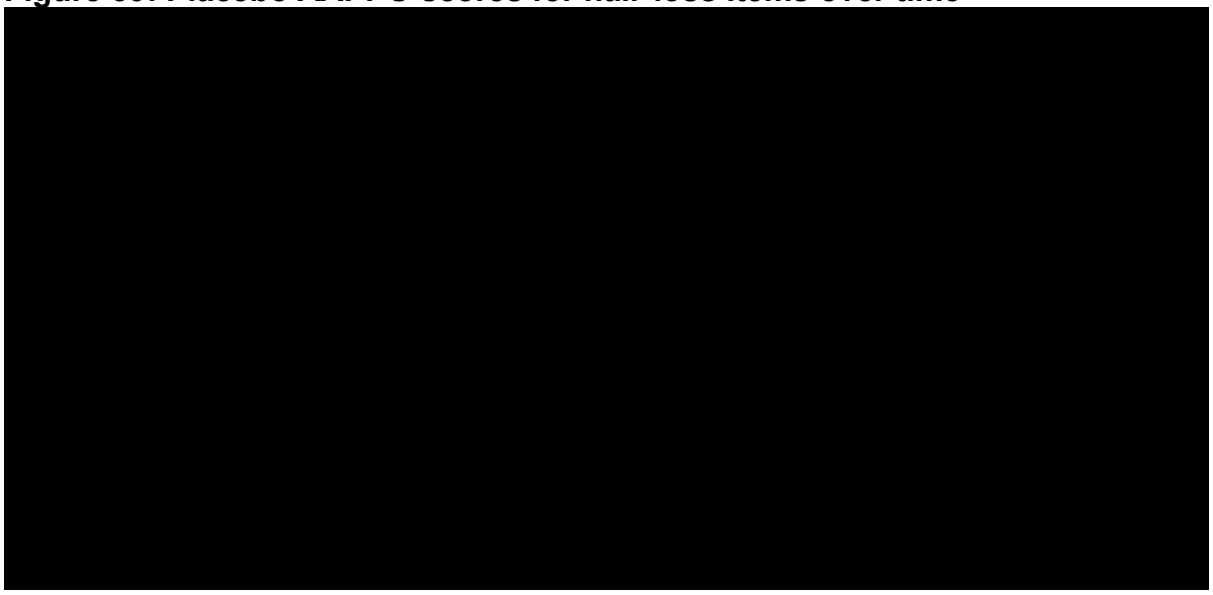
Figure 37: Improvement in AAPPO Hair Loss items from baseline up to Week 48 (among participants with score ≥2 at baseline in FAS)

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomisation.
Abbreviations: AAPPO; Alopecia Areata Patient Priority Outcomes; FAS, full analysis set; mg, milligram; S/E, standard error

The HRQoL results from each of the AAPPO hair-loss items are also presented in Figure 38 and Figure 39. The results demonstrate improvements in the quantity of hair loss across the different areas of the body for both ritlecitinib and placebo.

Figure 38: Ritlecitinib 50 mg AAPPO scores for hair-loss items over time

Figure 39: Placebo AAPPO scores for hair-loss items over time



Results over time from the ALLEGRO 2b/3 clinical trial for the AAPPO emotional symptoms and activity limitations over time are presented in Figure 40 and Figure 42. The scores at baseline for the activity domains for both ritlecitinib 50 mg and placebo at baseline are low, indicating lack of impact of the domain to patients' quality of life. Similarly, the average score for emotional symptoms at baseline was [redacted] for ritlecitinib

and placebo, suggesting that most recruited patients within the ALLEGRO 2b/3 trial are not largely impacted by these domains.

Figure 40: Ritlecitinib 50 mg AAPPO scores for emotional symptoms and activity limitations over time

Emotional symptoms: Mean of items 5,6,7,8 on the AAPPO. Activity limitations: Mean of items 9, 10, 11 on the AAPPO. Responses range from 0 to 4 with higher scores suggesting worse HRQoL.
Abbreviations: AAPPO, Alopecia Areata Patient Priority Outcomes

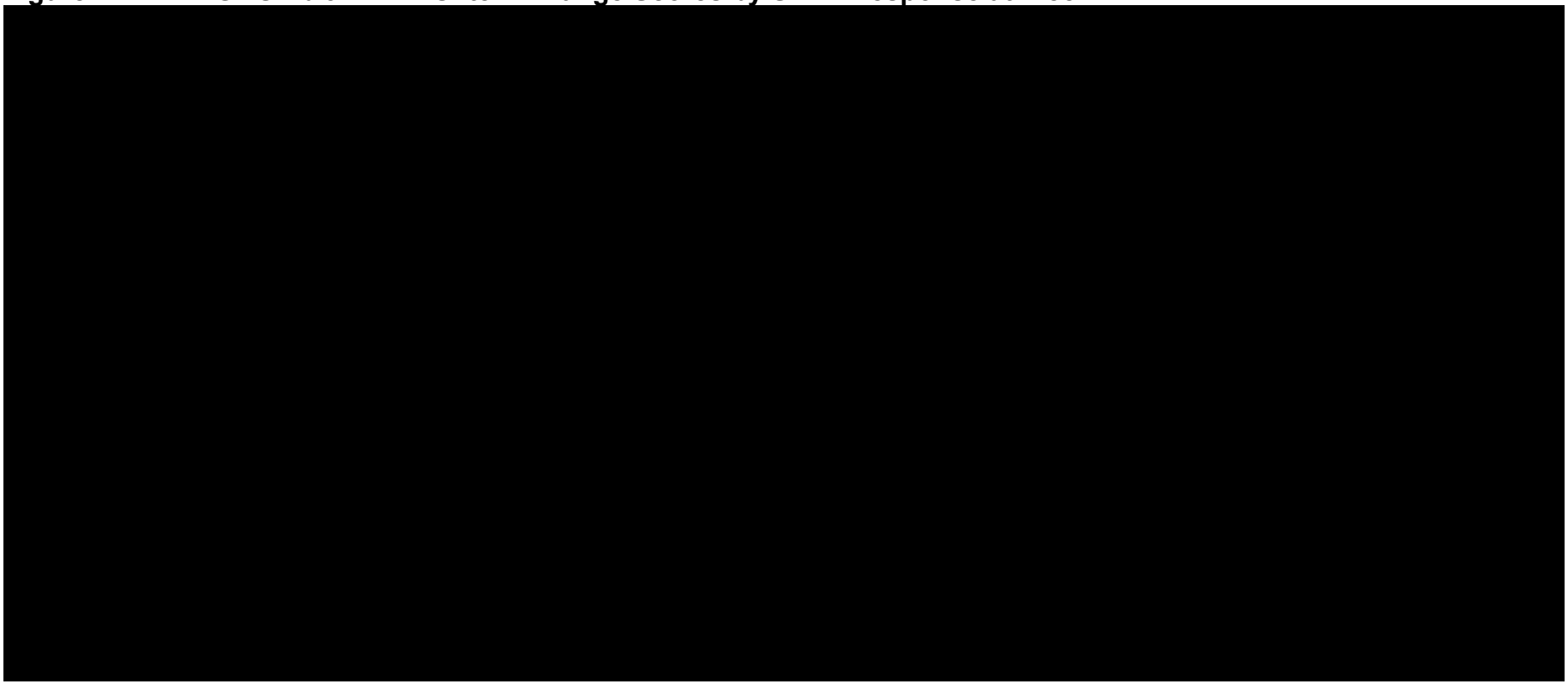
Figure 41: Placebo AAPPO scores for emotional symptoms and activity limitations over time

Emotional symptoms: Mean of items 5,6,7,8 on the AAPPO. Activity limitations: Mean of items 9, 10, 11 on the AAPPO. Responses range from 0 to 4 with higher scores suggesting worse HRQoL.
Abbreviations: AAPPO, Alopecia Areata Patient Priority Outcomes

To assess how the domains were impacted in patients who achieved a SALT score of ≤20 (responders), the mean change in each domain was assessed. At week 24, the largest changes were observed in patients according to the hair loss items, with nominal improvements seen in patients who responded to treatment in the activity limitations domain (Figure 42). The similar trends were observed at Week 48 (Figure 43), with greater increases in domain scores for responders compared to non-responders.

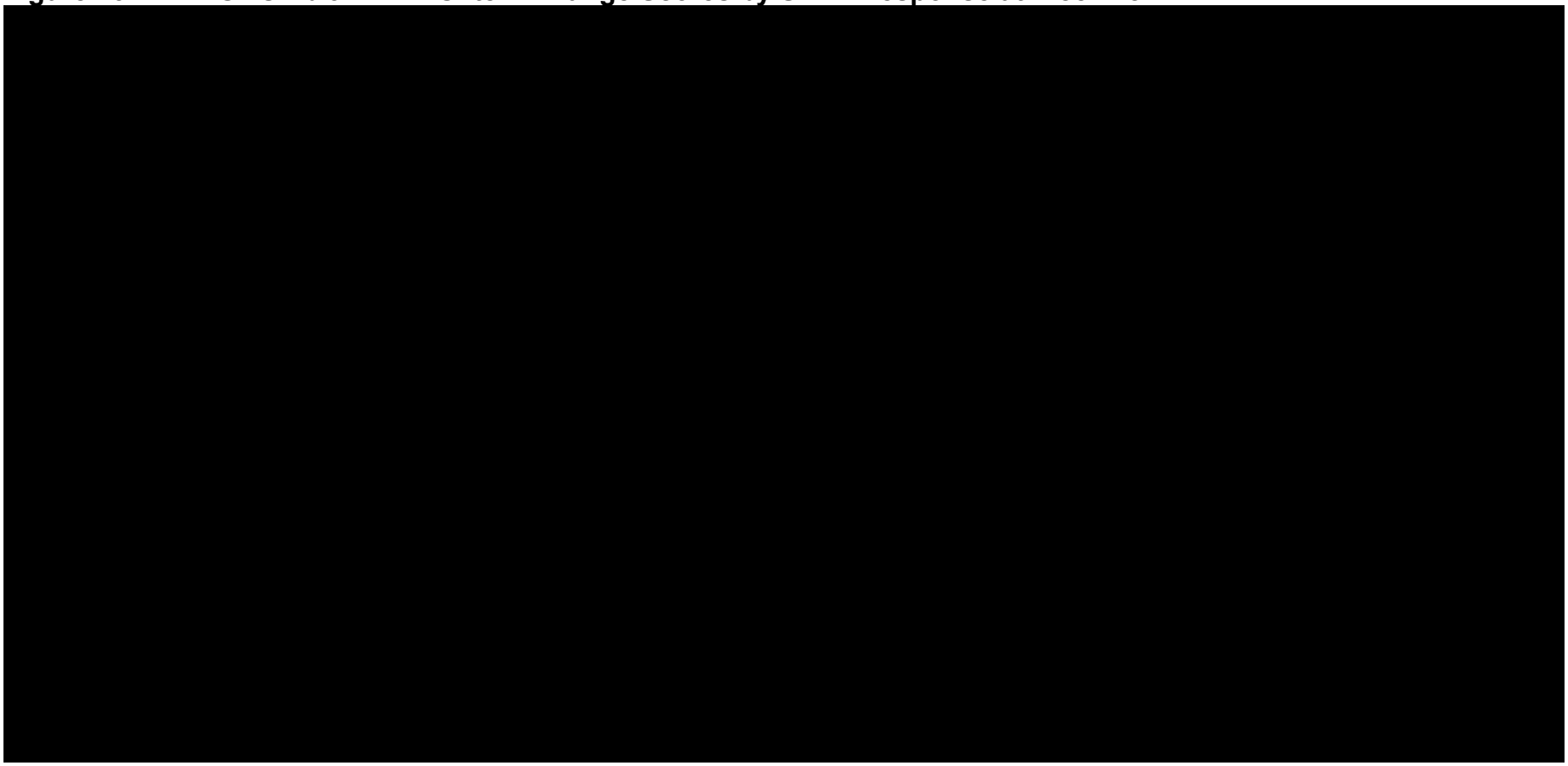
The AAPPO scores on both dimensions (emotional symptoms and activity limitations) show worse scores for people with more severe hair loss. The largest difference (and largest effect size) between groups was seen for the AAPPO emotional symptoms score. The effect size estimate suggests that this is a large effect (d=0.79).[176] The AAPPO Activity limitations score was associated with a moderate effect size (d=0.48) between SALT 0-10 and SALT 100.[176] The results suggest AAPPO is more sensitive than EQ-5D and SF-36 to variations in HRQoL in patients with AA.

Figure 42: ALLEGRO 2b/3: AAPPO Item Change Scores by SALT Response at Week 24[158]



Abbreviations: AAPPO = Alopecia Areata Patient Priority Outcomes; SALT = Severity of Alopecia Tool.

Figure 43: ALLEGRO 2b/3: AAPPO Item Change Scores by SALT Response at Week 48[158]



Abbreviations: AAPPO = Alopecia Areata Patient Priority Outcomes; SALT = Severity of Alopecia Tool.

B.3.4.1.6 Summary and interpretation of HRQoL from the ALLEGRO 2b/3 study

A substantial body of evidence exists that describes the burden that AA can have on patients' HRQoL. However, this was not completely reflected in the results from the ALLEGRO 2b/3 pivotal clinical trial.

In the ALLEGRO 2b/3 trial, significant changes in EQ-5D-5L or SF-36 were not seen between the ritlecitinib 50 mg group and placebo from baseline to Week 24. A possible explanation for this is that generic measures of HRQoL, such as EQ-5D-5L or SF-36, may not capture all the domains which are relevant to patients with AA and are therefore insensitive to HRQoL in this patient population. Whilst EQ-5D-5L and SF-36 can capture the depression and mental health of patients, respectively, other important aspects of HRQoL for patients are not captured. Therefore, both instruments lack content validity and potential responsiveness to changes in HRQoL for patients with AA.

AAPPO has been shown to be a reliable and valid condition specific patient reported outcome measure and appears to be more sensitive to the burden for patients with AA. However, results from the ALLEGRO 2b/3 trial showed mixed results. Whilst hair loss domains in AAPPO showed greater difference for patients, there was a positive but smaller trend in emotional and activity limitation domains.

A responder analysis of SF-36 and AAPPO outcomes was conducted that demonstrated a pattern of improved outcomes. The SF-36 analysis showed meaningful hair regrowth was associated with improvements in the emotional and psychosocial domains of HRQoL. The responder analysis for AAPPO showed a clear separation between responders and non-responders, particularly in the emotional and activity limitation domains with a consistent trend seen through to Week 48. This is more aligned with the described burden of disease in Section B.1.3.2.

Content validity is not the only consideration when interpreting these results. Most patients completing the EQ-5D-5L questionnaire in ALLEGRO 2b/3 reported the same score (no problems) across all domains at baseline, allowing no room for improvement up to Week 48, i.e., there is a high ceiling effect. Similarly, patients completing the EQ-VAS had similar, high mean score at baseline through to Week 48. The absolute SF-

36 scores reported in ALLEGRO 2b/3 study also showed a similar, high mean score at baseline through to Week 48. This is consistent with the results seen for AAPPO at baseline. Patients did not report frequent emotional symptoms and activity limitations in response to the relevant domains in the ALLEGRO 2b/3 study, limiting room for improvements to be observed. A similar observation is seen across anxiety and depression measured via HADS from ALLEGRO 2b/3 (Section B.3.4.1.4) which showed high baseline scores for these measures.

Given that the patients enrolled in ALLEGRO 2b/3 had AA for multiple years (mean 10.1 years since first diagnosis), this insensitivity may also be due to adaptation by patients who were enrolled to the study. That is, patients enrolled in the study may have learnt to cope with the way AA impacts their HRQoL. Across several studies, coping mechanisms are mentioned which offset the negative impact of AA on patients HRQoL the most extreme of which, after all other behaviours had been tried out, is acceptance.[11,167–171]

The higher HRQoL scores at baseline could also be due to eligibility criteria for the clinical study; patients with with "major psychiatric conditions" were excluded from the study (see clinical data limitations in Section B.2.13.3.2). The placebo-controlled period might also be too short to detect a meaningful difference (24 weeks).

Finally, the extent to which the heterogeneity of the condition (as outlined in B.1.3.2.3) impacts the interpretation of these results is unclear. Taken together the results from ALLEGRO 2b/3 are not aligned with the described burden of disease in Section B.1.3.2.

In developing the AAPPO PRO, a validation study was conducted in which scoring and measurement properties of the AAPPO were examined using baseline and 2-week follow-up data from a prospective study of 121 patients with AA.[178] This study revealed other potential reasons for only modest changes in the emotional symptoms and activity limitation scores in the ALLEGRO 2b/3 study.[178] As expected, in the validation study, the four AAPPO hair loss item mean scores were better for patients in the 25-49% SALT tertile compared with those in the highest SALT tertile (76-100%; $P < 0.01$). However, AAPPO emotional symptoms and activity limitations mean domain scores tended to be worse for participants in the 25–49% and 50-75% SALT

tertiles compared with those who had highest SALT scores (76-100%). This could be evidence of adaptation to the condition as mentioned above, or it could be due to greater daily emotional and activity-limiting burden to cosmetically conceal and manage smaller patchy areas of hair loss compared with patients with far greater or complete scalp hair loss (AT/AU).

Whilst the burden of AA is linked to hair loss, there is inconsistency in how this links with the extent of hair loss. AA appears to impact patients differently linking to the heterogeneity of the condition. Our engagement with PAGs suggested similar inconsistency stating impacts experienced by people with alopecia are the same, regardless of the subtype or extent of hair loss experienced. However, there were also comments that patients with minimal hair loss may experience fewer impacts and some felt the loss of identity experienced by people with AA increased with the extent of hair loss. This was particularly associated with loss of the eyebrows and eye lashes with AT or AU.

The PAG observations are supported by clinical feedback from the Therapeutic Delphi Panel in which 88% agreed that achieving a clinically meaningful response to treatment was associated with a meaningful improvement in HRQoL. However, in the interviews, clinicians reflected on both of these areas. The across patient variability was described; for example, one clinician stated that "*people with AT might be happy and not bothered with anything, while people with one or two patches would think their life is ruined*". In addition, for an individual patient, clinicians stated that generally there is a correlation between the extent of hair loss and HRQoL, i.e., patients with more hair loss would have a lower HRQoL. However, it was emphasised that the correlation is perhaps not always linear and can depend on a number of factors including the patient's acceptance of the disease, the location of the hair loss (i.e., whether it can be hidden) and how it affects their social life and hobbies. These findings alongside the results of the AAPPO validation and ALLEGRO 2b/3 PRO results highlight the complexity of correctly estimating the burden of HRQoL in AA.

Overall, despite ritlecitinib showing meaningful impacts on hair regrowth with an acceptable safety profile (Section B.2), HRQoL results reported using EQ-5D and SF-36 in the ALLEGRO 2b/3 trial do not adequately capture the detrimental burden

patients with AA experience, as described in B.1.3.2. Responder analysis of AAPPO suggest that as more HRQoL domains are incorporated, sensitivity to changes in HRQoL improves. We have outlined a number of potential reasons why HRQoL results were not accurately capturing the true burden of AA, including the potential impact of trial design and the extent to which generic and disease specific measures of HRQoL such as EQ-5D and SF-36 are insensitive in capturing the domains and nuances associated with the burden of AA on patients. Sections B.3.4.3 and B.3.4.4 outline the further steps we have taken to understand these observations further.

B.3.4.2. Mapping

EQ-5D-5L was measured directly in the ALLEGRO 2b/3 study, so mapping is not required to generate utilities.

The SF-36 instrument utilised for the ALLEGRO Phase 2b/3 trial, as discussed in Section B.3.4.1, is also impacted by potential trial design and or content validity concerns which does not adequately capture change in utility for patients with AA, and so has not been mapped to SF-6D.

The AAPPO is not a preference-based measure, so utilities cannot be derived, even indirectly. Given the insensitivities of preference-based measures such as the EQ-5D-5L and SF-36 (described in Section B.3.4.1.6), mapping the AAPPO to existing utility measures was not a viable option.

B.3.4.3. Health-related quality-of-life studies

An SLR was undertaken to identify and summarise the best available HRQoL evidence available for the treatment of AA, the methodology is summarised in Appendix H. The objective of the SLR was to assess the HRQoL and utility of patients with AA from interventional or RWE studies. Searches were performed in October 2021.

Three publications reporting utility data were identified and reported utility values based on HRQoL scales, including the EQ-5D-5L and Assessment of Quality of Life 8 dimensions (AQoL-8D):

- EQ-5D-5L results showed that quality of life decreased with increasing severity (mild 0.95 [0.14] vs moderate 0.93 [0.13] vs severe 0.87 [0.21]).[98] However, the

decrement were small despite using 5L, which is more sensitive to changes in HRQoL. Given the insensitivities of the EQ-5D as shown in B.3.4.1, these utility values were not suitable to be used in the cost-effectiveness model.

- The AQoL-8D scale is a generic instrument which enables comparison across diseases measures on a scale from 0 (death) to 1 (full health) meaning positive values reflect an improvement in HRQoL. Mean (SD) overall Assessment of Quality of Life-8 (AQoL-8D) score for patients with AA was 0.748 (0.206) at baseline.[126] At 3 months, a patient group treated with ciclosporin showed a trend greater improvement in HRQoL across 6 of 8 AQoL-8D dimensions, compared to those treated with placebo. However, results were not significantly different. This study was not suitable for use in the model as aggregate utilities for moderate to severe patients with AA were described, which are not relevant to the model.

- For patients with AA with<50% reduction in SALT score to ciclosporin in a preceding trial who were treated with sublingual tofacitinib in an open-label, roll-over clinical trial, change in HRQoL was assessed using the Assessment of AQoL-8D score. The mean change from baseline in Assessment of AQoL-8D score was -0.0148 (0.0515).[179] These utility values were not suitable for use in the cost-effectiveness model as only a mean reduction from baseline on treatment was described without disaggregation by SALT score.

The published utility values are presented in Table 45.

Table 45: Identified studies reporting utility scores

Study (year)	Country	Reported health state	Patient population	Data source	Average utility score (SD)
Burge *et al.* (2021)[98]	USA	Mild AA	Patients with mild AA	Measured with EQ5D-5L	0.95 (0.14)
		Moderate AA	Patients with moderate AA	Measured with EQ5D-5L	0.93 (0.13)
		Severe AA	Patients with severe AA	Measured with EQ5D-5L	0.87 (0.21)
	Australia	Moderate-to-severe AA at baseline	Moderate-to-severe,	Assessment in AQoL-8D	0.748 (0.206)

Lai *et al.* (2019)[126]			adults with AA		
		AT/AU	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.732 (0.2560)
		Patchy AA	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.773 (0.127)
		Female AA	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.738 (0.212)
		Male AA	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.791 (0.174)
		AA treated with placebo at 1 months	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.787
		AA treated with ciclosporin at 1 months	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.803
		AA treated with placebo at 2 months	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.781
		AA treated with ciclosporin at 2 months	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.820
		AA treated with placebo at 3 months	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.803
		AA treated with ciclosporin at 3 months	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	0.806
Lai *et al.* (2021)[179]	Australia	Non-responders to ciclosporin with moderate-to-severe AA & treated with sublingual tofacitinib	Moderate-to-severe, adults with AA	Assessment in AQoL-8D	Change from baseline: -0.0148 (0.0515).

Abbreviations: AA, Alopecia Areata; AT, Alopecia Totalis; AU, Alopecia Universalis; AQoL-8D, Assessment of Quality of Life 8 dimensions; EQ5D-5L, EQ-5D 5 levels; SD, standard deviation

Two studies in the literature identified through a targeted literature search evaluated the relationship between AA severity and EQ-5D based utilities. Bewley *et al.* (2022) report EQ-5D scores ranging from mild AA (0.90) to severe AA (0.78) among patients

from 5 European countries.[180] A similar study in the USA, reported EQ-5D scores for mild AA (0.95) and severe AA (0.87).[98] These studies did not include SALT scores to grade patients. It is possible that grading severity based on clinician judgment produced bias (that would be expected to result in larger decrements in overall HRQoL). However, the observed range of values across severity further support the EQ-5D lacking sensitivity in this population.

The DLQI is another measure commonly used in clinical practice to assess the HRQoL in patients with AA. However, a review of existing studies found that all studies which have used the DLQI only reported score data and had not used a mapping function to convert this data into utility values.[75, 78, 80, 92, 102, 104–108,181–185] Furthermore, all items included in the DLQI ask patients to reflect on how their skin problems (rather than hair) are affecting them. This makes it very difficult to interpret findings from the literature if modification is not specified.

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

B.3.4.4.1 Findings from the clinical trial HRQoL

As discussed in Section B.3.4.4.1, the clinical trial HRQoL data from the ALLEGRO 2b/3 trial does not demonstrate meaningful improvements in HRQoL and does not align with how the burden of AA impacts patients' mental and physical health as reported in Section B.1.3.2. Furthermore, mapping of AAPPO is also not viable as discussed in Section B.3.4.2. Given the EQ-5D data in the literature (Section B.3.4.3), along with what we have learned from our clinical trial (B.3.4.1) and the potential insensitivities as first outlined in B.1.3.2, we explored feedback from patients to gain a better understanding of utility estimates for AA patients.

B.3.4.4.2 Important elements of HRQoL in AA

Evidence exists that describes the substantial burden that AA can have on people. This includes the psychosocial burden but also a broader impact in terms of people's willingness to undertake daily activities. This evidence clearly suggests that the burden

and impact of AA extends beyond simply hair loss resulting in a much wider impact on HRQoL.

To understand the disconnect between the ALLEGRO 2b/3 HRQoL trial results and data on the burden of AA patients from the literature, nine PAG representatives provided feedback on the EQ-5D and advice was sought from clinicians on the SF-36 and EQ-5D via the Therapeutic Landscape Delphi panel.

Nine PAG representatives provided feedback on EQ-5D of which [redacted] said it is not a good measure for the assessment of HRQoL in adult patients with AA because it is too broad to capture the specific impacts faced by patients with AA in a meaningful way.[22] [redacted] participants added that if somebody with AA were to complete the EQ-5D, they would likely score within the normal range because the questions are not tailored suitably to this population. Similarly, of the three participants who were asked about EQ-5D-Y, [redacted] indicated that EQ-5D-Y is not a good measure for the assessment of HRQoL in adolescent patients with AA, citing similar reasons for why EQ-5D is not a good measure, namely, that the instrument lacked the specificity to accurately capture all relevant impacts of living with AA to HRQoL.

Advice from clinical experts in the UK was also sought via the Therapeutic Landscape Delphi panel to explore the use of SF-36 and EQ-5D for patients with AA.[21] Seven of the eight experts consulted said they thought the EQ-5D and SF-36 are not suitable to assess the HRQoL of patients with AA (7/8 clinicians, 88%). In particular, it was noted that EQ-5D is not suitable for adults or adolescents with severe AA (whom they would treat with systemic therapy) because it does not cover all impacts experienced by adults/adolescents with severe AA, it is too generic and it contains irrelevant domains such as physical activities such as dressing and walking.[21] All dermatologists with a specialist interest in hair disorders agreed that the mobility, pain and self-care domains in the EQ-5D were irrelevant for patients with AA (8/8 clinicians, 100%). Key elements of AA that are not captured from generic measures such as the EQ-5D are described in Figure 44.

For example, one clinical expert said:

"*EQ-5D was in my view utterly unsuited to assessment of AA. It focussed on physical activities such as dressing and walking rather than detailed impact on mental health- I've never had a patient with AA who has had trouble doing activities like dressing or walking but every single one has had significant problems with self-esteem, anxiety, low mood, reduced libido and a fear of lack of attractiveness – none of which was addressed in the scale other than "depression" which is far too broad a term for the problems encountered and which are after all the whole reason why we treat these patients.*"

Similarly, another clinical expert said:

"*It does not enquire about the type of activities that are impacted by AA, and those that are asked about are not relevant. 'Usual activities' is not specific enough. The health question [relating to usual activities] is ambiguous for an AA sufferer*."

The experts added that it does not cover important aspects of HRQoL including self-esteem, social impacts, happiness, impact on work-life for adults and bullying for adolescents. For example, when asked if any aspects of HRQoL are missing from EQ-5D which would be important to capture for patients with AA with ≥50% hair loss (including AT/AU), one clinician said:

"*AA does not physically limit patients' activities in the same way a painful condition like arthritis would. The main impact of the condition is psychological. I would want to know whether patients can live their lives normally i.e., I do / do not feel like leaving the house in the morning; I do / do not feel I can face my school friends / work colleagues / clients. I do / do not feel like I can have intimate / sexual relations with my partner etc.*"

Key elements of HRQoL for patients with AA are omitted from the EQ-5D questionnaire, based on feedback from PAG representatives (Figure 44).[22] These gaps were in the domains of social functioning, relationships, emotional, physical, appearance and financial (Figure 44). In the omission of these elements of HRQoL which are important to patients with AA, the EQ-5D lacks content validity.

Figure 44: Aspects of HRQoL not covered by the HRQoL generic measures[22]

Abbreviations: EQ-5D-Y, EQ-5D-Youth; EQ-5D-3L, EQ-5D-3-Level; PAG, patient advocacy group; QoL, quality of life

Although AA has been shown to significantly negatively impact HRQoL across most dimensions of the SF-36 questionnaire in French patients (n=60) compared to the general population, particularly in terms of patient mental health and vitality, results using generic measures such as the SF-36 are inconsistent.[96] Similarly, it has been reported that a range of factors severely impact social functioning in patients with AA not included in the SF-36, including the negative effect of appearance alterations related to AA, being self-conscious, social anxiety, barriers to taking part in physical activities/exercise and stigma.[72, 100, 103,186] As previously recognised, however, given the limitation of SF-36 for measuring HRQoL in patients with AA, the differentials in the studies are not considered representative of patient experience and likely underestimate the impact on HRQoL. Moreover, given the domains reported to be missing from the EQ-5D in Figure 44, there are also important domains not included in the SF-36, demonstrating that it also lacks content validity as a tool to value HRQoL of patients with AA.

The DLQI is another measure commonly used in clinical practice to assess quality of life in patients with AA. Seven cross-sectional studies from different countries showed

that greater hair loss was significantly associated with poorer HRQoL based on DLQI scores which suggests that the DLQI may be more sensitive to the HRQoL impacts of AA compared to other preference-based measures, such as the EQ-5D and SF-36.[75, 80, 102,104–107] For example, one cross-sectional study of patients with AA recruited from a hospital in Belgrade (n=60) found significant differences were observed for all DLQI domains by severity of scalp hair loss and proportion of total body hair loss.[80]

However, the aim of the DLQI questionnaire is to measure how much a skin problem has affected a patient's life over a weekly period, so it was not designed with AA in mind specifically, and as such was not used to collect HRQoL data in the ALLEGRO 2b/3 trial. Furthermore, DLQI refers to "skin" or "skin condition" in most of the items, which may contribute to some insensitivity to AA, which is predominantly a hair loss disease. Clinical experts commented via the Therapeutic Landscape Delphi panel that the DLQI/Children's Dermatology Life Quality Index (cDLQI) is more suitable than the EQ-5D (8/8 clinicians, 100%) and SF-36 (7/8 clinicians, 88%) and that the DLQI is considered as a good way to assess HRQoL in patients with AA (6/8 clinicians, 75%). However, there is no mapping from DLQI that is able to generate utilities. Moreover, four dermatologists with a specialist interest in hair disorders (4/8 clinicians, 50%) stated that it does have some limitations; with some domains remaining irrelevant, and as it does not capture the full emotional impact that AA has on patients (i.e., lack of construct validity), thus a more specific measure would be superior.

Given the limitations of the EQ-5D, the SF-36 and to a lesser extent DLQI/cDLQI to assess the HRQoL for patients with AA, the AAPPO tool was developed by Pfizer which has been shown to be the most sensitive tool to quantify the HRQoL of patients with AA. However, AAPPO cannot be used to determine HRQoL as it is not a preference-based measure.

There is a lack of construct validity for patients with AA given generic preference-based measures and condition specific preference-based measures are not appropriate either. Given this, there is a clear need for health state utilities which adequately describe the impacts that AA has on the HRQoL of patients in line with NICE guidance. Therefore, the logical next step was to try to connect what we have learned from the literature and our own clinical trials with what we have heard from

patients and clinicians. Consequently, based on the hierarchy of preferred HRQoL methods from NICE guidance,[187] a vignette methodology to derive utilities is the most appropriate approach for the submission.

B.3.4.4.3 Vignette study

Vignette study methodology

A vignette study was designed, in accordance with the hierarchy of preferred HRQoL methods in the NICE guidance,[187] to capture the full impact on HRQoL for patients who suffer from AA. Health state vignettes which are designed to describe patients' HRQoL in different stages of the disease can be developed, validated and then assessed using TTO methods. This approach allows utilities to be estimated for specific AA health states for different levels of hair loss and disease stages.

The impact of AA on caregivers of adolescents was also explored within the vignette study resulting in the implementation of a caregiver disutility for patients with a SALT score >50.

The study design consisted of three main parts.[23] In part 1, vignettes were designed to describe how key domains of HRQoL are affected by the disease for adult and adolescent patients and caregivers of patients with AA. These were informed by findings from three different sources:

1. Qualitative semi-structured interviews with adult and adolescent patients with AA and caregivers were conducted to describe the impact of AA on their HRQoL and wellbeing.
2. A detailed literature review was completed to describe the impact of AA on HRQoL in patients. An additional search was conducted to explore the HRQoL impacts on caregivers.
3. Retrospective analysis of data from the AAPPO, the SF-36 and the HADS from the ALLEGRO 2b/3 trial.

In part 2, a second round of interviews was conducted with adult and adolescent patients, caregivers, as well as healthcare professionals, to obtain feedback on the draft vignettes and to determine the accuracy of the descriptions.

In part 3, a cross-sectional study in which the vignettes were reviewed and rated by the UK general public using the TTO valuation technique was conducted. The TTO method explores the willingness of participants to trade years of life for changes in HRQoL, indicating the utility of each health state. Utilities were estimated for each of the patient and caregiver vignettes. There was a pause in the TTO interviews after the first 20 interviews to review the methods and ensure the exercise was being completed correctly; after which the remaining 100 TTO interviews were completed. More information on the vignette methodology and subsequent derived utilities can be found in Appendix H.

B.3.4.4.4 Patient and caregiver HRQoL

In addition to what has already been discussed in Section B.1.3.2, feedback from PAG representatives also reinforce the burden on carers and partners.[22] Participants noted that alopecia can negatively impact upon the relationships people have with their friends, families or partners. Although partners were generally described as supportive, it was reported that patients found it difficult to accept that their partner could still view them in the same way as they had prior to the onset of alopecia. Similarly, challenges associated with dating were mentioned, including struggles patients had with deciding if and when to inform dating partners of their condition and past negative experiences upon doing so. All PAG representatives who had experience with caregivers of adolescents with alopecia reported impacts upon the caregivers themselves. The most frequently reported impact was a feeling of helplessness, whereby caregivers wanted to help but they did not know how to do so.

[redacted]

Caregiver anxiety about the impact alopecia may have upon the future life of their child was reported by PAG representatives. Participants noted that caregivers often sought out treatments on behalf of their child. These caregivers were described as seeking

answers to questions regarding what treatments are available and from where. Other impacts reported by fewer PAG representatives included frustration and upset or distress. Parents may face challenges in getting their child to attend school. Finally, uncertainty and a strain on family dynamics were also reported.

Given the impact on caregiver whether defined as parent, carer or spouse it is appropriate to consider as part of the economic evaluation. The TTO utility weights of the full sample of people interviewed in the vignette study used in the model are presented in Table 46 and Figure 45.[23] These values are used in the cost effective model base case.

To calculate the caregiver disutility for patients with a SALT score >50, the utility for the UK population norm for people aged 35-44 (0.91),[175] aligned with the mean age of the sample in the vignette study, was subtracted from the carer utility.

Table 46: Summary of health state utility values and the caregiver utility value for the cost-effectiveness analysis

State	Utility value: mean (standard deviation)	Standard error	Range	95% confidence interval
N=120				
SALT 0-10	[redacted]	[redacted]	[redacted]	[redacted]
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT 21-49	[redacted]	[redacted]	[redacted]	[redacted]
SALT 50-100	[redacted]	[redacted]	[redacted]	[redacted]
N=57				
Carer utility for patients with a SALT score >50	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutility for patients with a SALT score >50	[redacted]	[redacted]	[redacted]	[redacted]

The caregiver health state has a total of n = 57 completed data points. All other health states all have n = 120 data points.
Caregiver disutility for patients with a SALT score >50 calculated by subtracting the utility for the UK population norm for people aged 35-44 (0.91)[175] from the carer utility for patients with a SALT score >50.
Abbreviations: SALT, Severity of Alopecia Tool

Figure 45: Full sample utility weights plot

Note: coloured points indicate mean value for each health state, bars indicate upper and lower 95% confidence limits, the caregiver health state has a total of n = 57 completed data points. All other health states all have n = 120 data points.

Although utilities were obtained for a SALT score of >50 and eyebrow or eyelash loss, a conservative assumption was made that all patients in the SALT >50 health state had a utility equal to the SALT 50-100 health state in the vignette study.

A sensitivity analysis was conducted to explore the impact of participant comprehension/understanding of the vignettes and TTO task on the health state valuations. The following exclusion criteria were applied:

- Any participant who values the mildest patient state as worse than most severe states
- Any participant who trades life and subsequently values all health states the same (apart from all states at 1 or -1)
- Any participant who TTO interviewers identified in interview notes as experiencing comprehension difficulties or being disengaged with the TTO task

One participant was excluded based on exclusion criterion one, and four participants were excluded based on exclusion criterion two. Criterion three identified three participants eligible for exclusion, with two already excluded based on criterion two,

resulting in a total of six participants excluded in this sensitivity analysis (n=114). Other general observations noted by interviewers included the influence of a participant's current level of hair loss on valuations. Individuals who had experienced hair loss often made reference to their own experience when valuing health states. It was unclear whether these individuals were less or more willing to trade life years to avoid the health states.

Utility weights of the sensitivity analysis can be found in Table 47. The overall pattern of the results remained unchanged, with the sensitivity analysis showing a slightly larger range from the mildest to most severe health state

Table 47: Summary of health state utility values and the caregiver utility value for the cost-effectiveness analysis (sensitivity analysis)

State	Utility value: mean (standard deviation)	Standard error	Range	95% confidence interval
N=114				
SALT 0-10	[redacted]	[redacted]	[redacted]	[redacted]
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT 21-49	[redacted]	[redacted]	[redacted]	[redacted]
SALT 50-100	[redacted]	[redacted]	[redacted]	[redacted]
N=57				
Carer utility for patients with a SALT score >50	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutility for patients with a SALT score >50	[redacted]	[redacted]	[redacted]	[redacted]

6 participants excluded (valuing all the states the same [N=5]; valuing mildest patient state as worse than most severe states [N=1]).
Caregiver disutility for patients with a SALT score >50 calculated by subtracting the utility for the UK population norm for people aged 35-44 (0.91)[175] from the carer utility for patients with a SALT score >50.
Abbreviations: SALT, Severity of Alopecia Tool

B.3.4.5. Adverse reactions

The adverse reactions to treatment as reported in ALLEGRO 2b/3 (Section 0) were included in the model to characterise the effect on HRQoL for patients experiencing the events.

The magnitude and duration of the disutility associated with the adverse events seen in ALLEGRO 2b/3 were collected from published literature and used to calculate the total QALY decrement. These were then applied to the proportion of patients

experiencing each event, according to their health state, as reported in Section B.3.3.5. Disutilities for adverse events and their duration were collected from published literature because trial data was not used to estimate utilities (as explained in Section B.3.4.4) and the vignette study used does not consider adverse events.

The TEAE disutility values and duration applied to patients receiving ritlecitinib and BSC are shown in Table 48. These inputs were obtained from published literature; it was conservatively assumed that the duration of each disutility would be 28 days. The total QALY decrement is a product of TEAE disutility and disutility duration.

Table 48: Disutility due to TEAEs

Adverse event	Disutility	Disutility duration (days)	Total QALY decrement	Source
Acne	0.07	28.00	0.02367	Assumed same as skin disorder - Stein *et al.* (2018)[188]
Diarrhoea	0.04	28.00	0.01320	Sullivan *et al.* (2004)[189]
Folliculitis	0.07	28.00	0.02367	Assumed same as skin disorder - Stein *et al.* (2018)[188]
Headache	0.03	28.00	0.00887	Sullivan *et al.* (2004)[189]
Nasopharyngitis	0.01	28.00	0.00343	Assumed same as ear and sense organ disorder - Sullivan *et al.* (2004)[189]
Rash	0.03	28.00	0.01083	TA403 NICE (2016)[190]
Upper respiratory tract infection	0.07	28.00	0.02333	Worbes-Cerezo *et al.* (2019)[191]
Urticaria	0.07	28.00	0.02367	Assumed same as skin disorder - Stein *et al.* (2018)[188]

Abbreviations: QALY, quality-adjusted life year; TEAE, treatment emergent adverse event

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

B.3.5.1. Intervention and comparators' costs and resource use

B.3.5.1.1 Ritlecitinib

The drug acquisition cost for ritlecitinib 50 mg per pack of 30 capsules is £[redacted] with the patient access scheme (PAS) discount applied. As lack of compliance is likely to lead to a delay to renew prescriptions of ritlecitinib, as verified by expert opinion,

wastage was modelled by applying compliance as observed in ALLEGRO 2b/3. Treatment compliance is assumed to be [redacted], equal to the compliance observed in the ALLEGRO 2b/3 trial.[147] Ritlecitinib will be used as a monotherapy so no costs for other medicated therapies are included. Administration costs are not included as oral drugs are assumed to have no administration costs.

Monitoring resource use for patients treated with ritlecitinib and BSC were based on findings from the HCRU Delphi panel (Table 49).[24] The ritlecitinib SmPC recommends that [redacted] (see Appendix C). [redacted] These costs for these procedures are assumed to be covered by the full blood count, lipid panel and liver panel that are performed in routine monitoring of patients and so have not been individually included in the treatment costs of ritlecitinib.

Table 49: Monitoring resource utilisation per 12 weeks

Event	Ritlecitinib	BSC
Full blood count	[redacted]	[redacted]
Lipid panel	[redacted]	[redacted]
Liver panel	[redacted]	[redacted]
Trichoscopy	[redacted]	[redacted]

Abbreviations: BSC, best supportive care

Monitoring costs for patients treated with ritlecitinib, including a full blood count, lipid panel, liver panel and trichoscopy, were sourced from NHS reference costs 2020/21 and are presented in Table 50.[192] No monitoring costs are incurred for patients treated with BSC as the cost of a trichoscopy is assumed to be zero since it would be performed during a routine appointment.

Table 50: Unit costs of routine monitoring interventions

Event	Cost, £	Reference
Full blood count	3.63	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. DAPS05 Haematology[192]
Lipid panel	1.85	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. DAPS04 Clinical biochemistry[192]

Liver panel	1.85	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. DAPS04 Clinical biochemistry[192]
Trichoscopy	-	Assumed zero as undertaken during a routine appointment

Abbreviations: NHS, National Health Service

B.3.5.1.2 BSC

BSC is defined as non-pharmacologic treatments, as described in Section B.1.3.3.4, and therefore there are no pharmacological treatment costs for BSC. As discussed in Section B.1.3.3.4, it is difficult to define exactly what this includes as there is a variance between specialist centres;[17] it is assumed it includes wigs and other prosthetic supports alongside routine appointments with healthcare professionals.

B.3.5.2. Health-state unit costs and resource use

Resource use every 12 weeks by health state was estimated using the HCRU Delphi panel (Appendix I).[24]

For the purposes of the economic model, HCRU categories were included in the economic model if the majority of respondents to the Delphi panel (≥5 respondents) agreed or were neutral to the frequency of resource use suggested in the second-round interview.

It was agreed by 7/8 respondents that patients receiving a JAK inhibitor would attend more NHS appointments/consultations than those receiving no pharmacological therapy. However, this is likely to relate to the additional standard monitoring for JAK inhibitors that is outlined in the SmPC. Therefore, to avoid double-counting of appointments with healthcare professionals, it is assumed that the tests required for ritlecitinib, according to the SmPC, are performed during these appointments.

Because, for ritlecitinib, there was a lack of consensus on the frequency of dermatology-related nurse visits in the primary care setting, the frequency of these visits was assumed to be equal to the frequency for patients treated with BSC given that resource use in patients treated with JAK inhibitors was assumed to be greater than in BSC. Similarly, it was assumed that psychological support consultations occurred at the same frequency in BSC and ritlecitinib treated patients despite

respondents indicating that resource use for psychological support would be lower with ritlecitinib treated individuals.

As the NHS funds up to 2 wigs per patient with AA per year, it was assumed that the maximum number of wigs that are provided was 2 per year. From the HCRU Delphi panel, it was found that synthetic wigs last for 3-11 months, whereas human wigs last from 6 months to 2 years. However, it was found that no human wigs are offered to patients with AA on the NHS. Moreover, as respondents reported that wigs are not provided to patients in the SALT ≤10 health state, it was assumed that no patients in this state would have a wig service despite respondents indicating that patients treated with ritlecitinib would have a wig service.

It was agreed by 6/8 respondents that resource use in adults and adolescents would be the same, so it is assumed that resource use for adults and adolescents is equal.

The resource use can vary for patients with a SALT score of 100 compared to patients with a SALT score of 50-99 as patient requirements can vary when patients have complete hair loss. Functionality therefore exists in the model for the user to select where either; health-state resource use for patients with a SALT score of ≥50 is equal to the health-state resource use for patients with a SALT score of 50-99, or patients with a SALT score of 100. In the base case, resource use in this patient group is assumed to be equal to patients with a SALT score of 50-99.

The resource use per cycle applied in the economic model is summarised below in Table 51 and Table 52.

Table 51: Ritlecitinib treatment resource utilisation

Event	SALT 100	SALT 50-99	SALT 21-49	SALT 11-20	SALT ≤10
Wigs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wig service for fitting/collection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Psychological support consultation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology nurse (outpatient setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related nurse visit (primary care setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related GP visits	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatologist outpatient visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; SALT, Severity of Alopecia Tool

Table 52: BSC treatment resource utilisation

Event	SALT 100	SALT 50-99	SALT 21-49	SALT 11-20	SALT ≤10
Wigs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wig service for fitting/collection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Psychological support consultation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology nurse (outpatient setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related nurse visit (primary care setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related GP visits	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatologist outpatient visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; SALT, Severity of Alopecia Tool

The unit costs of the associated events are presented in Table 53. As there were at least 5 clinicians who agreed with the duration of HCRU appointments for BSC, the duration was multiplied with the frequency of the appointments to capture the total resource use per 12 weeks. To avoid double-counting, it was assumed that the duration of appointments was the same for patients treated with BSC and ritlecitinib, given that ritlecitinib was associated with increased frequency of appointments with healthcare professionals.

Table 53: AA management unit costs

Event	Cost, £	Duration (minutes)	Net cost, £	Reference
Wigs	[redacted]	N/A	[redacted]	Expert opinion – average NHS spends on a synthetic wig[24]
Wig service for fitting/collection	20.50	[redacted]	[redacted]	PSSRU unit costs report 2021 - Estimated cost of a wig service with a consultant dermatologist (assumed to be a medical hospital based consultant), 10 minutes[193]
Psychological support consultation	221.52	[redacted]	[redacted]	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. Service code 656 - Clinical Psychology appointment, 45minutes.[192]
Dermatology nurse (outpatient setting)	17.00	[redacted]	[redacted]	PSSRU unit costs report 2021 - Hospital-based nurse Band 6, 20 minute appointment (duration based on expert opinion)[24,193]
Dermatology-related nurse visit (primary care setting)	11.00	[redacted]	[redacted]	PSSRU unit costs report 2021 - Nurse (GP practice), 15 minute appointment (duration based on expert opinion)[24,193]
Dermatology-related GP visits	39.23	[redacted]	[redacted]	PSSRU unit costs report 2021 - GP unit cost, 9.22 minute appointment (duration based on expert opinion)[24,193]

Dermatologist outpatient visit	171.93	[redacted]	[redacted]	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. Service code 330 - Dermatology (consultant led)[192]

Abbreviations: AA, alopecia areata; GP, general practitioner; NHS, National Health Service; PSSRU, Personal Social Services Research Unit

B.3.5.3. Adverse reaction unit costs and resource use

The costs of each adverse event in the model (Section B.3.3.5) were informed by the National Schedule of NHS Costs for 2020/21.[192]

Table 54: Adverse event costs

Adverse event	Cost, £
Acne	627.40
Diarrhoea	627.40
Folliculitis	230.22
Headache	627.40
Nasopharyngitis	230.22
Rash	627.40
Upper respiratory tract infection	231.65
Urticaria	627.40

B.3.5.4. Miscellaneous unit costs and resource use

B.3.5.4.1 Societal costs

Societal costs are included as a scenario analysis given the broader impact of AA beyond the NHS and PSS perspective. Societal costs are incorporated into the model as productivity losses through absenteeism and presenteeism and as OOP expenses.

Productivity costs

The percentage of patients with AA employed is 74.7%, assumed equal to the percentage of adult patients from the ALLEGRO 2b/3 clinical trial who were employed at baseline in the FAS.[147] The percentage of patients in part time work, hourly wage and average work hours per week were assumed equal to the general population in the UK. Therefore, the percentage of patients with AA in employment who work part time in the UK is 24.9%.[194] The average full-time and part-time hourly wages in the UK are £18.76 and £10.65, respectively.[195] The average full-time and part-time hours per week in the UK are 36.6 hours and 16.8 hours, respectively.[194]

The percent reduction in both full- and part-time work due to absenteeism and presenteeism is given in Table 55, based on analysis of responses from the ALLEGRO 2b/3 trial.[147] It was conservatively assumed that the reductions to productivity were the same in full time and part time employment.

Table 55: Percent reduction in full-time and part-time paid work

Reason for reduced productivity	SALT ≥50	SALT 21-49	SALT 11-20	SALT ≤10
Absenteeism (%)	[redacted]	[redacted]	[redacted]	[redacted]
Presenteeism (%)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SALT, severity of alopecia tool

These parameters were used to calculate the societal costs associated with patients with AA. Since the impacts are mostly observed on productivity, we have used the human capital approach. This was done by calculating the work hours lost per week per person by multiplying the average work hours per week by the percentage reduction in paid work due to absenteeism or presenteeism. The cost of lost work per week per person was then calculated by multiplying the average hourly wage by the work hours lost per week. This was then weighted by the proportion of patients in employment. The cost per week was multiplied by the percentage of patients in the model who are adults. This was then transformed to the cost per cycle and applied in the model.

Out-of-pocket costs

The indirect OOP costs for each SALT health state are calculated using the OOP cost and resource utilisation over a 12-week period.

The resource utilisation and unit costs were estimated using the HCRU Delphi panel, presented in Table 56 and Table 57.[24] It was conservatively assumed that patients purchase wigs privately at an equal rate to how often they can be obtained through the NHS.

Table 56: Use of out-of-pocket resources per 12 weeks

Purchase	SALT ≥50	SALT 21-49	SALT 11-20	SALT ≤10
Wigs	[redacted]	[redacted]	[redacted]	[redacted]
Semi-permanent tattooing	[redacted]	[redacted]	[redacted]	[redacted]

Eyebrow microblading	[redacted]	[redacted]	[redacted]	[redacted]
Scalp microblading	[redacted]	[redacted]	[redacted]	[redacted]
Hair fibres	[redacted]	[redacted]	[redacted]	[redacted]
Powders	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: OOP, out-of-pocket; SALT, Severity of Alopecia Tool

Table 57: Out-of-pocket costs

Out-of-pocket cost	Cost, £	Reference
Wigs	[redacted]	Expert opinion[24]
Semi-permanent tattooing	[redacted]	Expert opinion[24]
Eyebrow microblading	[redacted]	Expert opinion[24]
Scalp microblading	[redacted]	Expert opinion[24]
Hair fibres	[redacted]	Expert opinion[24]
Powders	[redacted]	Expert opinion[24]

B.3.6. Severity

Ritlecitinib does not meet the criteria for severity weighting.

B.3.7. Uncertainty

As already discussed in Sections B.1.3.1, B.1.3.2 and B.3.4, AA is a heterogenous condition characterised by a relapsing remitting disease course. There is a disconnect between what is demonstrated in ALLEGRO 2b/3, the HRQoL literature, and what PAGs and clinicians describe as the HRQoL burden to AA patients. The paucity of data means the understanding of how AA impacts patients with AA remains a process in evolution.

The nature of the condition itself may impact on the ability to generate high quality evidence. However, it is clear that a significant burden is associated with AA and the burden is linked to hair loss.

Ritlecitinib has proven efficacy with an acceptable safety profile (Section B.2). However, given observations of the literature, clinical trial data, and the patient and clinician feedback, generic measures of HRQoL such as EQ-5D are not appropriate

to capture the full HRQoL burden of AA. The appropriate next step was to use a vignette methodology to generate health state utilities.

B.3.8. Summary of base-case analysis inputs and assumptions

B.3.8.1. Summary of base-case analysis inputs

A summary of the base case analysis inputs can be found in Appendix J.

B.3.8.2. Assumptions

Table 58: Assumptions underpinning the cost-effectiveness model

Variable	Assumed value	Justification
Time horizon and cycle length		
Time horizon	Lifetime horizon	Aligned with NICE reference case, to capture all differences in costs and outcomes.[164]
Cycle length	12 weeks	12-week cycles to align with the schedule of clinical data collection.
Model structure		
Half cycle correction applied	Included in the base case	NICE reference case;[164] a half-cycle correction was applied to both costs and health outcomes in the semi-Markov model to align with conventional modelling standards.
Health states	On treatment SALT ≥50 On treatment SALT 21-49 On treatment SALT 11-20 On treatment SALT ≤10 BSC SALT ≥50 BSC SALT 21-49 BSC SALT 11-20 BSC SALT ≤10 Death	The population considered for the cost-effectiveness analysis are adolescents (aged ≥12 to <18 years) and adults (aged ≥18 years) with AA and a baseline SALT score of 50 or higher. SALT scores are a tool to measure the extent of disease and response to treatment; this measure was used in the ALLEGRO 2b/3 study and is used in clinical practice in the UK (all dermatologists with a specialist interest in hair disorders consulted use SALT score to define severity).[21] BSC is defined as non-pharmacologic treatments. The BSC comparator includes no drug treatment costs and efficacy is set to placebo results from the ALLEGRO 2b/3 trial. The health state SALT categories have been validated by clinical experts.
Model approach	Semi-Markov Model	Treatment effectiveness is captured by distinct SALT categories which map to resource use, costs, and patients' quality of life. Therefore, a semi-Markov structure is appropriate to capture response to treatment. Prior to the final discontinuation rule, patients are partitioned in SALT-based

Variable	Assumed value	Justification
		health states based on the ALLEGRO 2b/3 clinical trial, hence the "semi" portion of the model description.
Intervention and comparator		
Ritlecitinib study arms for inclusion	50mg ritlecitinib	Several doses of ritlecitinib have been studied in the clinical trials, however, the dose proposed for registration for the treatment of AA is the 50 mg once daily dose.
Comparator	BSC	BSC is the only relevant comparator for the patient population in the CEM, as outlined in Section B.1.3.3.1.
Clinical effectiveness		
In trial transition probabilities	Informed from the ALLEGRO 2b/3 study and ALLEGRO-LT study	This forms the most robust evidence base for ritlecitinib.
Transition probabilities between Week 48 and 96	Informed from the ALLEGRO-LT study	This forms the most robust evidence base for ritlecitinib.
Transition probabilities after Week 96	Patients remain in state unless they discontinue treatment.	Given that response was demonstrated to plateau in the ALLEGRO-LT trial (Section B.2.6.2), it is reasonable to assumed that patients remain in state.
Stopping rule	Interim + final stopping rule	A two-phase stopping rule is proposed because hair growth is not immediate; patients continue to reach higher thresholds of response beyond Week 24 and patients who show no clinically meaningful response at Week 24 can go on to respond by Week 48, as further described in Section B.3.2.3.2.
Discontinuation due to lack of response	If patients SALT score increases above 20 after Week 48, they will discontinue treatment	This would demonstrate a loss of response, and therefore patients would not continue with treatment.
Discontinuation for reasons other than lack of response after Week 48	Parametric curves are fitted to patients enrolled in the ALLEGRO-LT study after Week 48 who had a SALT score ≤20 at Week 48 and whose SALT score does not rise above 20.	This ensures patients who many discontinue ritlecitinib treatment for other reasons are captured, without double-counting discontinuation amongst those who would have discontinued due to loss of response.
Return to baseline after discontinuation for ritlecitinib	Following discontinuation of ritlecitinib to BSC at any time point in the model, it is assumed that patients gradually lose any prior improvement in SALT score and return to a SALT score >50.	This assumption was validated by dermatologists with a specialist interest in hair disorders, stating that most patients will progressively regress after discontinuing treatment.[17] See Section B.3.2.3.3 for further detail.
Spontaneous remission	The percentage of patients on BSC in spontaneous remission is assumed to be the same as the percentage of patients on placebo who had a SALT	As it is not known whether spontaneous remission is durable, it was assumed some patients lose spontaneous remission over time and an equal number of patients gain spontaneous remission over time

Variable	Assumed value	Justification
	score ≤10 at Week 24 in the ALLEGRO 2b/3 trial	
HRQoL		
Patient utility	[redacted]	Valued using the vignette study due to the limitations described with the available HRQoL data and insensitivity of generic measures in patients with AA,[23] as discussed in Section B.3.4.
Caregiver disutility	Included in the base case	The burden to patients described imposes significant burden to carers, as described in Section B.1.3.2.4.
Disutility	Various, sourced from published literature	Limited evidence available for the impact of disutilities specifically in patients with AA.
Adverse events	The model includes adverse events if they occurred during the clinical trial data collection period but will not include events that have increased rates overtime.	Best available evidence.
Costs and resource use		
Wastage	Wastage is modelling by consideration of compliance as observed in ALLEGRO 2b/3.	If patients are non-compliant with treatment, treatment outcomes will be impacted. Also, they will need to renew their prescription less regularly. Therefore, to accurately assess the cost-effectiveness of treatment, compliance is included.
Health state costs	Health state costs were calculated using a micro-costing approach	Allows for the identification of all relevant costs which are parametrised with up-to-date unit costs.

Abbreviations: BSC, best supportive care; NICE, National Institute for Health and Care Excellence; SALT, Severity of Alopecia Tool

B.3.9. Base-case results

B.3.9.1. Base-case incremental cost-effectiveness analysis results

Aggregate base case results for the cost-effectiveness of ritlecitinib compared with BSC are presented in Table 59. Over the lifetime horizon, treatment with ritlecitinib was associated with [redacted] QALYs at a total cost of [redacted]. The ICER of ritlecitinib compared with BSC is £10,877.75 per QALY.

Table 59: Base case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
BSC	[redacted]	[redacted]	[redacted]				
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	0.000	[redacted]	10,877.75

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life year

B.3.10. Exploring uncertainty

Deterministic and probabilistic sensitivity analyses are presented below to explore the level of uncertainty in the model results.

B.3.10.1. Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was used to assess the effect of parameter uncertainty on the ICER. The PSA works by drawing a value for each parameter from their assumed probability distributions 10,000 times and evaluating the ICER obtained with each iteration. Where the standard errors for the parameters are unknown, they are assumed to be 20% of the parameter value for the purposes of defining the distributions for each parameter. Mean incremental results were recorded and illustrated through an incremental cost-effectiveness plane (ICEP). A cost-effectiveness acceptability curve (CEAC) was plotted and is presented in Figure 48.

Aggregate probabilistic results for the cost-effectiveness of ritlecitinib compared with BSC are presented in Table 60. Over the lifetime horizon, treatment with ritlecitinib was associated with [redacted] QALYs at a total cost of [redacted]. The ICER of ritlecitinib compared with BSC is £11,708 per QALY.

Table 60: Probabilistic results

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	11,708

Ritlecitinib was more costly and more effective than BSC in 100% of iterations ran at a willingness-to-pay threshold of [redacted]. The ICEP is illustrated in Figure 46.

Figure 46: Incremental cost-effectiveness plane

Abbreviations: QALY, quality-adjusted life year

The CEAC is displayed in Figure 47 to illustrate the probability of ritlecitinib being cost-effective compared to BSC, at various willingness to pay thresholds. At willingness to pay thresholds above [redacted] ritlecitinib is likely to be more cost-effective than BSC.

Figure 47: Cost-effectiveness acceptability curve

B.3.10.2. Deterministic sensitivity analysis

A one-way sensitivity analysis (OWSA) was used to assess the effect of parameter variation on the ICER and net monetary benefit (NMB). The OWSA was performed using a standard error approach. Where the standard error was not available for a parameter, the standard error was assumed to be 20% of the mean value. Based on its mean and the standard error, the parameter was then varied using a 95% confidence interval based on the distribution of the parameter. Beta distributions were used for treatment compliance.

The results of the model were then evaluated using the upper and lower bounds for each parameter, fixing all other parameters' values and recording the overall NMB value. This measures which variables have the largest impact on the overall cost-effectiveness analysis results and provides justification for estimates of the model's robustness to parameter variation.

Figure 48 displays the tornado diagram for the incremental cost (£) per QALY gained by ritlecitinib compared to BSC. Results are most sensitive to the utility score of the SALT ≥ 50 health state, the treatment cost of the BSC SALT ≥ 50 health state and the SALT ≤10 utility. In all cases, the incremental cost (£) per QALY gained remains lower than £20,000.

Figure 48: Incremental cost (£) per QALY gained tornado diagram

B.3.10.3. Scenario analysis

Thorough sensitivity analysis of the model was performed, with results presented in Table 61. The scenario analysis found that the time horizon, BSC patients reverting to the SALT score ≥50 health state and including carer disutility had the biggest impact on the ICER of ritlecitinib relative to BSC. In all scenarios, the ICER of ritlecitinib relative to BSC is less than £20,000.

Table 61: Scenario analyses of the base case of the model

Model setting tests	Base case assumption	Scenario assumptions	ICER of ritlecitinib relative to BSC (£)
Base case	-	-	10,877.75
Perspective	Payer	Societal	7,871.39
Time horizon	Lifetime	5 years	13,725.14
Model death in first 48 weeks?	No	Yes	10,856.46
Age group	≥12 years	≥18 years	11,387.46
Age group	≥12 years	≥12 to <18 years	10,851.52
Stopping rule criteria	Interim+Final	Final Only	11,090.28
Final SALT score	SALT ≤ 20	SALT ≤ 10	11,317.11

Final stopping rule time point	48 weeks	36 weeks	10,952.96
BSC SALT 11-49 revert to SALT 50-100?	Yes	No	14,563.67
Allow ritlecitinib treatment discontinuers to have spontaneous remission?	Yes	No	11,907.19
Discontinue patients based on SALT score after 48 weeks	SALT ≤ 20	Do not discontinue	12,071.97
Extrapolation of LT data after 24 months	Stay in state	Last observation carried forwards	11,545.46
Extrapolation of LT data after 24 months	Stay in state	Average	12,381.31
Treatment discontinuation rate curve	Exponential	Weibull	10,618.57
Treatment discontinuation rate curve	Exponential	Gompertz	10,608.97
Treatment discontinuation rate curve	Exponential	Log-logistic	10,499.39
Treatment discontinuation rate curve	Exponential	Lognormal	10,369.84
SALT >50 HCRU assumption	SALT 50-99	SALT 100	10,896.55
Utility weight source	TTO Analysis	TTO Analysis (SA)	10,458.76
Include carer disutility	Yes	No	14,731.17
Disutility weight source	TTO Analysis	TTO Analysis (SA)	12,251.19
Spontaneous remission probability	1.54%	10.00%	12,172.33

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LT, long term; SA, sensitivity analysis; SALT, Severity of Alopecia Tool; TTO, time-trade off

B.3.11. Subgroup analysis

Subgroups by age were explored for adults (≥18 years) and adolescents (≥12 years and <18 years) to account for any differences in response to treatment with ritlecitinib. For the adult subgroup, data collected in adults only were used to determine transitions between health states from ALLEGRO 2b/3 and ALLEGRO-LT for patients who were

of the relevant age. For the adolescent group, due to smaller patient numbers, it was assumed that the efficacy in adults would be the same as in adults and adolescents.

Aggregate base case results for the cost-effectiveness of ritlecitinib compared with BSC for adults are presented in Table 62. Over the lifetime horizon, treatment with ritlecitinib was associated with [redacted] QALYs at a total cost of [redacted]. The ICER of ritlecitinib compared with BSC is £11,387 per QALY.

Table 62: Base case results for adults (≥18 years)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
BSC	[redacted]	[redacted]	[redacted]				
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	11,387

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life year

Aggregate base case results for the cost-effectiveness of ritlecitinib compared with BSC for adolescents are presented in Table 63. Over the lifetime horizon, treatment with ritlecitinib was associated with [redacted] QALYs at a total cost of [redacted]. The ICER of ritlecitinib compared with BSC is £10,852 per QALY.

Table 63: Base case results for adolescents (≥12 years and <18 years)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
BSC	[redacted]	[redacted]	[redacted]				
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	10,852

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life year

B.3.12. Benefits not captured in the QALY calculation

Given the use of a vignette, there are unlikely notable aspects missing from the QALY calculation.

As the vignettes were focussed on health states only and did not differentiate between treatment (i.e., ritlecitinib versus BSC), preferences in route of administration or for BSC the impact of camouflage or other coping strategies may not have been captured. Ritlecitinib alleviates the unmet need for a licensed oral treatment with good efficacy and acceptable safety profiles for patients with AA. The oral route of administration for ritlecitinib is preferable for some patients, might potentially reduce unnecessary visits to hospital or specialist setting, along with once daily dosing which potentially aids compliance.

B.3.13. Validation

B.3.13.1. Validation of cost-effectiveness analysis

As already discussed in B.1.1 and B.1.3.3, the approach with the model development was to develop a clear and comprehensive understanding of current clinical practice and the therapeutic landscape for patients with AA in the UK. This was achieved by conducting a therapeutic landscape Delphi panel based on current guidance and the opinions of key clinical experts.[21] In terms of comparators most relevant to ritlecitinib for patients with severe AA, clinicians with a special interest in hair disorders advised that there are a number of off-licence systemic pharmacological therapies for AA (as detailed in Section B.1.3.3). However, they are rarely used for patients with severe AA (and less so with adolescents), due to broad concerns around safety, limited evidence for their efficacy, and not being widely available across the UK. Therefore, the vast majority of patients with severe AA are not receiving any pharmacological treatment. Insights from the clinician validation are that dermatologists outside of specialist centres would not prescribe systemic treatment and patients would go straight to being treated with BSC defined as non-pharmacological therapy.[17]

The model has subsequently undergone thorough internal and external validation. Validation of the model by both internal and external health economists involved review of:

- Formulae
- Consistency with the model decision problem
- VBA implementation

- Inputs
- Model functionality

In addition, the model structure, clinical inputs and assumptions were validated, initially with three UK clinical dermatologists with a specialist interest in hair disorders. Clinicians supported the proposed methodology to parameterise the clinical effectiveness of ritlecitinib and BSC, including (for ritlecitinib) the stopping rule and patients transitioning sequentially through the health states with a greater SALT score until reaching 'BSC SALT ≥50' following treatment discontinuation.

A second round of validation interviews with three UK clinical dermatologists with a specialist interest in hair disorders (including the two clinicians interviewed in the first round) took place to revalidate model structure, the use of clinical data from the ALLEGRO 2b/3 study, inputs and assumptions and ensure they reflected what would be observed in practice.

B.3.14. Interpretation and conclusions of economic evidence

The cost effectiveness analysis shows that for patients with severe AA, including AT/AU, ritlecitinib 50mg taken once daily is a cost-effective use of NHS resources. The results from the base case analysis show that ritlecitinib is associated with [redacted] when compared to ritlecitinib over a lifetime time horizon, and is cost-effective with an ICER compared to BSC of £10,878.

In all of the scenario analyses and sensitivity analyses, ritlecitinib remained cost-effective at the cost per QALY of £30,000 threshold. Subgroup analyses also demonstrated that ritlecitinib is cost effective when considering adults and adolescents separately. The results from the PSA confirm the deterministic results and show that in [redacted].

B.3.14.1. Relevance and generalisability

The economic evaluation is based on the adult and adolescent patient population of ALLEGRO 2b/3 and ALLEGRO-LT trials which evaluated the efficacy and safety of ritlecitinib in patients ≥12 years of age with severe AA, defined as a SALT score of

≥50. Clinical feedback supports the use of SALT to evaluate treatment outcomes and a SALT score of ≤20 represents a meaningful outcome for patients. The characteristics of the population of ALLEGRO 2b/3 and ALLEGRO-LT are considered generalisable to England and Wales, based on best available evidence (Section B.2.3.1.2 and B.2.3.2.2).

Extensive feedback has been sought via a therapeutic Delphi panel (Section B.1.1) consisting of UK clinicians with a special interest in hair disorders to support model development. The modelled treatment pathway and inputs have been designed and selected based on these and follow up discussions to be fully reflective of clinical practice in England and Wales.

B.3.14.2. Strengths of the economic evaluation

The key strengths of the economic analysis are:

- No cost effectiveness studies of interventions in AA were identified to inform the economic analysis presented in this submission (Appendix G). Therefore, a *de novo* economic model was developed to address the decision problem which reflects original and novel research alongside feedback from NICE scientific advice.
- Efficacy and safety was based on the ALLEGRO 2b/3 trial a large randomised controlled trial and supported by ALLEGRO-LT, a long term study of AA patients.
- Key components of the analysis were validated by clinical experts and the economic analysis also incorporates (where relevant) input from representatives of patient advisory groups.
- The efficacy for both arms was drawn from the same trial limiting heterogeneity in the data while the outcomes evaluated are representative of clinical outcomes deemed meaningful by dermatologists with a special interest in hair disorders.
- Extrapolation of data is supported by the follow up of AA patients in the ALLEGRO-LT trial allowing best fit calculations to represent long term outcomes to be evaluated and reduce decision uncertainty.
- Scenario analyses and an OWSA were performed to test the impact of parameter uncertainty on results. A range of sensitivity analyses have been explored to test structural and parametric uncertainty.

B.3.14.3. Limitations of the economic evaluation

As with all economic analysis, there are some limitations. The main limitations are that:

- There is significant paucity of data in AA both from an economic and clinical perspective. In addition, there have been no previous appraisals of treatments for patients with AA and no treatment analogues to learn from. Therefore, significant engagement with clinicians with a special interest in hair disorders and PAGs was undertaken to understand the condition and its treatment pathway in the UK.
- EQ-5D and SF-36 data were reported in the clinical trial but did not show a significant difference between treatment groups. AAPPO was also not appropriate as it is not a preference-based measure. Therefore, utilities were obtained from a vignette study and may be considered less generalisable across other diseases and treatments. However, according to the hierarchy of preferred HRQoL methods in the NICE guidance,[187] this is an appropriate solution.
- There is limited evidence to parametrise best supportive care, with the ALLEGRO 2b/3 study placebo-controlled only until 24 weeks. To address this, scenario analysis has been considered to understand the impact of different underlying rates of spontaneous remission.
- HCRU resource data was not available in the literature, therefore a study was conducted to ascertain HCRU costs and activity for AA. In the absence of published evidence, this is the best quality evidence that could be considered.

B.3.14.4. Conclusions from the economic evidence

There is substantial unmet need in severe AA both for adults and adolescents. Until recently there have been no licensed or approved treatments and there is significant paucity of clinical and resource impact data. A *de novo* economic model was developed in to assess the cost-effectiveness of ritlecitinib relative to BSC for the treatment of AA. The model uses data from the relevant RCT, ALLEGRO 2b/3, and long-term extension study, ALLEGRO-LT, to inform the model as well as published sources and clinical expert elicitation. The economic evidence also includes extensive insight from PAGs to support understanding of HRQoL for patients with AA. Uncertainty in the model was explored through extensive deterministic, probabilistic

and scenario analysis. The economic evidence shows that in adults and adolescent patients with severe AA including AT/AU, ritlecitinib 50mg given once daily represents a cost-effective and valuable treatment for the UK NHS.

B.4 References

1. 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.
2. Guo H, Cheng Y, Shapiro J, *et al.* The role of lymphocytes in the development and treatment of alopecia areata. *Expert Rev Clin Immunol* 2015. 11: 1335–51.
3. Fricke VAC & Miteva M. Epidemiology and burden of alopecia areata: a systematic review. *Clin Cosmet Investig Dermatol* 2015. 8: 397–403.
4. Harries M, Macbeth A, Holmes S, *et al.* The epidemiology of alopecia areata: a population-based cohort study in UK primary care. *British Journal of Dermatology* 2021.
5. Benigno M, Anastassopoulos KP, Mostaghimi A, *et al.* A large cross-sectional survey study of the prevalence of alopecia areata in the United States. *Clin Cosmet Investig Dermatol* 2020. 13: 259–266.
6. Mesinkovska N, King B, Mirmirani P, *et al.* Burden of Illness in Alopecia Areata: A Cross-Sectional Online Survey Study. *Journal of Investigative Dermatology Symposium Proceedings* 2020. 20: S62–S68.
7. Messenger AG, McKillop J, Farrant P, *et al.* British Association of Dermatologists' guidelines for the management of alopecia areata 2012. *Br J Dermatol* 2012. 166: 916–26.
8. Harries MJ, Sun J, Paus R, *et al.* Management of alopecia areata. *BMJ* 2010. 341: c3671.

9. Strazzulla LC, Wang EHC, Avila L, *et al.* Alopecia areata: Disease characteristics, clinical evaluation, and new perspectives on pathogenesis. *J Am Acad Dermatol* 2018. 78: 1–12.

10. Xu L, Liu KX & Senna MM. A Practical Approach to the Diagnosis and Management of Hair Loss in Children and Adolescents. *Front Med (Lausanne)* 2017. 4: 112.

11. 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. *Journal of Patient-Reported Outcomes* 2020. 4: 76.

12. Rodgers A. Why Finding a Treatment for Alopecia Areata Is Important: A Multifaceted Perspective. *J Investig Dermatol Symp Proc* 2018. 19: S51–S53.

13. Layegh P, Arshadi HR, Shahriari S, *et al.* A comparative study on the prevalence of depression and suicidal ideation in dermatology patients suffering from psoriasis, acne, alopecia areata and vitiligo. *Iranian Journal Of Dermatology* 2010. 13:

14. Liu LY, King BA & Craiglow BG. Alopecia areata is associated with impaired health-related quality of life: a survey of affected adults and children and their families. *Journal of the American Academy of Dermatology* 2018. 79: 556-558. e1.

15. Pfizer data on file. Primary market research to inform GPA development for Ritlecitinib in AA. 2022.

16. Macbeth AE, Holmes S, Harries M, *et al.* PLACEHOLDER (DRAFT MANUSCRIPT): The associated burden of mental health conditions in alopecia areata: A UK population-based cohort study using primary care data. 2021.

17. Pfizer data on file. UK interviews with dermatologists with a specialist interest in hair disorders for the use of ritlecitinib for patients with AA. 2022.

18. Tosti A, Bellavista S & Iorizzo M. Alopecia areata: A long term follow-up study of 191 patients. *Journal of the American Academy of Dermatology* 2006. 55: 438–441.

19. Açıkgöz G, Özmen İ, Cayırlı M, *et al.* Pulse methylprednisolone therapy for the treatment of extensive alopecia areata. *Journal of dermatological treatment* 2014. 25: 164–166.

20. National Institute for Health and Care Excellence. Clinical Knowledge Summaries: alopecia areata. 2018. at <https://cks.nice.org.uk/topics/alopecia-areata/>

21. Pfizer data on file. Elicitation of expert opinion to aid understanding of current therapeutic landscape in the UK for people with alopecia areata. Final report. 2022.

22. Pfizer data on file. Qualitative research in alopecia. Targeted literature review and patient advocacy group representative interviews in the United Kingdom. 2022.

23. Pfizer data on file. Vignette study for utility estimation in Alopecia Areata. 2022.

24. Pfizer data on file. Understanding health care resource utilisation of alopecia areata in the UK. 2022.

25. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) advice letter (EMEA/H/SA/3875/1/2018/HTA/III). 2018.

26. Wyrwich KW, Kitchen H, Knight S, *et al.* The alopecia areata Investigator's Global Assessment (AA-IGA^{TM}) scale: a measure for evaluating clinically meaningful success in clinical trials. *Br J Dermatol* 2020. 183: 702–709.

27. Wambier CG & King B. *Rethinking the classification of alopecia areata*. (2019).

28. Pratt CH, King LE Jr, Messenger AG, *et al.* Alopecia areata. *Nat Rev Dis Primers* 2017. 3: 17011.

29. Alopecia Areata. *NORD (National Organization for Rare Disorders)* at <https://rarediseases.org/rare-diseases/alopecia-areata/>

30. Thomis DC & Berg LJ. Peripheral expression of Jak3 is required to maintain T lymphocyte function. *J Exp Med* 1997. 185: 197–206.

31. Telliez J-B, Dowty ME, Wang L, *et al.* Discovery of a JAK3-Selective Inhibitor: Functional Differentiation of JAK3-Selective Inhibition over pan-JAK or JAK1-Selective Inhibition. *ACS Chemical Biology* 2016. 11: 3442–3451.

32. 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.

33. Bao Y, Zheng J, Han C, *et al.* Tyrosine kinase Btk is required for NK cell activation. *J Biol Chem* 2012. 287: 23769–78.

34. Perera E, Yip L & Sinclair R. Alopecia areata. *Hair Ther Transplant* 2014. 4: 118.

35. Bertolini M, Zilio F, Rossi A, *et al.* Abnormal interactions between perifollicular mast cells and CD8+ T-cells may contribute to the pathogenesis of alopecia areata. *PLoS One* 2014. 9: e94260.

36. Atherly LO, Brehm MA, Welsh RM, *et al.* Tec kinases Itk and Rlk are required for CD8+ T cell responses to virus infection independent of their role in CD4+ T cell help. *J Immunol* 2006. 176: 1571–81.

37. 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. 85: 379–387.

38. Gilhar A, Etzioni A & Paus R. Alopecia areata. *N Engl J Med* 2012. 366: 1515–25.

39. Xing L, Dai Z, Jabbari A, *et al.* Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. *Nat Med* 2014. 20: 1043–9.

40. Ito T, Ito N, Saatoff M, *et al.* Maintenance of hair follicle immune privilege is linked to prevention of NK cell attack. *J Invest Dermatol* 2008. 128: 1196–206.

41. Simakou T, Butcher JP, Reid S, *et al.* Alopecia areata: A multifactorial autoimmune condition. *Journal of autoimmunity* 2019. 98: 74–85.

42. Suarez-Farinas M, Ungar B, Noda S, *et al.* Alopecia areata profiling shows TH1, TH2, and IL-23 cytokine activation without parallel TH17/TH22 skewing. *J Allergy Clin Immunol* 2015. 136: 1277–87.

43. Guttman-Yassky E, Ungar B, Noda S, *et al.* Extensive alopecia areata is reversed by IL-12/IL-23p40 cytokine antagonism. *J Allergy Clin Immunol* 2016. 137: 301–304.

44. Xu H, Jesson MI, Seneviratne UI, *et al.* PF-06651600, a dual JAK3/TEC family kinase inhibitor. *ACS chemical biology* 2019. 14: 1235–1242.

45. Dai Z, Chen J, Chang Y, *et al.* Selective inhibition of JAK3 signaling is sufficient to reverse alopecia areata. *JCI insight* 2021. 6:

46. Divito SJ & Kupper TS. Inhibiting Janus kinases to treat alopecia areata. *Nature Medicine* 2014. 20: 989–990.

47. Smith SE, Neier SC, Reed BK, *et al.* Multiplex matrix network analysis of protein complexes in the human TCR signalosome. *Sci Signal* 2016. 9: rs7.

48. Suchonwanit P, Kositkuljorn C & Pomsoong C. Alopecia Areata: An Autoimmune Disease of Multiple Players. *Immunotargets Ther* 2021. 10: 299–312.

49. Fonda-Pascual P, Vano-Galvan S, Garcia-Hernandez MJ, *et al.* Alopecia Areata Sisaipho: Clinical and Therapeutic Approach in 13 Patients in Spain. *Int J Trichology* 2016. 8: 99–100.

50. Cranwell WC, Lai VW, Photiou L, *et al.* Treatment of alopecia areata: An Australian expert consensus statement. *Australas J Dermatol* 2019. 60: 163–170.

51. Islam N, Leung PSC, Huntley AC, *et al.* The autoimmune basis of alopecia areata: A comprehensive review. *Autoimmunity Reviews* 2015. 14: 81–89.

52. Altmeyers encyclopedia. Alopecia areata universalis. 2020. https://www.altmeyers.org/en/dermatology/alopecia-areata-universalis-118559 [Last accessed November 2021].

53. Finner AM. Alopecia areata: Clinical presentation, diagnosis, and unusual cases. *Dermatol Ther* 2011. 24: 348–354.

54. Olsen E, Hordinsky M, McDonald-Hull S, *et al.* Alopecia areata investigational assessment guidelines. *Journal of the American Academy of Dermatology* 1999. 40: 242–246.

55. Olsen EA, Hordinsky MK, Price VH, *et al.* Alopecia areata investigational assessment guidelines–Part II. *Journal of the American Academy of Dermatology* 2004. 51: 440–447.

56. Kiszewski AE, Bevilaqua M & De De Abreu LB. Mesalazine in the treatment of extensive alopecia areata: a new therapeutic option? *International Journal of Trichology* 2018. 10: 99.

57. Grenier P-O & Veillette H. Treatment of alopecia universalis with oral alitretinoin: A case report. *JAAD Case Reports* 2017. 3: 140–142.

58. Ibrahim O, Bayart CB, Hogan S, *et al.* Treatment of alopecia areata with tofacitinib. *JAMA dermatology* 2017. 153: 600–602.

59. Bertolini M, McElwee K, Gilhar A, *et al.* Hair follicle immune privilege and its collapse in alopecia areata. *Experimental dermatology* 2020. 29: 703–725.

60. Schwartzberg PL, Finkelstein LD & Readinger JA. TEC-family kinases: regulators of T-helper-cell differentiation. *Nature Reviews Immunology* 2005. 5: 284–295.

61. Walker SA & Rothman S. Alopecia areata: a statistical study and consideration of endocrine influences. *Journal of Investigative Dermatology* 1950. 14: 403–413.

62. Gip L. Alopecia areata: A follow up investigation of outpatient material. *Acta Derm Venereol (Stockh)* 1969. 49: 180–188.

63. Hordinsky MK. Overview of alopecia areata. *J Investig Dermatol Symp Proc* 2013. 16: S13-5.

64. Tan E, Tay Y-K, Goh C-L, *et al.* The pattern and profile of alopecia areata in Singapore – a study of 219 Asians. *International Journal of Dermatology* 2002. 41: 748–753.

65. Toussi A, Barton VR, Le ST, *et al.* Psychosocial and psychiatric comorbidities and health-related quality of life in alopecia areata: A systematic review. *Journal of the American Academy of Dermatology* 2021. 85: 162–175.

66. Gupta M. Incidence of psychiatric disorders in dermatological patients. *Journal of the European Academy of Dermatology and Venereology* 2003. 17: 624–626.

67. Pfizer data on file. Ritlecitinib OMA Engagement 19th January 2022. 2022.

68. Lee HH, Gwillim E, Patel KR, *et al.* Epidemiology of alopecia areata, ophiasis, totalis, and universalis: A systematic review and meta-analysis. *J Am Acad Dermatol* 2020. 82: 675–682.

69. Pfizer data on file. Summary of SÆfetyWorks Analysis Methods and Results. 2018.

70. Oakley A & Bell H. Alopecia Areata. *DermNet NZ*

71. Winnette R, Martin S, Harris N, *et al.* Development of the Alopecia Areata Patient Priority Outcomes Instrument: A Qualitative Study. *Dermatol Ther (Heidelb)* 2021. 11: 599–613.

72. Macey J, Kitchen H, Aldhouse NV, *et al.* Dermatologist and Patient Perceptions of Treatment Success in Alopecia Areata and Evaluation of Clinical Outcome Assessments in Japan. *Dermatology and therapy* 2021. 11: 433–447.

73. Winnette R, Martin S, Harris N, *et al.* Development of the Alopecia Areata Patient Priority Outcomes Instrument: A Qualitative Study. *Dermatology and Therapy* 2021. 11: 599–613.

74. Wyrwich KW, Kitchen H, Knight S, *et al.* Development of clinician-reported outcome (ClinRO) and patient-reported outcome (PRO) measures for eyebrow, eyelash and nail assessment in alopecia areata. *American journal of clinical dermatology* 2020. 21: 725–732.

75. 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.

76. Chelidze K & Lipner SR. Nail changes in alopecia areata: an update and review. *Int J Dermatol* 2018. 57: 776–783.

77. Han JJ, Li SJ, Joyce CJ, *et al.* Association of resilience and perceived stress in patients with alopecia areata: A cross-sectional study. *Journal of the American Academy of Dermatology* 2021.

78. Yu N-L, Tan H, Song Z-Q, *et al.* Illness perception in patients with androgenetic alopecia and alopecia areata in China. *Journal of psychosomatic research* 2016. 86: 1–6.

79. Finlay AY & Khan G. Dermatology Life Quality Index (DLQI)—a simple practical measure for routine clinical use. *Clinical and experimental dermatology* 1994. 19: 210–216.

80. Janković S, Perić J, Maksimović N, *et al.* Quality of life in patients with alopecia areata: a hospital-based cross-sectional study. *Journal of the European Academy of Dermatology and Venereology* 2016. 30: 840–846.

81. Chren M-M. The Skindex instruments to measure the effects of skin disease on quality of life. *Dermatologic clinics* 2012. 30: 231–236.

82. Holmes S, Harries M, Macbeth AE, *et al.* People with alopecia areata have an increased burden of atopic and autoimmune comorbidity: a large UK population-based matched control study in primary care. in

83. Lee S, Lee H, Lee CH, *et al.* Comorbidities in alopecia areata: a systematic review and meta-analysis. *Journal of the American Academy of Dermatology* 2019. 80: 466-477. e16.

84. Koo JYM, Shellow WVR, Hallman CP, *et al.* ALOPECIA AREATA AND INCREASED PREVALENCE OF PSYCHIATRIC DISORDERS. *International Journal of Dermatology* 1994. 33: 849–850.

85. Okhovat J-P, Marks DH, Manatis-Lornell A, *et al.* Association between alopecia areata, anxiety, and depression: a systematic review and meta-analysis. *Journal of the American Academy of Dermatology* 2019.

86. Chu SY, Chen YJ, Tseng WC, *et al.* Psychiatric comorbidities in patients with alopecia areata in Taiwan: a case–control study. *British Journal of Dermatology* 2012. 166: 525–531.

87. Colon EA, Popkin MK, Callies AL, *et al.* Lifetime prevalence of psychiatric disorders in patients with alopecia areata. *Compr Psychiatry* 1991. 32: 245–51.

88. Ruiz-Doblado S, Carrizosa A & Garcia-Hernandez MJ. Alopecia areata: psychiatric comorbidity and adjustment to illness. *Int J Dermatol* 2003. 42: 434–7.

89. Sellami R, Masmoudi J, Ouali U, *et al.* The relationship between alopecia areata and alexithymia, anxiety and depression: a case-control study. *Indian J Dermatol* 2014. 59: 421.

90. Rajoo Y, Wong J, Cooper G, *et al.* The relationship between physical activity levels and symptoms of depression, anxiety and stress in individuals with alopecia Areata. *BMC Psychol* 2019. 7:

91. Gupta MA & Gupta AK. Depression and suicidal ideation in dermatology patients with acne, alopecia areata, atopic dermatitis and psoriasis. *The British journal of dermatology* 1998. 139: 846–850.

92. Vélez-Muñiz R d C, Peralta-Pedrero ML, Jurado-Santa Cruz F, *et al.* Psychological Profile and Quality of Life of Patients with Alopecia Areata. *Skin Appendage Disorders* 2019. 5: 293–298.

93. Christensen T, Yang JS & Castelo-Soccio L. Bullying and Quality of Life in Pediatric Alopecia Areata. *Skin Appendage Disord* 2017. 3: 115–118.

94. Liakopoulou M, Alifieraki T, Katideniou A, *et al.* Children with alopecia areata: psychiatric symptomatology and life events. *J Am Acad Child Adolesc Psychiatry* 1997. 36: 678–84.

95. Sinclair RD. Alopecia areata and suicide of children. *The Medical Journal of Australia* 2014. 200: 145.

96. Dubois M, Baumstarck-Barrau K, Gaudy-Marqueste C, *et al.* Quality of life in alopecia areata: a study of 60 cases. *J Invest Dermatol* 2010. 130: 2830–3.

97. Hay RJ, Johns NE, Williams HC, *et al.* The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions. *J Invest Dermatol* 2014. 134: 1527–1534.

98. Burge RT, Anderson P, Austin J, *et al.* 26158 The patient-reported burden of alopecia areata by current severity: A real-world study in the US. *Journal of the American Academy of Dermatology* 2021. 85: AB86.

99. Assessment of Quality of Life. AQoL Transformations. at <https://www.aqol.com.au/index.php/transformations>

100. Liu LY, King BA & Craiglow BG. Health-related quality of life (HRQoL) among patients with alopecia areata (AA): a systematic review. *Journal of the American Academy of Dermatology* 2016. 75: 806-812. e3.

101. Rencz F, Gulacsi L, Pentek M, *et al.* Alopecia areata and health-related quality of life: a systematic review and meta-analysis. *Br J Dermatol* 2016. 175: 561–71.

102. Ghajarzadeh M, Ghiasi M & Kheirkhah S. Depression and quality of life in Iranian patients with Alopecia Areata. *Iranian Journal of Dermatology* 2011. 14: 140–143.

103. Chernyshov PV, Tomas-Aragones L, Finlay AY, *et al.* Quality of life measurement in alopecia areata. Position statement of the European Academy of Dermatology and Venereology Task Force on Quality of Life and Patient Oriented Outcomes. *Journal of the European Academy of Dermatology and Venereology* 2021. 35: 1614–1621.

104. Abedini R, Hallaji Z, Lajevardi V, *et al.* Quality of life in mild and severe alopecia areata patients. *International journal of women's dermatology* 2018. 4: 91–94.

105. Al-Mutairi N & Eldin O. Clinical profile and impact on quality of life: Seven years experience with patients of alopecia areata. *Indian Journal of Dermatology, Venereology and Leprology* 2011. 77: 489.

106. Qi S, Xu F, Yang Q, *et al.* Profile of alopecia areata in 655 Chinese patients. in *JOURNAL OF THE AMERICAN ACADEMY OF DERMATOLOGY* (MOSBY-ELSEVIER 360 PARK AVENUE SOUTH, NEW YORK, NY 10010-1710 USA, 2010). 62: AB75–AB75.

107. Abideen F, Valappil AT, Mathew P, *et al.* Quality of life in patients with alopecia areata attending dermatology department in a tertiary care centre-A cross-sectional study. *Journal of Pakistan Association of Dermatologists* 2018. 28: 175–180.

108. Zhang M & Zhang N. Quality of life assessment in patients with alopecia areata and androgenetic alopecia in the People's Republic of China. *Patient Prefer Adherence* 2017. 11: 151–155.

109. Gakhar A, Gupta S, Singla R, *et al.* Impact on Quality of Life of Chronic Paediatric Dermatoses on their Family Members – A Study from Tertiary Care Hospital in North India. *International Journal of Pharmaceutical Quality Assurance* 2021. 270–274.

110. Putterman E, Patel DP, Andrade G, *et al.* Severity of disease and quality of life in parents of children with alopecia areata, totalis, and universalis: a prospective, cross-sectional study. *Journal of the American Academy of Dermatology* 2019. 80: 1389–1394.

111. Senna M, Ko J, Tosti A, *et al.* Alopecia Areata Treatment Patterns, Healthcare Resource Utilization, and Comorbidities in the US Population Using Insurance Claims. *Advances in therapy* 2021. 38: 4646–4658.

112. Li SJ, Mostaghimi A, Tkachenko E, *et al.* Association of out-of-pocket health care costs and financial burden for patients with alopecia areata. *JAMA dermatology* 2019. 155: 493–494.

113. Kim BJ, Uk min S, Park KY, *et al.* Combination therapy of cyclosporine and methylprednisolone on severe alopecia areata. *Journal of dermatological treatment* 2008. 19: 216–220.

114. Spano F & Donovan JC. Alopecia areata. *Canadian Family Physician* 2015. 61: 757.

115. MacDonald Hull SP, Wood ML, Hutchinson PE, *et al.* Guidelines for the management of alopecia areata. *British Journal of Dermatology* 2003. 149: 692–699.

116. Rossi A, Muscianese M, Piraccini BM, *et al.* Italian Guidelines in diagnosis and treatment of alopecia areata. *Giornale italiano di dermatologia e venereologia: organo ufficiale, Societa italiana di dermatologia e sifilografia* 2019. 154: 609–623.

117. Delamere FM, Sladden MM, Dobbins HM, *et al.* Interventions for alopecia areata. *Cochrane Database Syst Rev* 2008. CD004413. doi:10.1002/14651858.CD004413.pub2

118. Trueb RM & Dias MFRG. Alopecia Areata: a Comprehensive Review of Pathogenesis and Management. *Clin Rev Allergy Immunol* 2018. 54: 68–87.

119. 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. *Journal of the American Academy of Dermatology* 2020. 83: 123–130.

120. The All Party Parliamentary Group on Skin. Mental Health and Skin Disease report (2020). at <https://www.appgs.co.uk/mental-health-and-skin-disease-report-2020/>

121. 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.

122. 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.

123. Madani S & Shapiro J. Alopecia areata update. *Journal of the American Academy of Dermatology* 2000. 42: 549–566.

124. Phan K, Ramachandran V & Sebaratnam DF. Methotrexate for alopecia areata: A systematic review and meta-analysis. *Journal of the American Academy of Dermatology* 2019. 80: 120-127.e2.

125. Açıkgöz G, Çalışkan E, Tunca M, *et al.* The effect of oral cyclosporine in the treatment of severe alopecia areata. *Cutaneous and Ocular Toxicology* 2014. 33: 247–252.

126. 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.

127. NHS. Wigs and fabric supports on the NHS. 2020. at <https://www.nhs.uk/nhs-services/help-with-health-costs/wigs-and-fabric-supports-on-the-nhs/>

128. Johnson A & Montgomery K. NHS Wig Provision in England. A report into NHS England's provision of wigs to Alopecia patients. 2017.

129. U.S. Food and Drug Administration. Patient-Focused Drug Development Public Meeting for Alopecia Areata (Report). 2017. at <https://www.fda.gov/media/112100/download>

130. 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.

131. 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. *Journal of the American Academy of Dermatology* 2015. 72: 640–650.

132. Wiseman MC, Shapiro J, MacDonald N, *et al.* Predictive model for immunotherapy of alopecia areata with diphencyprone. *Arch Dermatol* 2001. 137: 1063–8.

133. 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.

134. Lester RS, Knowles SR & Shear NH. The risks of systemic corticosteroid use. *Dermatologic clinics* 1998. 16: 277–288.

135. Waljee AK, Rogers MA, Lin P, *et al.* Short term use of oral corticosteroids and related harms among adults in the United States: population based cohort study. *bmj* 2017. 357:

136. Rice JB, White AG, Scarpati LM, *et al.* Long-term systemic corticosteroid exposure: a systematic literature review. *Clinical therapeutics* 2017. 39: 2216–2229.

137. Villani A, Fabbrocini G, Ocampo-Candiani J, *et al.* Review of oral minoxidil as treatment of hair disorders: in search of the perfect dose. *Journal of the European Academy of Dermatology and Venereology* 2021. 35: 1485–1492.

138. Chartaux E & Joly P. Long-term follow-up of the efficacy of methotrexate alone or in combination with low doses of oral corticosteroids in the treatment of alopecia areata totalis or universalis. in (2010). 137: 507–513.

139. Firooz A & Fouladi D. Methotrexate plus prednisolone in severe alopecia areata. *American Journal of Drug Discovery and Development* 2013. 3: 188–193.

140. Alkhalifah A, Alsantali A, Wang E, *et al.* Alopecia areata update: part II. Treatment. *J Am Acad Dermatol* 2010. 62: 191–202, quiz 203–4.

141. Strazzulla LC, Wang EHC, Avila L, *et al.* Alopecia areata: an appraisal of new treatment approaches and overview of current therapies. *Journal of the American Academy of Dermatology* 2018. 78: 15–24.

142. Ranchoff RE, Bergfeld WF, Steck WD, *et al.* Extensive alopecia areata. Results of treatment with 3% topical minoxidil. *Cleveland Clinic Journal of Medicine* 1989. 56: 149–154.

143. Randolph M & Tosti A. Oral minoxidil treatment for hair loss: A review of efficacy and safety. *Journal of the American Academy of Dermatology* 2021. 84: 737–746.

144. Sharma AN, Michelle L, Juhasz M, *et al.* Low-dose oral minoxidil as treatment for non-scarring alopecia: a systematic review. *International Journal of Dermatology* 2020. 59: 1013–1019.

145. Ramos PM, Anzai A, Duque-Estrada B, *et al.* Consensus on the treatment of alopecia areata–Brazilian Society of Dermatology. *Anais Brasileiros de Dermatologia* 2021. 95: 39–52.

146. Pfizer. *Study To Evaluate The Efficacy And Safety Profile Of PF-06651600 And PF-06700841 In Subjects With Alopecia Areata*. (clinicaltrials.gov, 2020). at <https://clinicaltrials.gov/ct2/show/NCT02974868>

147. Pfizer. *A Phase 2b/3 randomised, double blind, placebo-controlled,dose ranging study to investigate the efficacy and safety of PF-06651600 in adult and adolescent alopecia areata (AA) subjects with 50% or greater scalp hair loss.* (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT03732807>

148. Pfizer. *Placebo-controlled safety study of ritlecitinib (PF-06651600) in adults with alopecia areata (Allegro2a)*. (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT04517864>

149. Pfizer. *Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT)*. (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT04006457>

150. King B, Guttman-Yassky E, Peeva E, *et al.* Safety and Efficacy of Ritlecitinib and Brepocitinib in Alopecia Areata: Results from the Crossover Open-Label Extension of the ALLEGRO Phase 2a Trial. *JID Innovations* 2022. 2: 100156.

151. Clinicaltrials.gov. PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-2b/3). 2021. at <https://clinicaltrials.gov/ct2/show/NCT03732807>

152. Data on file. *B7981015 Final Protocol 2020. PHASE 2B/3 RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED, DOSE-RANGING STUDY TO INVESTIGATE THE EFFICACY AND SAFETY OF PF-06651600 IN ADULT AND ADOLESCENT ALOPECIA AREATA AA) SUBJECTS WITH 50% OR GREATER SCALP HAIR LOSS*. (2020).

153. King B, Zhang X, Harcha WG, *et al.* From the International ALLEGRO Phase 2b/3 Randomized, Double-Blind, Placebo-Controlled Study (NCT03732807). 20.

154. Pfizer data on file. *Clinical study report: study B7981015 (ALLEGRO 2B/3)*. (2021).

155. Pfizer data on file. *Top Line Report for Study B7981015*. (2021).

156. Sinclair R, Lesiak A & Mehlis B. Long-term safety and efficacy of ritlecitinib in adults and adolescents with alopecia areata: interim results from the ALLEGRO-LT phase 3, open-label study. Presented at: European Academy of Dermatology and Venereology (EADV) Congress; September 7-10, 2022; Milan, Italy. Oral presentation. 2022.

157. Miettinen O & Nurminen M. Comparative analysis of two rates - PubMed. at <https://pubmed.ncbi.nlm.nih.gov/4023479/>

158. Law E, Sherif B & N.J W. Improvements in patient-reported outcomes based on scalp hair regrowth among patients with alopecia areata: analysis of the

ALLEGRO 2b/3 trial. Presented at: EADV September 7-11, 2022; Milan, Italy. Poster P1062./. 2022.

159. Pfizer data on file. *Clinical study report: study B7931005 (ALLEGRO)*. (2019).

160. Winthrop KL. The emerging safety profile of JAK inhibitors in rheumatic disease. *Nat Rev Rheumatol* 2017. 13: 234–243.

161. Pfizer. *A PHASE 2a, RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED STUDY INVESTIGATING THE SAFETY OF RITLECITINIB (PF-06651600) IN ADULT PARTICIPANTS WITH ALOPECIA AREATA*. (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT04517864>

162. Pfizer. *A PHASE 2A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE EFFICACY AND SAFETY PROFILE OF PF-06651600 AND PF-06700841 IN SUBJECTS WITH MODERATE TO SEVERE ALOPECIA AREATA WITH A SINGLE-BLIND EXTENSION PERIOD AND A CROSS-OVER OPEN LABEL EXTENSION PERIOD*. (clinicaltrials.gov, 2020). at <https://clinicaltrials.gov/ct2/show/NCT02974868>

163. Report B. Ethical Principles and Guidelines for the Protection of Human Subjects of Research. The National Commission for Protection of Human Subjects of Biomedical and Behavioral Research. *OPRR Reports* 1979. 4–8.

164. National Institute for Health and Care Excellence. The reference case | Guide to the methods of technology appraisal 2013 | Guidance. 2013. at <https://www.nice.org.uk/process/pmg9/chapter/the-reference-case>

165. Briggs A, Sculpher M & Claxton K. *Decision modelling for health economic evaluation*. (Oup Oxford, 2006).

166. ONS. National Life Tables. Based on data for the years 2018-2020. 2020. at <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables>

167. Davey L, Clarke V & Jenkinson E. Living with alopecia areata: an online qualitative survey study. *British Journal of Dermatology* 2019. 180: 1377–1389.

168. Welsh N & Guy A. The lived experience of alopecia areata: A qualitative study. *Body Image* 2009. 6: 194–200.

169. Barkauskaite R & Serapinas D. Therapeutic implications of psychological state in patients with alopecia areata: A qualitative study. *Dermatologic Therapy* 2020. 33: e14269.

170. Wolf JJ & Hudson Baker P. Alopecia Areata: Factors That Impact Children and Adolescents. *Journal of Adolescent Research* 2019. 34: 282–301.

171. Matzer F, Egger J & Kopera D. Psychosocial Stress and Coping in Alopecia Areata: A Questionnaire Survey and Qualitative Study Among 45 Patients. *Acta dermato-venereologica* 2011. 91: 318–27.

172. Yang Y, Brazier J & Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *The European Journal of Health Economics* 2015. 16: 927–939.

173. Song HJ, Park H, Park S-Y, *et al.* Estimation of Health Utilities Based on the Response to Treatment in Atopic Dermatitis: a Population-based Study. *Clinical Therapeutics* 2019. 41: 700–713.

174. Matusiak Ł, Bieniek A & Szepietowski JC. Psychophysical aspects of hidradenitis suppurativa. *Acta dermato-venereologica* 2010. 90:

175. Kind P, Hardman G & Macran S. *UK population norms for EQ-5D*. (1999).

176. Pfizer data on file. Does the EQ-5D measure the full impact of alopecia areata on patients' quality of life? 2022.

177. Mostaghimi A, Napatalung L, Sikirica V, *et al.* Patient perspectives of the social, emotional and functional impact of alopecia areata: a systematic literature review. *Dermatology and therapy* 2021. 11: 867–883.

178. Wyrwich KW, Winnette R, Bender R, *et al.* Validation of the Alopecia Areata Patient Priority Outcomes (AAPPO) Questionnaire in Adults and Adolescents with Alopecia Areata. *Dermatology and Therapy* 2021. doi:10.1007/s13555-021-00648-z

179. Lai VWY, Bokhari L & Sinclair R. Sublingual tofacitinib for alopecia areata: a roll-over pilot clinical trial and analysis of pharmacokinetics. *International Journal of Dermatology* 2021. 60: 1135–1139.

180. Bewley A, Galvan S, Johansson E, *et al.* Measuring the Burden of Alopecia Areata with the European Quality of Life-5 Dimensions (EQ-5D): Results from a real-world survey in 5 European countries. *Value in Health*

181. Atış G, Tekin A, Ferhatoğlu ZA, *et al.* Type D personality and quality of life in alopecia areata and vitiligo patients: A cross-sectional study in a Turkish population. *TURKDERM-Turkish Archives of Dermatology and Venereology* 2021. 55: 87–91.

182. 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. *Journal of Investigative Dermatology Symposium Proceedings* 2013. 16: S49–S50.

183. Lambert J, Bostoen J, Geusens B, *et al.* A novel multidisciplinary educational programme for patients with chronic skin diseases: Ghent pilot project and first results. *Archives of dermatological research* 2011. 303: 57–63.

184. Venten I, Hess N, Hirschmuller A, *et al.* Treatment of therapy-resistant Alopecia areata with fumaric acid esters. *European journal of medical research* 2006. 11: 300.

185. Nasimi M, Ghandi N, Torabzade L, *et al.* Alopecia Areata-Quality of Life Index questionnaire (reliability and validity of the Persian version) in comparison to Dermatology Life Quality Index. *International Journal of Trichology* 2020. 12: 227.

186. Masmoudi J, Sellami R, Ouali U, *et al.* Quality of life in alopecia areata: a sample of tunisian patients. *Dermatol Res Pract* 2013. 2013: 983804.

187. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. 2022. at <https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741>

188. Stein EM, Yang M, Guerin A, *et al.* Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States. *Health and Quality of Life Outcomes* 2018. 16: 1–12.

189. Sullivan PW, Valuck R, Saseen J, *et al.* A comparison of the direct costs and cost effectiveness of serotonin reuptake inhibitors and associated adverse drug reactions. *CNS drugs* 2004. 18: 911–932.

190. National Institute for Health and Care Excellence. TA403: Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer. 2016. at <https://www.nice.org.uk/guidance/ta403>

191. Worbes-Cerezo M, Nafees B, Lloyd A, *et al.* Disutility Study for Adult Patients with Moderate to Severe Crohn's Disease. *Journal of health economics and outcomes research* 2019. 6: 47.

192. NHS England. National Cost Collection for the NHS - National schedule of NHS costs 2020/21. at <https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/>

193. PSSRU. Unit Costs of Health & Social Care 2021. 2021. at <https://kar.kent.ac.uk/92342/25/Unit%20Costs%20Report%202021%20-%20Final%20version%20for%20publication%20%28AMENDED2%29.pdf>

194. ONS. A01 Summary of Labor Market. 2022.

195. ONS. Annual Survey of Hours and Earnings. 2021.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over [ID4007]

Summary of Information for Patients (SIP)

January 2023

File name	Version	Contains confidential information	Date
Ritlecitinib_Company evidence submission_SIP [Final]	1.0	No	09 January 2023

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Response:
Generic name: Ritlecitinib Brand name: TBC

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

Response:
Ritlecitinib is a treatment for adults and adolescents who are 12 years of age or older with severe alopecia areata. Alopecia areata, also called AA, is a long-term condition that causes hair loss because the body's own defence system attacks normal tissue.[1] In people with AA, hair loss can happen in different parts of the body, in different patterns and of different severity.[2-4]

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

Response:
Ritlecitinib is not yet approved for adults and adolescents who are 12 years of age or older with severe AA. The marketing authorisation application has been submitted by Pfizer and a decision is pending. See Table 3, section B.1.2

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

Response:

Not applicable.

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

Response:

Alopecia areata (AA)

AA is the second most common hair loss condition that causes hair loss without any scarring of the skin.[1] AA is an autoimmune disease, which means that it is caused by the body's own defence system attacking normal tissue.[5] People with AA often have other autoimmune or skin conditions as well.[6] The severity of AA symptoms usually varies over time, with the symptoms being milder for some time before returning worse than before.[5] Almost all people with AA have more than 1 episode of AA, which is when the condition is considered more severe. After an AA episode, hair can grow back but this can take time and usually happens between 3 months to several years after the AA episode.[1] AA can be worsened by psychological stress.[6-9]

The symptoms of AA include hair loss in different parts of the body. People with AA can have different locations, patterns and severity of hair loss.[2-4] Different types of AA vary by the amount of hair lost, as described below:

- People with **severe AA** have at least 50% of hair loss on their scalp.[1, 10-13]
- People with **alopecia totalis, also called AT,** have complete or nearly complete hair loss on their scalp.
- People with **alopecia universalis, also called AU,** have complete or nearly complete hair loss that affects the entire face, scalp and body.

The hair loss in AA is caused by the body's immune cells attacking hair follicles.[3, 5] This causes inflammation in the hair follicles, leading to the hair falling out.[3, 5] Hair follicles are not destroyed, which means that the hair can grow from the follicles again. Approximately half of people with AA also have damage to their fingernails, such as having tiny dents in the nails.[14, 15] This is also known as pitting. More than a third of people with AA have nail damage in both their fingernails and toenails.[14, 15] Nail damage is thought to be linked to more severe AA.[16-18]

Prevalence of different types of AA

AA affects approximately 2% of people in the world.[16, 19, 20] It can affect people of any gender, race or age.[21] In the United Kingdom (UK), AA affects approximately 0.58% of adults.[22] This means that

about 400,000 people in the UK have AA, based on the calculations that include the UK population size of 67,330,000.

Although the number of people with severe AA in the UK is not known, about 0.09% of people in the United States of America (USA) are estimated to have moderate to severe AA.[23] If the prevalence of severe AA was similar in the UK, this would mean that about 60,597 people in the UK would have moderate to severe AA, based on the calculations that include the UK population size of 67,330,000. AT and AU are less common, together affecting between about 0.06% and 0.1% of people in the world.[24]

How does AA affect the patients' quality of life?

AA is a complex condition that affects different people in different ways. These include physical (i.e. hair loss or symptoms secondary to hair loss such as itching and burning), emotional (i.e. grief and distress) and functional ways (i.e. avoiding social activities or impact work and school). Several studies have assessed how AA affects patients' quality of life. Overall, coping with AA is a daily challenge for 85% of people with AA.[14]

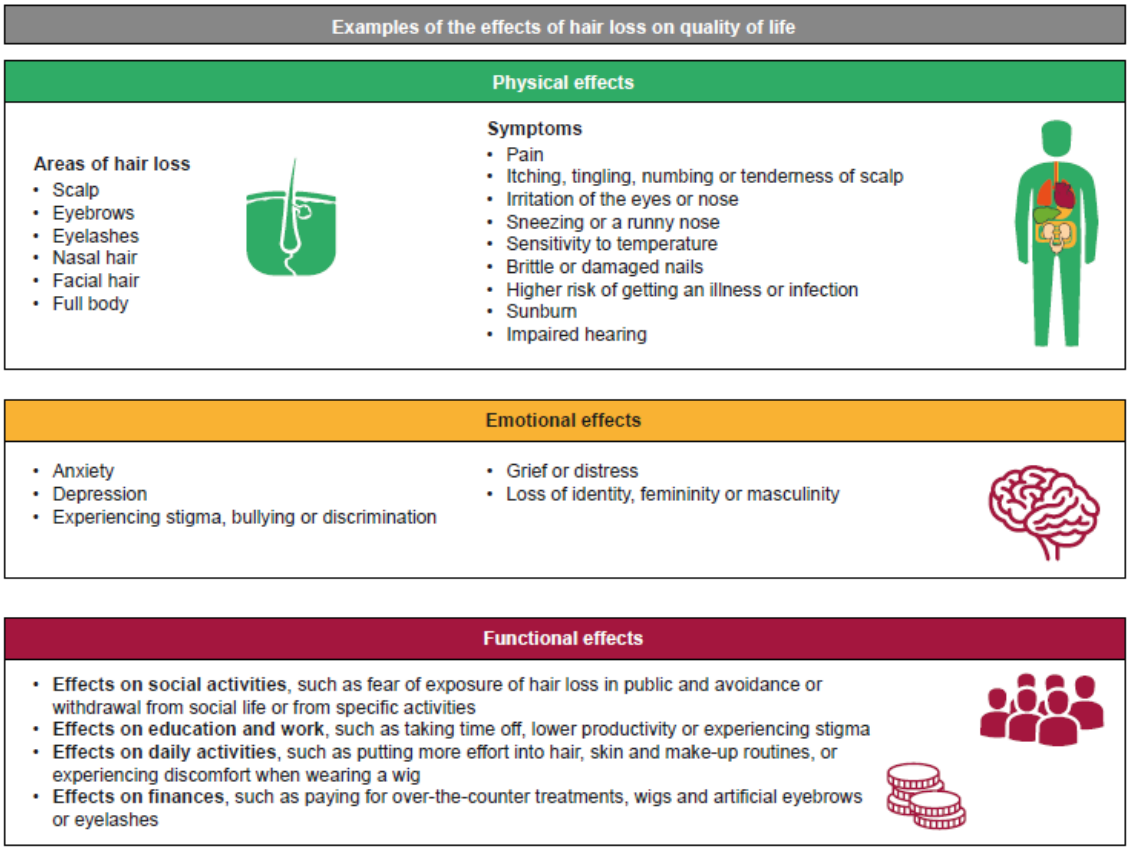
AA appears to have a bigger effect on emotional, psychological and social well-being of people than on their physical well-being. For example, studies found that people with AA are more likely to have anxiety, decreased self-esteem and changed body image,[21, 22, 25-28] and are more likely to avoid people and activities they would usually enjoy.[21, 22, 25-28] The effects of AA on patients' quality of life are thought to increase with AA severity.[8] This has been shown by studies using tests that measure quality of life in people with certain conditions.[29, 30]

In extreme cases patients may suffer from suicidal thoughts or are at risk of taking their own life.[31,32] In a study of 300 patients comparing similar skin conditions, including psoriasis, vitiligo and acne, AA was found to have higher numbers of suicidal ideation. The study also found that the risk of suicide increases with the severity of hair loss. [33]

However, the burden of AA on a patient's quality of life may be difficult to capture. Some studies have shown little evidence that the effects of AA on patients' quality of life increase with AA severity.[34, 35] The majority of such studies measured the effect of AA on quality of life using generic quality of life tests that are not specific to any one condition. The conflicting results of such studies and those using tests specific to certain conditions[34, 35] suggest that generic tests may not capture the full impact AA has on patients.[1, 26, 36, 37]

To understand the literature more fully, insights from people with AA and their clinicians were collected to improve understanding of how severe AA affects patients' quality of life.[38] To do this, 8 clinicians with a special interest in hair disorders and 10 people with severe AA who represented 6 different Patient Advocacy Groups (PAGs) in the UK were interviewed.[38] PAG representatives said that severe AA had a physical, psychological and emotional impact on their quality of life. They also said that they found the emotional and psychological symptoms to be the most bothersome. Figure 1 shows a summary of the effect of severe AA on different aspects of quality of life, as described by PAG representatives in interviews.

Figure 1: Example effects of severe AA on the patients' health-related quality of life, as described by PAG representatives[7]

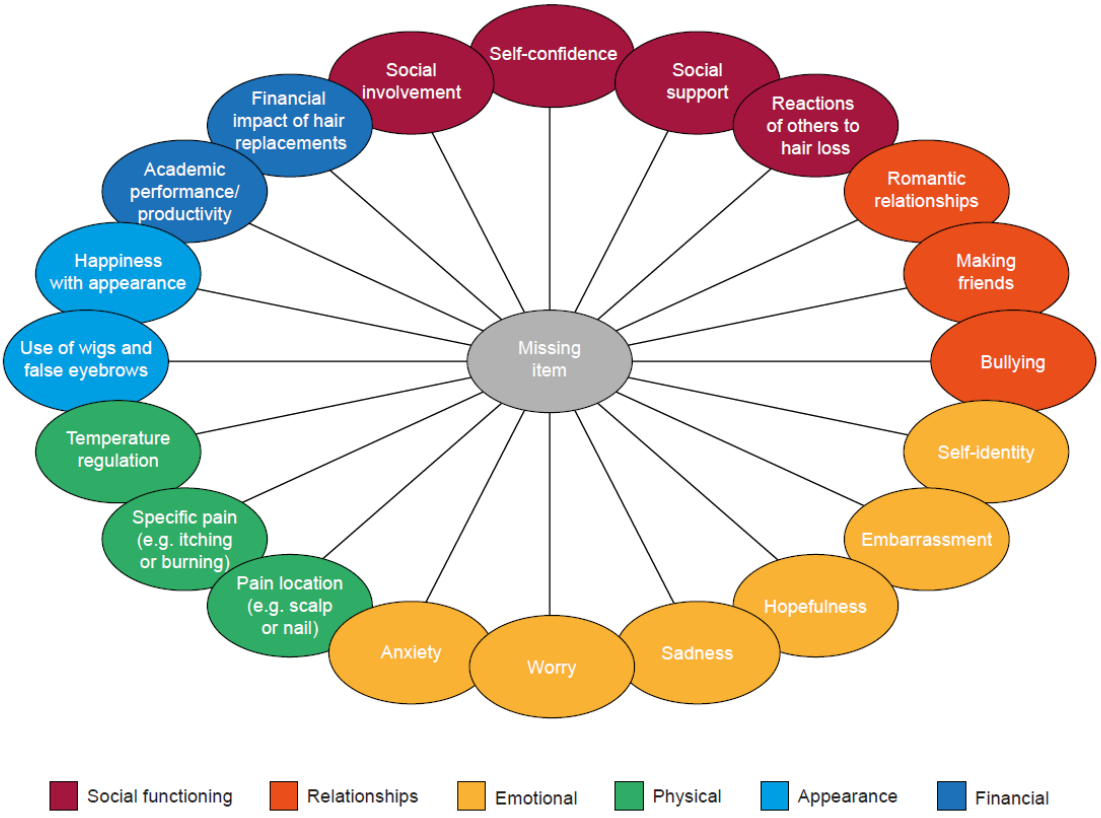


Interviews with PAG representatives confirmed that AA affects many aspects of patients' quality of life, but that many currently used tests do not capture the full effects, are not specific enough and could miss factors that are important to people with AA.[7] Figure 2 below shows the key areas of health-related quality of life for people with AA that are missing from the current generic tests to measure the effect of AA on the quality of life, according to PAG representatives.

Summary of effects of AA on patients' quality of life:

- AA appears to have a bigger effect on emotional, psychological and social well-being of affected people than on their physical well-being.
- The effects of AA on patients' quality of life are thought to increase with AA severity, although not all studies support this.
- Many currently used tools may not capture the full effects of AA on patients' quality of life because they are not specific enough and could miss factors that are important to people with AA.

Figure 2: Aspects of quality of life that are not covered by generic quality of life tests, as described by PAG representatives[7]



Given that the currently available tests may not capture all the impact of AA on patients' and caregivers' quality of life, researchers did a type of study called a vignette study.[39] This means that the researchers described different scenarios of how AA affects different aspects of quality of life in adults and adolescents with AA, as well as in their caregivers. Then, the participants provided feedback on how accurate the descriptions were. The participants included adults and adolescents with AA, as well as their caregivers or healthcare professionals. Finally, members of the UK public were asked to review different scenarios of how AA affects quality of life and decide whether they would consider swapping years lived with quality of life. This vignette approach is one of a number suggested by the National Institute for Health and Care Excellence (NICE) to understand how much burden a condition such as AA has on patients' lives. The results are used in the economic model NICE uses to assess the value of a medicine and is further discussed in section 3i.

How does AA affect the quality of life of families and caregivers of people with AA?

There is limited information about how AA affects the quality of life of families and caregivers (e.g. parents or friends) of people with AA. However, a study of 229 people found that 69.9% of family members of adults with AA and 87.2% of family members of children with AA have a decreased health-related quality of life.[40] The most common impact of AA is on the emotional and social well-being of families and caregivers, with more hair loss in children with AA being linked with worse well-being.[41]

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?

Response:

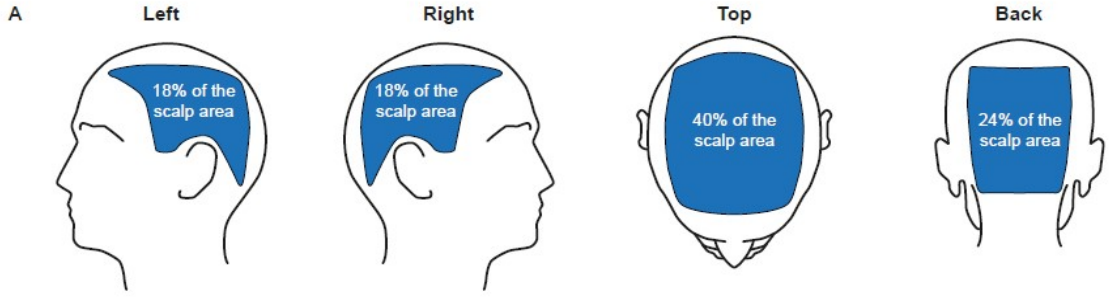
Diagnosis and assessment of severe AA

To diagnose AA, doctors need to do a physical check-up of a person with suspected AA and review their medical history.[28, 42] Currently, there is no standard tool for assessing the severity of AA. However, a tool most commonly used to measure the severity of AA is the Severity of Alopecia Tool, also called SALT.[7, 43]

When using SALT to assess AA severity, doctors measure the percentage of hair loss in 4 different areas of the scalp. Then, they use these measurements to calculate the total SALT score.[43] SALT scores range from 0 to 100, with a lower SALT score meaning less scalp hair loss. A SALT score of 0 means that a person has no scalp hair loss, whereas a SALT score of 100 means that they have complete scalp hair loss summarised in Figure 3.

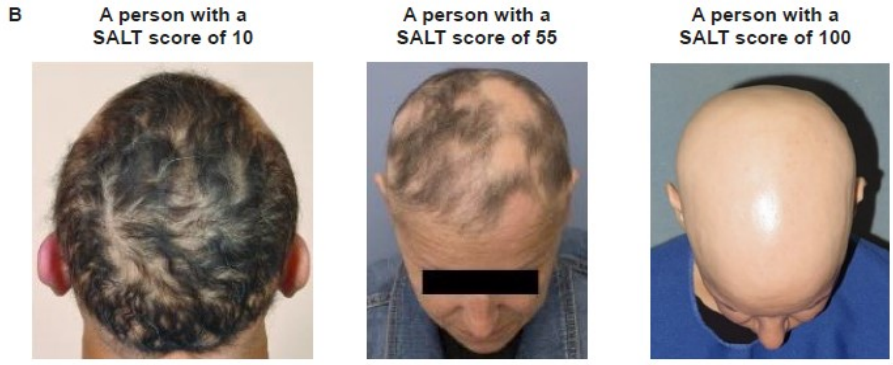
A SALT score of 50 or more is used as a diagnosis of severe AA, which means that people have lost approximately 50% of the hair on their scalp. SALT score cannot be used to measure the effects of AA on quality of life. To better assess the impact of AA on quality of life, doctors also use other tools that capture the opinions of people with AA. You can find more information about those other questionnaires in section 3f.

Figure 3: Example of how SALT scores are calculated[44]



Severity of Alopecia Tool (SALT) score is calculated in 3 steps.

1. First, doctors measure the percentage of hair loss in each scalp area: left, right, top and back of the scalp.
2. Then, they multiply the percentage of hair loss in a scalp area by the size of that scalp area.
3. Finally, they add the 4 values together to get a SALT score for the whole scalp.



Diagnostic test before taking ritlecitinib

Before people start taking ritlecitinib, doctors need to do some tests and measurements. These include:

- measuring the number of cells called platelets, which help blood to clot, and lymphocytes, which are a type of white blood cell, in the blood.
- monitoring people with severe AA during and after treatment with ritlecitinib to see if they develop any signs or symptoms of an infection.
- Checking if people have tuberculosis, also known as TB, or an infection that causes liver inflammation, also known as viral hepatitis.
- Checking the skin of people who have a higher risk of skin cancer.

If a person gets a new infection while taking ritlecitinib, doctors should closely monitor them and diagnose the infection using a test for people with a weakened immune system and treat them accordingly

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

- **What is the treatment pathway for this condition** and **where in this pathway the medicine is likely to be used**? Please use diagrams to accompany text where possible. Please give emphasis to the **specific setting and condition being considered** by NICE in this review. For example, by **referencing current treatment guidelines**. It may be relevant to show **the treatments people may have before and after the treatment under consideration in this SIP**.
- Please also consider:
    - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
    - are there any **drug–drug interactions and/or contraindications** that commonly cause challenges for patient populations? If so, please explain what these are.

Response:

Current treatments for AA

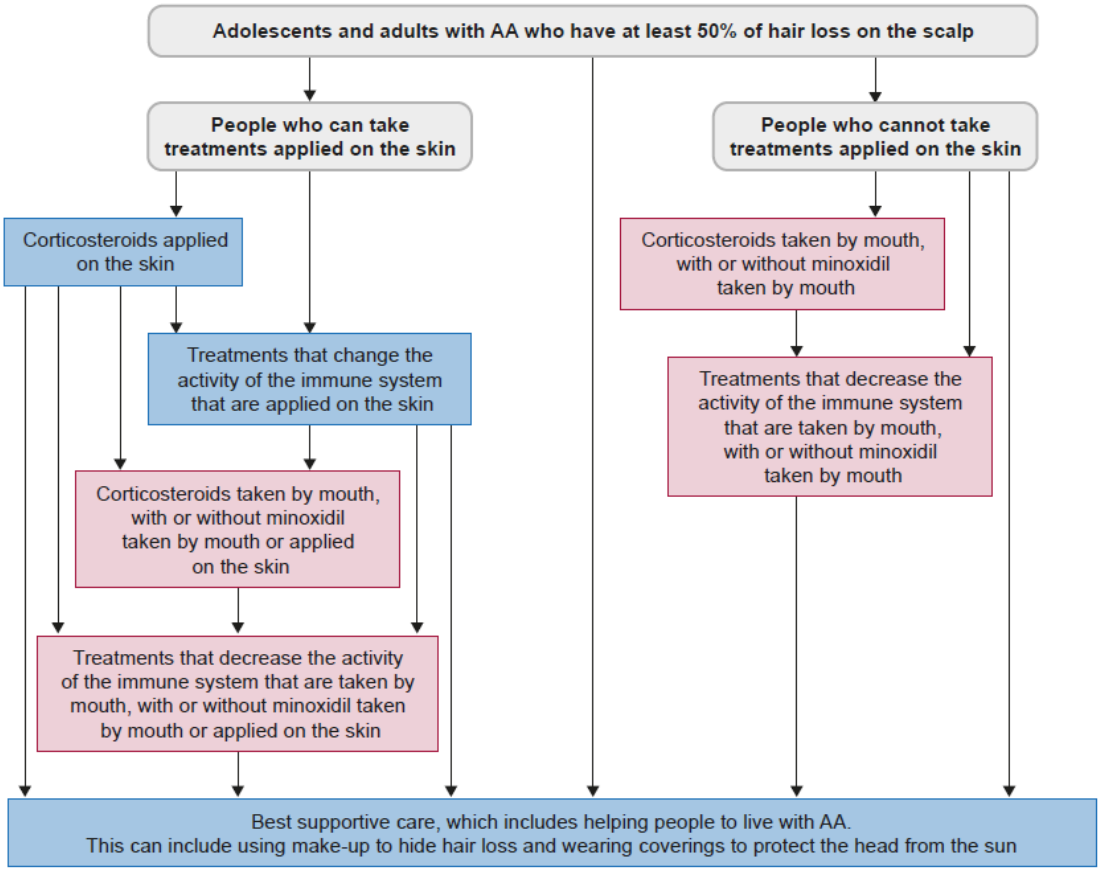
There are currently no up-to-date guidelines on how to treat AA in the UK. The most recent United Kingdom guidelines were published by the British Association of Dermatologists (experts in skin conditions) in 2012.

Doctors treat severe AA with medicines that are not currently approved to be used in this condition. These treatments are called off-licence treatments. Off-licence treatments for AA are used to stop any hair loss from getting worse by calming down the body's immune system or to help the hair to grow back. However, off-licence treatments have not been tested for use in people with AA or severe AA in clinical studies. Therefore, it is not known how safe or effective they are in treating this condition. Figure 4 shows some of the off-licence treatments currently used for treating AA, based on interviews with UK clinicians with a special interest in hair disorders.[45]

Currently, few off-licence treatment options are available for people with AA or severe AA, and not everyone with AA can receive them. A study in the UK found that 46% of 6,765 people with AA did not receive treatment.[42] This could be because of the lack of treatment options available or because treatments are not prescribed by all clinicians. Interviews with clinicians have revealed that some dermatologists are reluctant to prescribe off-licence treatments, especially to younger age groups,[46] making them harder for people with AA to access. For people who do receive treatment that helps to stop hair loss or helps the hair to grow back, the hair loss usually gets worse again after they stop taking the treatment.[38, 47] In the interviews, clinicians and people with severe AA shared that more treatment options are needed.[38, 45]

People with severe AA may try to hide their hair loss by using wigs, head covers, false eyelashes and semi-permanent make-up. However, a study has found that wearing a wig improved the confidence or self-esteem of only 23% of 338 people with AA, and 65% of 338 people worried about being able to afford a new wig.[48] Some people also fear other people finding out that they wear a wig or have difficulties applying wigs and other covers.[28] This shows that these solutions are not effective enough to help people with severe AA to manage their condition or to improve their quality of life.

Figure 4. Overview of treatments used in severe AA



Grey: patient population
Blue: treatments that work in a specific part of the body
Red: treatments that work in the whole body. The only treatments that work in the whole body are off-label treatments and are more commonly given to adults than to adolescents

Ritlecitinib as a treatment for severe AA

Ritlecitinib is being developed as a treatment for severe AA in adults and adolescents who are 12 years of age or older. Ritlecitinib works in the whole body. If ritlecitinib is approved to treat severe AA, it will provide an approved treatment option for adults and the only approved treatment option for adolescents with severe AA.

Clinical studies of ritlecitinib have shown that it is effective in achieving hair regrowth compared with a placebo.[49] Ritlecitinib also has an acceptable safety profile.[49, 50] The number of participants with side effects was similar for those who took ritlecitinib and those who took the placebo.[49, 50] Overall, ritlecitinib improved some but not all measurements of quality of life.[49] You can find more information about the safety and efficacy of ritlecitinib in sections 3e-g.

2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

Response:

The impact of AA and severe AA on quality of life is summarised in section 2a.

A study of patient preferences showed that people with AA are willing to accept an increased risk of side effects (including blood clots, serious infections or cancer) if a treatment improved the chances of their scalp hair growing back from 0% to 50%.[4] Other hair regrowth, such as that on eyebrows and eyelashes, was less important to people with AA.[4] The patient preferences shown in this study suggest that hair regrowth is important to people with AA and that hair loss in AA likely affects their quality of life. This is consistent with several other studies showing that AA has a significant effect on the emotional and psychological well-being of people with AA, as summarised in section 2a.

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

Response:

Ritlecitinib has been developed for treatment of severe AA in adults and adolescents who are 12 years of age or older.

Normally, the immune system protects the body from infections. AA is an autoimmune condition, in people with AA, the immune system gets confused and mistakenly attacks the hair follicles. This causes inflammation in the hair follicles, leading to the hair falling out. Hair follicles are not destroyed, which means that the hair can grow from the follicles again. Ritlecitinib is designed to stop certain proteins in the immune system from working. This calms down the body's immune system and decreases inflammation at the hair follicle. This can help the hair to grow back.

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes / No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

Response:

Ritlecitinib should be taken on its own as a capsule by mouth.

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

Response:

How to take ritlecitinib?

Ritlecitinib is taken as a 50 mg capsule by mouth once daily. Ritlecitinib can be taken with or without food. Taking ritlecitinib as a capsule by mouth means that people with severe AA can take the treatment by themselves and without experiencing any pain or fear of needles.

People with severe AA must stop taking ritlecitinib if they get certain serious side effects. You can find more information about these serious side effects in section 3g.

Pausing ritlecitinib for up to 6 weeks should not cause significant hair loss in the areas where hair has grown back following treatment with ritlecitinib.

Ritlecitinib has several advantages:

- It is more effective at helping the hair to grow back than no treatment.[51] You can find more details in section 3e.
- It can be taken as a capsule by mouth once a day. This means that people with AA need to take treatment less frequently than if they take other treatments for AA, such as those applied on the skin, which are often used twice daily.
- Many other AA treatments are received as an injection. Taking ritlecitinib as a capsule by mouth means that people with AA can take the treatment by themselves and without experiencing any pain.
- Ritlecitinib can also be taken at home so people with AA do not need to visit clinicians like they do for some other treatments. This decreases the burden of treatment on the patient. It also potentially decreases the burden on caregivers if they no longer need to support people with severe AA with attendance to, and transport to and from clinics.

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

Response:

Before a drug is approved for people to take, researchers do clinical studies to find out how it works and understand its safety profile. Several clinical studies of ritlecitinib for treating severe AA have been completed. Key studies are listed below.

ALLEGRO 2a[52,53]

- ALLEGRO 2a is a completed phase 2a study. In a phase 2a study, researchers study how different doses of a drug works and look at the safety profile in a small number of participants with a certain condition.
- In this study, researchers assessed the effectiveness and safety of ritlecitinib and brepocitinib in people with severe AA. Brepocitinib is not a focus of this summary.
- ALLEGRO 2a was carried out in 3 countries: Australia, Canada and the USA.
- The ClinicalTrials.gov identifier for this study is NCT02974868.
- ALLEGRO 2a study involved 142 participants with AA who have at least 50% scalp hair loss.
- The study started in December 2016 and ended in May 2019.
- The main aim of the study was to assess the effectiveness and safety of ritlecitinib and brepocitinib in people with severe AA.

ALLEGRO 2b/3 study[51]

- ALLEGRO 2b/3 is a completed phase 2b/3 study. In a phase 2b/3 study, researchers study how a drug works and look at the safety profile in a large number of participants with a certain condition.
- In this study, the researchers assessed the effectiveness and safety of different doses of ritlecitinib compared with a placebo in people with severe AA.
- ALLEGRO 2b/3 was carried out in 18 countries: Argentina, Australia, Canada, Chile, China, Colombia, the Czech Republic, Germany, Hungary, Japan, Mexico, Poland, Russia, Spain, South Korea, Taiwan, the UK and the USA.
- The ClinicalTrials.gov identifier for this study is NCT03732807.

- ALLEGRO 2b/3 involved 718 participants with severe AA (as determined by a SALT score of at least 50) who had no thick hair regrowth on the scalp during the 6 months before they joined the study.
- The study started in December 2018 and was completed in December 2020.
- The main aim of the study was to assess the percentage of participants whose SALT score decreased to 20 or less during 24 weeks of treatment with ritlecitinib or placebo.

ALLEGRO-LT[50]

- ALLEGRO-LT is an ongoing phase 3 study. In a phase 3 study, researchers study how a drug works and look at the safety profile in a large number of participants with a certain condition.
- In this study, the researchers assessed the long-term safety and effectiveness of ritlecitinib compared with a placebo in people with severe AA.
- ALLEGRO-LT is being carried out in 17 countries: Argentina, Australia, Canada, Chile, China, Colombia, the Czech Republic, Germany, Japan, Mexico, Poland, Russia, Spain, South Korea, Taiwan, the UK and the USA.
- The ClinicalTrials.gov identifier for this study is NCT04006457.
- ALLEGRO-LT involves 449 participants with AA from the ALLEGRO 2b/3 study as well as new study participants with a SALT score of at least 25 who had no thick hair regrowth on the scalp during the 6 months before they joined the study.
- The study started in July 2019 and is ongoing.
- The main aim of the study is to assess the long-term safety of ritlecitinib in people with AA over 36 months of treatment.

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

Response:

How effective is ritlecitinib in treating severe AA?

<u>ALLEGRO 2b/3 study</u>[51]

In this study of people with severe AA, the researchers measured the participants' hair loss by calculating a SALT score at the start of the study, at week 24 and finally at week 48. Patients who started the study had a SALT score of at least 50. This means that they had lost more than 50% of their hair. To measure how effective ritlecitinib was, the researchers counted the number of participants whose SALT score decreased to a score of 20 or less or to a score of 10 or less during treatment.

After 24 weeks of treatment, just under a quarter of participants in the ritlecitinib group had a SALT score of 20 or less, compared with participants in the placebo group. This means that the participants who responded had hair on at least 80% of their scalp.

Similarly, there were more participants who took ritlecitinib who had a SALT score of 10 or less after 24 weeks of treatment, compared with those who took a placebo. This means that the participants who responded had hair on at least 90% of their scalp.

The researchers also found that the number of participants whose SALT scores decrease to 20 or below, or 10 or below, increased over 48 weeks of treatment.

Finally, the researchers measured the changes in participants' eyebrow and eyelash hair during the study. They found that more participants taking ritlecitinib than those taking the placebo had substantial eyebrow and eyelash hair regrowth after 24 and 48 weeks of treatment.

ALLEGRO-LT study[50]

In this ongoing study, the researchers are assessing the long-term safety and effectiveness of ritlecitinib compared with a placebo in people with severe AA.

The data gathered so far show that ritlecitinib has long-term effectiveness that increases over time until at least 24 months of treatment. Similar to the results in ALLEGRO 2b/3, data from ALLEGRO-LT shows that during the treatment period (patients taking ritlecitinib for 24 months) there was an increase in the number of participants who regrew their hair compared with those taking placebo.

Are any outcomes more important to patients than others and why?

A study that analysed treatment preferences in 201 people with AA from the USA and European Union (EU) showed that the most important outcome for people with AA was a higher chance (from 0% to 50%) of having most or all of the hair on their scalp grow back after 24 weeks of treatment.[54] This was more important than avoiding treatment-related risks. The results of the ALLEGRO 2b/3 study show that the safety and effectiveness of ritlecitinib align with the preferences of people with AA.

Are there data limitations that affect how to interpret the results?

One limitation that affects how the results can be interpreted is the relatively short duration of the ALLEGRO 2b/3 study, in which ritlecitinib and a placebo were compared for only 24 weeks. Another limitation is that participants with psychiatric conditions were excluded from the ALLEGRO 2b/3 study. This means that the study was not able to show whether ritlecitinib affects health-related quality of life linked to psychiatric conditions.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

Response:

How effective did people with severe AA think that ritlecitinib was?

In addition to measuring the number of participants with SALT score improvement after treatment with ritlecitinib in ALLEGRO 2b/3 study, the researchers wanted to learn how effective people with severe AA thought ritlecitinib was in reducing hair loss. They wanted to understand the patients' perspective, rather than only clinicians' perspective.

To do this, the participants were asked to fill in 2 different questionnaires:

- **Patient Global Impression of Change, or PGI-C**. The participants completed the PGI-C questionnaire before and at certain time points during their treatment. The PGI-C questionnaire asked the participants how their overall condition had changed since the start of the study. The PGI-C questionnaire helped the researchers to learn how effective the participants thought the treatment was and whether reporting higher improvement was linked to a greater decrease in SALT scores.

  In the ALLEGRO 2b/3 clinical study, the results of the PGI-C questionnaire after 24 weeks of treatment showed that almost half of the participants who took ritlecitinib 50 mg felt that their symptoms had improved moderately or greatly compared with a small proportion (~10%) than those who took the placebo.[51] After 24 weeks of treatment, about two-thirds of participants taking ritlecitinib were satisfied with the amount of hair regrown. In comparison, only one-third of participants taking the placebo were satisfied with hair regrowth.

  In ALLEGRO-LT, the results of the PGI-C questionnaire after 24 months of treatment with ritlecitinib showed that the proportion of participants who felt that their symptoms had improved moderately or greatly increased over time.[50]

- **Patient Satisfaction with Hair Growth, or P-Sat.** In the P-Sat questionnaire, the participants were asked about how satisfied they were with the hair that had regrown after they had started taking the treatment. Of the participants who took ritlecitinib, approximately two thirds said they were satisfied with the amount of hair regrowth after 24 weeks of treatment, compared with a smaller proportion (~16%) of participants who took the placebo. Similar results were seen for satisfaction with the quality of new hair.

How does ritlecitinib affect the quality of life of people with AA, and their families and caregivers?

In the ALLEGRO 2b/3 clinical study, health-related quality of life after treatment with ritlecitinib was assessed using several questionnaires:[51]

- **Alopecia Areata Patient Priority Outcomes, or AAPPO**. AAPPO measured the participants' opinions about the changes in hair loss, emotional symptoms and the activities they were able to do. The results showed that there were more participants taking ritlecitinib with improved hair loss scores than there were participants taking the placebo after 24 weeks of treatment. There were around a quarter of participants taking ritlecitinib 50 mg who achieved a response of either "no hair loss" or "mild scalp hair loss" after 24 weeks of treatment, compared with a small number (less than 10%) of those taking the placebo. However, both participants taking ritlecitinib and those taking the placebo noted only minimal improvements in emotional symptoms or the activities they were able to do after

24 weeks of treatment. It is possible that this is because the participants had few emotional symptoms at the start of the study.

After the researchers analysed the main AAPPO data, they did additional analyses to find out if there were any patterns in the data. They found out that participants whose SALT scores decreased to 20 or below or 10 or below during the study, also had an improvement in emotional symptoms and the activities they were able to do, compared with other participants. This means that the health-related quality of life improved for the participants whose SALT scores decreased to 20 or below or 10 or below during the study.

- **Hospital Anxiety and Depression Scale, or HADS**. The HADS questionnaire was used to measure any changes in anxiety and depression the participants may have had during treatment. The researchers saw only small changes in anxiety and depression scores. But, given that very few people had anxiety or depression before taking the treatment, the researchers were not able to determine whether ritlecitinib affects them compared with a placebo.

- **EQ-5D.** EQ-5D measured the effects of AA on mobility, self-care, usual daily activities, pain, discomfort, anxiety and depression. The EQ-5D questionnaire did not show any effects of AA on quality of life. After the researchers had analysed the main EQ-5D data, they also did additional analyses to find out if there were any patterns in the data. They found that EQ-5D scores were similar for the participants whose SALT scores decreased to 20 or less during treatment and other participants. It is possible that the EQ-5D questionnaire did not show any effects of AA and ritlecitinib on quality of life because the aspects of life assessed with EQ-5D don't capture fully the effects of AA on patients. This has been previously suggested by PAG representatives in interviews[7], and supported by clinicians[45].

- **Short-Form 36, or SF-36.** The SF-36 assesses changes in different aspects of physical and mental health. The SF-36 showed that overall, AA did not affect quality of life, possibly because of the lack of effects on physical health. After the researchers had analysed the main SF-36 data, they also did additional analyses to find out if there were any patterns in the data. They found that the participants whose SALT scores decreased to 20 or below or to 10 or below during treatment had a greater improvement in mental health scores than other participants. However, all participants had similar scores for their physical health. Like EQ-5D, SF-36 may also not show all effects of AA and ritlecitinib on quality of life because it does not measure all the relevant aspects of quality of life.

- **Work Productivity and Activity Impairment, or WPAI.** WPAI is a questionnaire that assesses the ability of people to work, how much work time they miss and how much their work is impaired by a condition. The researchers found that the participants who had more hair regrowth also had a higher improvement in WPAI scores, which means that their work was less affected by AA. The results also showed that AA did not cause the participants to miss work, but that the participants whose SALT scores decreased to 20 or below or to 10 or below during 48 weeks at the study missed less work time and had less work impairment and were able to do their job more effectively than other participants.

Overall, ritlecitinib improved some but not all measurements of quality of life. Generic questionnaires such as EQ-5D and SF-36 may not capture all of the burden AA patients experience, and this is supported by PAG representatives and clinicians. The design of the study and the patients included may have also played a part in explaining the mixed results on health-

related quality of life. The results suggest capturing the health-related quality of life burden for patients with severe AA is challenging. Other approaches may need to be considered to fully capture the impact severe AA has on patients.

To address this researchers did a type of study called a vignette study.[39] This means that the researchers described different scenarios of how AA affects different aspects of quality of life in adults and adolescents with AA, as well as in their caregivers. Then, the participants provided feedback on how accurate the descriptions were. The participants included adults and adolescents with AA, as well as their caregivers or healthcare professionals. Finally, members of the UK public were asked to review different scenarios of how AA affects quality of life and decide whether they would consider swapping years lived with quality of life. This vignette approach is one of a number suggested by the National Institute for Health and Care Excellence (NICE) to understand how much burden a condition such as AA has on patients' lives.[55]

What quality of life measures should be considered in future studies?

Based on the insights gathered in interviews with PAG representatives with severe AA and clinicians, future studies should investigate the effects on emotional and psychological quality of life.[7, 46]

What is important to people with AA when considering safety and efficacy of treatments?

A study of 201 adults with AA from the USA and EU tried to answer this question. In the, participants looked at scenarios showing different safety and efficacy of a treatment[4] and chose which one they preferred.

The study showed that people with AA were willing to accept a higher level of risk of side effects if a treatment worked better. Their key interest was a higher chance of having most or all of their scalp hair grow back. The risk they were least likely to accept regardless of how well a treatment worked, was a higher risk of cancer in the following 3 years. However, 23.4% of participants were willing to accept a higher risk of blood clots, and 18.5% were willing to accept a higher risk of serious infection during 3 years of treatment.

It should be noted that preferences varied between different groups, with adults from the EU being more willing to accept the risk of serious infection and blood clots than adults in the USA. Similarly, eyelash hair growing back was more important for female participants than for male participants. Additionally, a study with 120 adolescents with severe AA from the USA and EU showed that adolescents with severe AA were willing to accept a higher risk of cancer if the treatment had a 50% chance of meaningfully helping their hair to grow back after 24 weeks of treatment.

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

Response:

Ritlecitinib safety

This section is a summary of the medical problems the participants had during the clinical studies. Information about the safety of ritlecitinib in adolescents and adults with severe AA comes from the ALLEGRO 2b/3 and ALLEGRO-LT studies.[49, 50]

The medical problems in these studies are called 'adverse events'. An adverse event is any unfavourable and unintended sign, symptom or disease that a participant may have during a study. An adverse event is considered 'serious' when it is life-threatening, causes persistent or significant disability, birth defect or requires hospital care. Doctors keep track of all the adverse events that happen in studies, even if they do not think the adverse events might be related to the study drug.

The safety and tolerability of ritlecitinib in participants with severe AA was studied ALLEGRO 2b/3 and ALLEGRO-LT studies. The results showed that:

- Ritlecitinib was well-tolerated over short-term and long-term (up to 48 weeks)
- Most adverse events were mild, went away on their own over time, and rarely caused the participants to stop their treatment or leave the study.
- The most common adverse events were cold, also known as nasopharyngitis, headache, acne, diarrhea, nose and throat infection, nausea, inflammation of hair follicles and itchy red and dry skin.
- There were 2.4% of participants who had serious adverse events during 24 weeks of treatment. This was 28 out of 1165 participants. The number of participants who had serious adverse events was similar in the ritlecitinib and placebo groups.
- The results suggest that the safety of ritlecitinib is similar for adults and adolescents with severe AA.

The tables below show the safety results for ALLEGRO 2b/3 and ALLEGRO-LT studies. Different doses of ritlecitinib were studied in the ALLEGRO clinical trials, however, the 50mg dose has been chosen for treatment when ritlecitinib is approved.

	Placebo	Ritlecitinib				
	Placebo (out of 131 participants)	200/50 mg (out of 131 participants)	200/30 mg (out of 129 participants)	50 mg (out of 130 participants)	30 mg (out of 132 participants)	10 mg (out of 62 participants)
The number of participants who stopped the treatment or left the study due to adverse events	2 (1.5%)	4 (3.1%)	None	2 (1.5%)	4 (3.0%)	0
Number of participants with serious adverse events	2.3 % (3/131)	3.1% (4/131)	None	None	0.8% (1/132)	3.2% (2/62)
Number of participants with adverse events	71.0% (93/131)	73.3% (96 /131)	70.5% (91/129)	75.4% (98/130)	72.7% (96 /132)	69.4% (43/62)

Most common adverse events:						
Nose and throat infection	7.6% (10/131)	12.3% (16/131)	7.8% (10/129)	6.2% (8/130)	8.3% (11/132)	3.2% (2/62)
Cold	6.1% (8/131)	11.5% (15/131)	14.0% (18/129)	(10.0% (13/130)	12.1% (16/132)	9.7% (6/62)
Headache	8.4% (11/131)	8.4% (11/131)	7.8% (10/129)	9.2% (12/130)	15.2% (20/132)	17.7% (11/62)

Table 1. Safety results (most common adverse events) in ALLEGRO 2b/3 after 24 weeks

The researchers also monitored serious adverse events the participants had during treatment. The serious adverse events the participants had during 48 weeks of treatment were:

- Suicidal behaviour
- Painful swelling of the appendix
- Spontaneous abortion
- Blood poisoning that happens when bacteria and their toxins circulate in the blood and lead to organ damage
- Breast cancer
- Breast cancer type that begins in the milk-producing glands in the breast
- Chemical poisoning
- Clot in a blood vessel in the lungs
- Gut disease
- Heavy menstrual bleeding
- Itchy, red and dry skin
- Pus in the cavity between the lungs and the membrane that surrounds it
  Unexplained blindness, paralysis or other nervous system symptoms

The ALLEGRO-LT study is ongoing, but the safety data up until 24 months of treatment are already available. The safety results of ritlecitinib in ALLEGRO-LT are similar to those reported in ALLEGRO 2b/3. Table 2 below summarizes the adverse events, serious adverse events and adverse events that caused the participants to stop the treatment during the 24 months of treatment with ritlecitinib. The table shows only the most common adverse events that happened in at least 10% of participants. There were other adverse events but those happened in fewer participants.

Table 2. Safety results in ALLEGRO-LT after 24 months

	Ritlecitinib (out of 447 participants)
The number of participants who stopped the treatment or left the study due to adverse events	22 (4.9%)
Number of participants with serious adverse events	18 (4.0%)
Number of participants with adverse events	350 (78.3%)
Most common adverse events:	
Headache	73 (16.3%)
SARS-CoV-2 test positive	60 (13.4%)
Acne	52 (11.6%)

How can adverse events be managed?

Most adverse events with ritlecitinib are mild and likely to go away on their own over time.

However, people with severe AA must stop taking ritlecitinib if they:

- get a serious or an opportunistic infection while taking ritlecitinib. An opportunistic infection affects people with a weaker immune system more often or more severely than people with a healthy immune system. People who get a serious or an opportunistic infection should stop taking ritlecitinib until the infection is controlled.
- have blood abnormalities, such as having problems with certain cells in the blood. They should pause or stop taking ritlecitinib.

What are the benefit and risk trade-offs?

The benefit and risk trade-offs of ritlecitinib, which described what is important to people with AA when considering safety and efficacy of treatment, are summarised in section 3f.

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration
-

Response:

There are currently no NICE approved treatments for AA. There is one licensed therapy that is currently being assessed by NICE separately to ritlecitinib. All other treatments are used off-licence and there are concerns over their efficacy and safety. If ritlecitinib was approved for people to take, it would provide another treatment option for adults with severe AA and it would be the only option for adolescents with severe AA. Ritlecitinib can be taken by mouth as a capsule once daily, which makes it easy for people with AA to take.

Ritlecitinib addresses the unmet need for a treatment for severe AA that is effective and has few side effects. Clinical studies of ritlecitinib have shown that it is effective in decreasing hair loss compared with a placebo.[49, 50] Ritlecitinib also has a favourable safety profile, and the number of participants with adverse events is similar for those who took ritlecitinib and those who took the placebo.[56] Overall, ritlecitinib improved some but not all measurements of quality of life.[49]

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers?
- Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration
- What is the impact of any disadvantages highlighted compared with current treatments

Response:

In ALLEGRO 2b/3 and ALLEGRO-LT studies, the most common adverse events for the participants taking ritlecitinib were cold, headache, acne, diarrhoea, nose and throat infection, nausea, inflammation of hair follicles and itchy red and dry skin. There were 2.4% of participants who had serious adverse events during 24 weeks of treatment. This was 28 out of 1165 participants. You can find more information about adverse events and serious adverse events in section 3g.

3i) Value and economic considerations

Introduction for patients:

Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model.

In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on:

- The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?)
- If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)?
- How the condition, taking the new treatment compared with current treatments affects your quality of life.

Response:

A health economic model has been developed to assess the value of ritlecitinib to the national health service (NHS) and personal social services in the UK. Personal social services include services provided by local organizations for people with different physical or mental health needs. The health economic model considers several factors that are important in assessing how a drug impacts people's lives, as well as its financial effects and the cost to the NHS in patient care (known as cost-effectiveness).

The health economic model was created based on feedback from NICE and clinicians with a special interest in hair disorders. It estimated how cost-effective ritlecitinib is for people with AA compared with the standard of care from the perspective of the UK's NHS. Standard of care in this health economic model means the participants receive no pharmacological treatment. The health economic model was based on the results from the ALLEGRO 2b/3 and ALLEGRO-LT studies, and it took into consideration adults and adolescents with AA with a baseline SALT score of at least 50, which means that they have severe AA. The study data were used for the first 144 weeks of the health economic model. From then on, response rates from participants were estimated using a mathematical approach called extrapolation. The health economic model considered the cost of ritlecitinib, number of hospital visits and medical tests, the costs linked to the treatment of adverse events, as well as the costs related to how much and how well people are able to work.

Based on the assumptions, the health economic model showed that the value of ritlecitinib when taken as 50mg capsule once daily is lower than the threshold that NICE would normally consider a

drug to be cost-effective. Ritlecitinib may therefore be considered an appropriate use of NHS resources for people with severe AA.

Uncertainty

Until recently, there were no effective treatments for severe AA. Because of this, the care pathway for people with severe AA, as well as the costs related to health care are not well known. Therefore, several assumptions made in the health economic model. These included assumptions about the cost of treatments, adverse events and administration and monitoring, and the costs linked to health-related quality of life, as well as the percentage of people who are likely to stop taking the treatment. These assumptions mean the outcomes of the economic model have some uncertainty.

The health economic model tested what impact this uncertainty has on whether ritlecitinib is a good use of NHS resources. The assumption that had the greatest effect on the value of ritlecitinib was the number used to represent the quality of life for patients with greater than 50% hair loss (i.e. severe hair loss). However, even when several assumptions were changed several times using a computer, ritlecitinib was found to be a good use of NHS resources in every one of the 10,000 times the model was repeated.

Additionally, there are certain benefits and advantages of ritlecitinib that are not captured by the model. For example, ritlecitinib is expected to have more effects on costs than those taken into consideration by the model. This is because the model focuses on the costs related to healthcare only. However, studies have shown that AA affects the ability of people to work, or to work well.[57] Furthermore, an epidemiological study in the UK has shown that AA more commonly affects people from urban and deprived areas.[42] This means that the people with the greatest need for treatment for AA, whether pharmacological or cosmetic, are less likely to be able to pay for it. The model does not take this into account.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.
If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

Response:

There is currently only one approved treatment for adults with severe AA and none currently approved by NICE. All other treatments are used off-licence and there are concerns over their efficacy and safety. If ritlecitinib was approved for people to take, it would provide another treatment option for adults with severe AA, and it would be the only option for adolescents with severe AA. The key benefits and the innovation of ritlecitinib are summarised in section 3h.

3k) Equalities

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

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

Response:

Inequalities in access to current treatment and solutions:

There is currently only one treatment approved for adults with severe AA in the UK and none for adolescents with severe AA. Some people with severe AA may receive off-licence treatments that are applied to the skin and change how the immune system works. However, these are not widely available in the UK and can be time-consuming because people need to visit a clinic several times over several months to get these treatments. This means that these treatments aren't suitable for everyone with severe AA.

Other inequalities include the access to wigs for people with AA. Some local NHS organizations have a limited number of wigs available, whereas others have none.[7, 58] Additionally, some NHS trusts consider AA to be a cosmetic issue rather than a medical problem. For this reason, they provide no wigs to people with AA or only do so at a cost.[58] Finally, some people with AA cannot wear wigs: for example, if they also wear hearing aids.[58]

AA also affects people of different races, ethnicities, or socioeconomic backgrounds differently. Research has found that AA is more common in people with skin of colour, especially in people of Asian backgrounds.[42] From a socioeconomic perspective, people from more deprived backgrounds in urban areas are more likely to have AA and less likely to be able to afford high quality wigs.[42]

Inequalities in access to ritlecitinib:

Ritlecitinib is planned to be free at the point of need and the price will be agreed by the department of health and social services. NICE will determine whether the price is a good use of NHS resources.

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc.
Where possible, please provide open access materials or provide copies that patients can access.

Response:

Further information on NICE and the role of patients**:**

- Public Involvement at NICE: https://www.nice.org.uk/about/nice-communities/nice-and-the-public/public-involvement
- NICE's guides and templates for patient involvement in health technology assessments: https://www.nice.org.uk/about/nice-communities/nice-and-the-public/public-involvement/support-for-vcs-organisations/help-us-develop-guidance/guides-to-developing-our-guidance

- EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/
- EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf
- National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/
- INAHTA: http://www.inahta.org/
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

Further information on AA:

- Alopecia UK: https://www.alopecia.org.uk/
- National Alopecia Areata Foundation: https://www.naaf.org/
- Support Groups – British Hair and Nail Society: https://bhns.org.uk/index.php?/support_groups_new.html/
- British Association of Dermatologists: https://www.bad.org.uk/
- List of UK hair loss charities: https://www.belgraviacentre.com/

4b) Glossary of terms

Response:

AA – alopecia areata

AAPPO – Alopecia Areata Patient Priority Outcomes

AAQ – Alopecia Areata Quality of Life

AT – alopecia totalis

AU – alopecia universalis

DLQI – Dermatology Life Quality Index

EU – European Union

HADS – Hospital Anxiety and Depression Scale

NHS – National Health Service

P-Sat – Patient Satisfaction with Hair Growth

PAG – Patient Advocacy Group

PGI-C – Patient Global Impression of Change

SALT – Severity of Alopecia Tool

SF-36 – Short-Form 36

UK – United Kingdom

USA – United States of America

WPAI – Work Productivity and Activity Impairment

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

Response:

1. Pratt CH, King LE, Jr., Messenger AG, Christiano AM, Sundberg JP. Alopecia areata. Nat Rev Dis Primers. 2017;3:17011.
2. Cranwell WC, Lai VW, Photiou L, Meah N, Wall D, Rathnayake D, et al. Treatment of alopecia areata: an Australian expert consensus statement. Australas J Dermatol. 2019;60(2):163-70.
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(2):51-60.
4. Fonda-Pascual P, Vano-Galvan S, Garcia-Hernandez MJ, Camacho F. Alopecia areata sisaipho: clinical and therapeutic approach in 13 patients in spain. Int J Trichology. 2016;8(2):99-100.
5. Guo H, Cheng Y, Shapiro J, McElwee K. The role of lymphocytes in the development and treatment of alopecia areata. Expert Rev Clin Immunol. 2015;11(12):1335-51.
6. Hordinsky MK. Overview of alopecia areata. J Investig Dermatol Symp Proc. 2013;16(1):S13-5.
7. Pfizer data on file. Qualitative research in alopecia. Targeted literature review and patient advocacy group representative interviews in the United Kingdom. 2022.
8. Tan E, Tay YK, Goh CL, Chin Giam Y. The pattern and profile of alopecia areata in Singapore--a study of 219 Asians. Int J Dermatol. 2002;41(11):748-53.
9. Toussi A, Barton VR, Le ST, Agbai ON, Kiuru M. Psychosocial and psychiatric comorbidities and health-related quality of life in alopecia areata: a systematic review. J Am Acad Dermatol. 2021;85(1):162-75.
10. European Medicines Agency. Committee for medicinal products for human use (CHMP) advice letter (EMEA/H/SA/3875/1/2018/HTA/III). 2018.
11. National Institute for Health and Care Excellence. Clinical knowledge summaries: alopecia areata 2018 [Available from: https://cks.nice.org.uk/topics/alopecia-areata/.
12. Wambier CGK, B. Rethinking the classification of alopecia areata. 2019.
13. Wyrwich KW, Kitchen H, Knight S, Aldhouse NVJ, Macey J, Nunes FP, et al. The Alopecia Areata Investigator Global Assessment scale: a measure for evaluating clinically meaningful success in clinical trials. Br J Dermatol. 2020;183(4):702-9.

14. Mesinkovska N, King B, Mirmirani P, Ko J, Cassella J. Burden of illness in alopecia areata: a cross-sectional online survey study. J Investig Dermatol Symp Proc. 2020;20(1):S62-S8.
15. Roest YBM, van Middendorp HT, Evers AWM, van de Kerkhof PCM, Pasch MC. Nail involvement in alopecia areata: a questionnaire-based survey on clinical signs, impact on quality of life and review of the literature. Acta Derm Venereol. 2018;98(2):212-7.
16. Villasante Fricke AC, Miteva M. Epidemiology and burden of alopecia areata: a systematic review. Clin Cosmet Investig Dermatol. 2015;8:397-403.
17. You HR, Kim SJ. Factors associated with severity of alopecia areata. Ann Dermatol. 2017;29(5):565-70.
18. Zeberkiewicz M, Rudnicka L, Malejczyk J. Immunology of alopecia areata. Cent Eur J Immunol. 2020;45(3):325-33.
19. Lee HH, Gwillim E, Patel KR, Hua T, Rastogi S, Ibler E, et al. Epidemiology of alopecia areata, ophiasis, totalis, and universalis: A systematic review and meta-analysis. J Am Acad Dermatol. 2020;82(3):675-82.
20. Vu BK, Tuson H, Harricharan S, et al. Epidemiology of alopecia areata across global regions-a systematic literature review [Abstract]. ISPOR Europe; Nov 6-9; Vienna, Austria2022.
21. Strazzulla LC, Wang EHC, Avila L, Lo Sicco K, Brinster N, Christiano AM, et al. Alopecia areata: Disease characteristics, clinical evaluation, and new perspectives on pathogenesis. J Am Acad Dermatol. 2018;78(1):1-12.
22. Xu L, Liu KX, Senna MM. A practical approach to the diagnosis and management of hair loss in children and adolescents. Front Med (Lausanne). 2017;4:112.
23. Benigno M, Anastassopoulos KP, Mostaghimi A, Udall M, Daniel SR, Cappelleri JC, et al. A large cross-sectional survey study of the prevalence of alopecia areata in the United States. Clin Cosmet Investig Dermatol. 2020;13:259-66.
24. Pfizer data on file. Summary of sÆfety works analysis methods and results. 2018.
25. Aldhouse NVJ, Kitchen H, Knight S, Macey J, Nunes FP, Dutronc Y, 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(1):76.
26. Fricke VAC, Miteva M. Epidemiology and burden of alopecia areata: a systematic review. Clin Cosmet Investig Dermatol. 2015;8:397-403.
27. Harries MJ, Sun J, Paus R, King LE, Jr. Management of alopecia areata. BMJ. 2010;341:c3671.
28. Messenger AG, McKillop J, Farrant P, McDonagh AJ, Sladden M. British Association of Dermatologists' guidelines for the management of alopecia areata 2012. Br J Dermatol. 2012;166(5):916-26.
29. Chernyshov PV, Tomas-Aragones L, Finlay AY, Manolache L, Marron SE, Sampogna F, et al. Quality of life measurement in alopecia areata. Position statement of the European Academy of Dermatology and Venereology Task Force on quality of life and patient oriented outcomes. J Eur Acad Dermatol Venereol. 2021;35(8):1614-21.
30. Okhovat JP, Marks DH, Manatis-Lornell A, Hagigeorges D, Locascio JJ, Senna MM. Association between alopecia areata, anxiety, and depression: a systematic review and meta-analysis. J Am Acad Dermatol. 2019.
31. Gupta MA, Gupta AK. Depression and suicidal ideation in dermatology patients with acne, alopecia areata, atopic dermatitis and psoriasis. Br J Dermatol. 1998;139(5):846-50.
32. Velez-Muniz RDC, Peralta-Pedrero ML, Jurado-Santa Cruz F, Morales-Sanchez MA. Psychological Profile and Quality of Life of Patients with Alopecia Areata. Skin Appendage Disord. 2019;5(5):293-8.
33. Layegh P, Arshadi H, Shahriari S. A comparative study on the prevalence of depression and suicidal ideation in dermatology patients suffering from psoriasis, acne, alopecia areata and vitiligo. Iranian Journal Of Dermatology 2010;13.

34. Burge R, Anderson P, Austin J, et al. The patient-reported burden of alopecia areata by current severity: a real-world study in the United States [Poster 26158]. American Academy of Dermatology (AAD); March 19-23; San Francisco, CA2021.
35. Ghajarzadeh M GM, Kheirkhah S. Depression and quality of life in Iranian patients with Alopecia Areata. Iranian Journal of Dermatology. 2011;14:140–3.
36. 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(4):806-12.
37. 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(3):561-71.
38. Pfizer data on file. Qualitative research in alopecia. Targeted literature review and patient advocacy group representative interviews in the United Kingdom. 2022.
39. Pfizer data on file. Vignette study for utility estimation in Alopecia Areata. 2022.
40. Liu LY, King BA, Craiglow BG. Alopecia areata is associated with impaired health-related quality of life: a survey of affected adults and children and their families. J Am Acad Dermatol. 2018;79(3):556-8 e1.
41. Putterman E, Patel DP, Andrade G, Harfmann KL, Hogeling M, Cheng CE, et al. Severity of disease and quality of life in parents of children with alopecia areata, totalis, and universalis: A prospective, cross-sectional study. J Am Acad Dermatol. 2019;80(5):1389-94.
42. 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. Br J Dermatol. 2022;186(2):257-65.
43. Olsen E, Hordinsky M, McDonald-Hull S, Price V, Roberts J, Shapiro J, et al. Alopecia areata investigational assessment guidelines. National Alopecia Areata Foundation. J Am Acad Dermatol. 1999;40(2 Pt 1):242-6.
44. Olsen EA, Hordinsky MK, Price VH, Roberts JL, Shapiro J, Canfield D, et al. Alopecia areata investigational assessment guidelines--Part II. National Alopecia Areata Foundation. J Am Acad Dermatol. 2004;51(3):440-7.
45. Pfizer data on file. Elicitation of expert opinion to aid understanding of current therapeutic landscape in the UK for people with alopecia areata. Final report. 2022.
46. Pfizer data on file. UK interviews with dermatologists with a specialist interest in hair disorders for the use of ritlecitinib for patients with AA. 2022.
47. Rossi A, Muscianese M, Piraccini BM, Starace M, Carlesimo M, Mandel VD, et al. Italian Guidelines in diagnosis and treatment of alopecia areata. Giornale italiano di dermatologia e venereologia: organo ufficiale, Societa italiana di dermatologia e sifilografia. 2019;154(6):609-23.
48. Montgomery K, White C, Thompson A. A mixed methods survey of social anxiety, anxiety, depression and wig use in alopecia. BMJ Open. 2017;7(4):e015468.
49. Pfizer. A Phase 2b/3 randomised, double blind, placebo-controlled,dose ranging study to investigate the efficacy and safety of PF-06651600 in adult and adolescent alopecia areata (AA) subjects with 50% or greater scalp hair loss. (clinicaltrials.gov, 2022) 2022 [Available from: https://clinicaltrials.gov/ct2/show/NCT03732807.
50. Pfizer. Long-term PF-06651600 for the treatment of alopecia areata (ALLEGRO-LT). (clinicaltrials.gov, 2022) 2022 [Available from: https://clinicaltrials.gov/ct2/show/NCT04006457.
51. Pfizer data on file. Clinical study report. Study B7981048: alopecia areata benefit-risk trade-off study. 2022.
52. Pfizer. Study to evaluate the efficacy and safety profile of PF-06651600 and PF-06700841 in subjects with alopecia areata. 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT02974868.
53. King B, Guttman-Yassky E, Peeva E, Banerjee A, Zhu L, Zhu H, et al. Safety and Efficacy of Ritlecitinib and Brepocitinib in Alopecia Areata: Results from the Crossover Open-Label Extension of the ALLEGRO Phase 2a Trial. JID Innovations. 2022;2(6):100156.
54. Pfizer data on file. Clinical study report: study B7981015 (ALLEGRO 2B/3). 2021.

55. 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.
56. Pfizer data on file. Clinical overview module 2.5 (EU). 2022.
57. Fridman M, Ray M, Shy M, et al. The association of alopecia areata-related emotional symptoms with work productivity and daily activity among patients with alopecia areata [abstract]. EADV; September 29-October 2; Virtual Congress2021.
58. Johnson AM, K. NHS Wig Provision in England. A report into NHS England's provision of wigs to Alopecia patients. 2017.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over [ID4007]

Clarification questions

January 2023

Notes for company

Highlighting in the template

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

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

Section A: Clarification on effectiveness data

Literature searching and systematic literature review

A1. CS Appendix D, Table 1, page 2. In Table 1, interventions and comparators are listed together. Please clarify whether ritlecitinib was considered the intervention and the other treatments as comparators for this SLR, or whether all treatments listed were considered potential interventions in a much broader review of AA treatments.

All treatments listed in Appendix D (Table 1, page 2) were considered as potential pharmacological options to treat patients with alopecia areata (AA) of any severity in any geographical location. The treatments included in the search ranged from oral and topical medications (corticosteroids, immunomodulators, Janus kinase inhibitors (JAK) inhibitors, biologics, etc.), to variously administered therapies (contact immunotherapy, platelet rich plasma injections, etc.). A broader review question suited the objective of the systematic literature review (SLR) and it provided a comprehensive overview of the current evidence on all available treatments for AA. It's important to note that the availability of treatments varies by geographical location particularly contact immunotherapy and the availability and use of off licence treatments more generally such as JAK inhibitors and immunosuppressants.

A2. CS Appendix D, Table 1, page 2. Please clarify whether any form of non-pharmacological clinical management (potentially including placebo treatment) was considered a comparator in the SLR.

Non-pharmacological clinical management was not considered as an intervention or comparator in the clinical SLR. All studies must have included an active pharmacological treatment arm for inclusion. However, studies that compared pharmacological treatment to placebo were included in the review. Moreover, within studies, non-pharmacological treatment may have been used alongside investigational treatment and it would be an add on to pharmacological treatment rather than an alternative to it. For example, in the ALLEGRO 2b/3 study, patients were able to continue using non-pharmacological clinical management such as wigs alongside the investigational treatments (ritlecitinib or placebo).

A3. CS Appendix D, Figure 1, page 5. Step 1 in Figure 1 notes that title and abstract screening was technology aided. Please clarify which technology aided title and abstract screening and data extraction, and how it was used in the process.

The option of using software tools or platforms at different steps of the systematic review process was discussed in the initial stages of planning and designing the original SLR. This included using DistillerSR® for the management of the SLR during title and abstract review step or applying digitisation software for extracting data presented in graphs. However, during the SLR execution, technology-aided review and/or data extraction was not required. Therefore, all the review steps and data extraction were conducted manually, following the methodology described briefly below and in Appendix D, pages 2 to 21.

Publications identified through the systematic review were evaluated in a three-step process to assess whether they should be included for data extraction.

Step 1: Abstract review:

All publications were reviewed against the inclusion/exclusion criteria based on their abstract and title by two reviewers independently. Any conflicts between the two reviewers were resolved by a third independent reviewer. All papers included by the reviewers at the end of this stage were retained for Step 2. Papers excluded at this

abstract review stage were disregarded and the rejection reason was documented for use in the PRISMA flow diagram.

Step 2: Full-text review:

Publications included after abstract review (from Step 1) were obtained for a full review of the text. All papers included after the full-text review were retained for data extraction. A record was kept of papers excluded at this stage along with a clear justification for their exclusion.

Two reviewers screened all citations and full-text articles and any discrepancies in their decisions were resolved by a third independent reviewer. To inform the data collection phase of the review, details for the inclusion/exclusion criteria were consulted during this step. This ensured that all decisions regarding the inclusion and exclusion of studies were kept consistent throughout the review process.

Step 3: Data extraction:

Data from included studies (from Step 2) were extracted into a pre-defined Excel-based template, ensuring that data were extracted uniformly and were comparable across studies. Two analysts independently extracted data and their results were checked and reconciled by a third independent analyst.

When the update to the clinical SLR was conducted, screening and data extraction were both carried out manually by two independent reviewers in Excel spreadsheets. No other technology was used. The methodology for abstract review, full-text review and data extraction was performed in the same way as for the original SLR.

A4. CS Appendix D. Please clarify which studies were included in the SLR, to address the review question, and provide an explicit number of included studies for which ritlecitinib is the intervention. Please explain any discrepancies with the number reported in Figures 2 (131 studies reported in 134 papers) and 3 (a further 25 studies reported in 29 papers).

As indicated in Table 7, Appendix D, 163 publications were identified in the SLR after combining the findings of the original SLR (which identified 134 publications) and the updated SLR (which identified 29 publications). Amongst these 163 publications, 156

unique studies were identified; 131 studies were identified in the original SLR and 25 studies were identified in the SLR update.

Figure 2, Appendix D graphically presents findings of the original clinical SLR, which searched databases from January 2000 until October 2021. Figure 3, Appendix D presents the findings of the updated SLR, which searched databases from October 2021 to September 2022.

As depicted in **Error! Reference source not found.** (which is an excerpt of Table 7, Appendix D), two studies contained ritlecitinib as the intervention.

Table 1: Studies included from the SLR that contained ritlecitinib as an intervention

	Study	Citations	Intervention	Comparator	Population	Disease severity
RCTs and non-RCTs						
1	ALLEGRO 2a	King 2021[1] Guttman-Yassky 2022[2] Winnette 2022[3]	Ritlecitinib (JAK)	Brepocitinib (JAK) Placebo	Patients with severe AA	AA with ≥50% scalp hair loss
2	ALLEGRO 2b/3	King 2021[4] Mesinkovska 2022[5] Sinclair 2022[6] King 2022[7] Guttman-Yassky 2022[8] Hordinsky 2022[9]	Ritlecitinib (JAK)	Placebo	Patients with severe AA	AA with ≥50% scalp hair loss (including patients with alopecia totalis and alopecia universalis)

AA; Alopecia areata, RCT; Randomised Controlled Trial

A5. CS Appendix D, page 62. Please clarify how many reviewers performed the assessments of study quality for the ALLEGRO 2b/3 and ALLEGRO-LT studies, whether these were checked/adjudicated, and the process for assessing overall study quality based on responses to checklist items for each study.

A single reviewer performed the assessment of study quality for the ALLEGRO 2b/3 and ALLEGRO-LT studies. The assessment was performed by following the suggested quality assessments for parallel group randomised controlled trials (RCTs)

and non-randomised studies provided in the NICE process and methods guide for single technology appraisal and highly specialised technology evaluation.[10]

A6. CS Appendix D, Table 11. Please clarify why items 6(b), 9, 10, 11 and 12 were omitted from the CASP Cohort Study Checklist when it was adapted for use with the ALLEGRO-LT study in Table 11.

For the ALLEGRO-LT study, the quality assessment was performed in line with the NICE process and methods guide for single technology appraisal and highly specialised technology evaluation. The tables used to complete the quality assessment were taken from the suggested table formats provided by NICE in the NICE User guide for when there is more than one study to assess – paragraph 2.5.4, Page 18 (https://www.nice.org.uk/process/pmg24/resources/single-technology-appraisal-and-highly-specialised-technologies-evaluation-user-guide-for-company-evidence-submission-template-pdf-72286715419333) "Single technology appraisal and highly specialised technologies evaluation: User guide for company evidence submission template". Items 6(b), 9, 10, 11 and 12 are omitted from the outline suggested in the NICE process and methods guide,[10] and as such they were not performed for the quality assessment of ALLEGRO-LT.

Clinical effectiveness evidence and statistical analysis

A7. Priority question. CS page 63. Please clarify why the ALLEGRO-2a safety study data is not presented in this submission, given that it would be useful to appraise all available safety data on ritlecitinib, regardless of whether it accompanied any effectiveness data. Please present safety data from the ALLEGRO-2a safety study.

The ALLEGRO-2a safety study (ALLEGRO B7981037) is an ongoing study, data wasn't available to present at the time of submission. An interim CSR is now available and has been provided. At the time of the interim analyses no new safety signal or changes have been observed.

A8. **CS, page 64. Please clarify whether "*within >30 days*" should be "*within 30 days*" or "*>30 days*" in the following passage: "*participants from the Phase 2a***

***& 2b/3 studies were eligible to enrol within >30 days between their last dose in their prior study and first visit in ALLEGRO-LT*”.**

This sentence should read as follows; Moreover, participants from the Phase 2a & 2b/3 studies were eligible to enrol (with >30 days between their last dose in their prior study and first visit in ALLEGRO-LT) regardless of their SALT score.

A9. Priority question. CS, Table 8, pages 65-66. Please clarify why patients with depression and suicide ideation were excluded from ALLEGRO 2b/3. Please provide the number of patients excluded from the study for this reason.

Many clinical trials exclude patients with psychological illnesses.[11] There is a responsibility to protect vulnerable groups within the clinical trial setting as stated in the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) Guideline – General Considerations for Clinical Studies E8(R1); Important principles of ethical conduct of clinical studies and the protection of participants, including special populations, have their origins in the Declaration of Helsinki and should be observed in the conduct of all human clinical investigations. These principles are stated in other ICH guidelines, in particular, ICH E6-Good Clinical Practice. As further described in the E6 guideline, the investigator and sponsor have responsibilities for the protection of study participants together with the Institutional Review Board/Independent Ethics Committee.[12]

There were 5 patients documented as “screen failures” i.e., excluded based on ‘any psychiatric condition including recent or active suicidal ideation or behaviour that meets any of the following criteria’ (as listed in CS.Table 8). Investigators may not have approached patients that did not meet these criteria but there isn’t any empiric evidence of this.

A10. CS, Table 16, page 81. Please clarify which placebo arm is which in rows 8 and 14.

The results refer to the patients switching from placebo to active treatment. In row 8; 22/65 (33.9) is for the placebo arm which switched to 200/50mg ritlecitinib and 12/64 (18.8) is for the placebo arm which switched to 50mg ritlecitinib.

In row 15; 16/65 (24.6) is for the placebo arm which switched to 200/50mg ritlecitinib and 9/64 (14.1) is for the placebo arm which switched to 50mg.

A11. CS, Table 16, page 81. Please provide the baseline SALT score, time since diagnosis and duration of current episode, for the 2 patients in the placebo group who attained a SALT score of <20% and <10% at week 24.

For clarity, the [redacted] participants in the placebo group with SALT ≤10 at week 24 were the same [redacted] participants recorded as achieving SALT ≤20 at week 24.

[redacted] had a baseline SALT [redacted], AA duration since first diagnosis of [redacted] years and duration of current episode was [redacted] years.

The [redacted] had a baseline SALT score of [redacted], AA duration since first diagnosis of [redacted] years and duration of current episode was [redacted] years. For this participant, the duration of current episode is longer than the duration since first diagnosis because the formal diagnosis came after the start of the current episode.

A12. CS, B.2.7: The CS states, "*The differences between each ritlecitinib group and placebo in the proportion of response based on SALT ≤20 at Week 24 were consistent across most pre-specified subgroups (age, BMI, weight, gender, race, region, severity of disease, duration since diagnosis, duration of current episode, prior pharmacological treatment for AA) for all doses.*" The CS then describes exceptions for race and AA severity. However, Appendix E appears to show [redacted] Please clarify whether differences in treatment effectiveness were seen by [redacted].

We have provided sub-group information to indicate that the response to ritlecitinib, based on SALT≤20, is consistent across sub-groups rather than provide evidence of within subgroup differences. We advise caution on overinterpreting the results of these analyses as they are exploratory subgroup analyses for the purpose of hypothesis generation. In this instance, the subgroups are generally small and any variations from the overall effect estimate for SALT≤20 response is likely due to

chance. However, the Company did conduct an internal longitudinal concentration response (LCR) analysis to explore the relationship between these patient characteristics and SALT response.[13] The LCR analysis performed on the raw SALT score included a multivariate analysis where a full LCR model was developed using stepwise covariate modelling approach of forward inclusion/backward elimination. The analysis concluded that AA severity status (i.e., AT/AU at baseline) was the only important covariate in the model, where lower efficacy was expected in the AT/AU population. This is clinically plausible and could be expected since patients are starting from a more severe baseline (SALT95-100) to achieve a response of SALT ≤20. AT/AU population also has relatively longer duration of disease for the current episode compared to remaining AA population. Age, gender, race, region, disease duration since first AA diagnosis, disease duration of the current AA episode, weight, and prior pharmacological treatment were not identified as significant covariates impacting the efficacy of ritlecitinib. Future studies that are adequately powered to detect differences within subgroups is required to draw stronger conclusions

A13. CS, Appendix E. Please provide p-values for the difference between the two groups for all reported subgroups.

P-values were not calculated in the *a priori* analysis of the ALLEGRO-2b/3 subgroup analysis but have been provided in an updated Appendix E.[14] The Company maintains that subgroup analyses are hypothesis generating only, were not adequately powered, and every effort should be made to avoid overinterpretation of the p-values as clinically or statistically significant.

A14. CS, page 115, states "*As it is a small molecule there is no anticipated immunogenicity and so it is unlikely to generate antidrug antibodies which may potentially result in loss of efficacy over time.*" Please clarify if antidrug antibodies were measured in any of the clinical trials and if so please report the findings.

Antidrug antibodies were not measured in any of the clinical trials. It is not standard practice to measure antidrug antibodies for small molecular weight molecules, they are non-immunogenic and as a result are highly unlikely to generate antibodies.

A15. CS, page 118 states that "*treating a patient with ritlecitinib sooner can lead to a better response.*" Please clarify if this statement is consistent with the subgroup analysis in Appendix E for time since diagnosis and duration of current episode.

See Answer A12 – subgroups are exploratory, hypothesis generating and subject to high degrees of uncertainty. We have provided subgroup data in Appendix E to show the treatment effect of ritlecitinib (SALT≤20) is consistent across subgroups not to show within subgroup differences which require additional research.

The company submission is written in the context of patient preference in receiving earlier treatment before their severity increases. For example, it will take longer to get to SALT≤20 from AT/AU (SALT95-100) than SALT50. This is not unreasonable as interactions with clinicians and patient groups support the view that earlier treatment "can lead to a better response".

[redacted]

A16. CS, page 8 states that "*There is currently only one licensed systemic treatment option for adult patients with severe AA*". However, the rest of Section B.1 consistently states that there are no licensed therapies available within the UK. Please explain the discrepancy.

The CS, page 8 states that "*There* ***is currently only one licensed systemic treatment option for adult patients with severe AA, none for adolescents 12 years and over, and none that are currently approved by National institute for Health and Care Excellence (NICE)".*** The intention was to indicate that, since no medicine is currently recommended by NICE, the one available treatment is not publicly available (free at the point of need) within the UK. This is what is meant in the rest of section B.1 with the consistent statement that no licensed therapies are available in the UK. We acknowledge that there may be opportunities to purchase the one currently licensed therapy through private insurance and/or out of pocket expense. In addition, there may be opportunities under individual funding requests (IFR) or individual patient funding requests (IPFR, Wales).

A17. Appendix E: Please clarify whether and why certain criteria such as AT/AU status and prior pharmacological treatment can be considered treatment effect

modifiers. Additionally, please provide the same forest plot as in Figure 1 but after 48 weeks.

See Answer A12 – subgroups are exploratory, hypothesis generating and subject to high degrees of uncertainty. We have provided subgroup data in Appendix E to show the treatment effect (SALT≤20) is consistent across subgroups not to show within subgroup differences which require further research.

Although we continue to advise caution, our further analysis using a longitudinal concentration response model as described in A12 does indicate a potential treatment effect of specifically AT/AU. This is clinically plausible and could be expected since patients are starting from a more severe baseline (SALT95-100) to achieve a response of SALT ≤20. AT/AU population also has relatively longer duration of disease for the current episode compared to remaining AA population. Once again, we advise caution in drawing any conclusions from this. Future research is needed which is adequately powered to detect differences within subgroups of interest.

Appendix E has been updated to include ritlecitinib response of SALT≤20 at 48 weeks. A copy of the forest plot is below [redacted]; the confidence intervals at week 48 are wide due to small sample sizes resulting in high levels of uncertainty. Please note we are unable to provide p-values for this data as the 48-week data is not placebo-controlled. We continue to advise caution in interpreting these results.

[redacted]

n: Number of participants with SALT ≤ 20 in each subgroup; N: Number of participants with valid data in each subgroup.
Participants in the non-AT/AU category had a SALT score of <100% at Baseline (regardless of the category in the AA history CRF), and participants in the AT/AU category had a SALT score of 100% at Baseline (regardless of the category in the AA history CRF).
Confidence Interval is based on normal approximation.
Cases with missing data at Week 48 due to reasons unrelated to COVID-19 are considered as non-response.
Cases with missing data at Week 48 due to COVID-related reasons are excluded.

A18. CS, Section B.2.3.1 and CS Appendix D, Sections D2.2 and D2.3. Please clarify how patients were identified and recruited for the ALLEGRO 2b/3 and ALLEGRO-LT studies.

This question below was discussed with EAG and the below response was deemed sufficient to answer the question.

Selection of Study Population for ALLEGRO 2b/3 (B7981015): To be eligible to enrol in the ALLEGRO 2b/3 (B7981015) study, participants must have had AA with ≥50% hair loss of the scalp (SALT score ≥50) at both Screening and Baseline visits, without evidence of terminal hair regrowth within the previous 6 months and with the current episode of hair loss lasting ≤10 years.

Screening occurred within 35 days prior to the first dose of study intervention, to confirm that participants met selection criteria for the study. Photographs were taken at the Screening Visit to verify eligibility (AA with ≥50% hair loss of the scalp).

The key inclusion and exclusion criteria are listed in the ALLEGRO 2b/3 (B7981015) CSR (page 34).

Selection of study population for ALLEGRO-LT (B7981032): De novo participants and participants originating from Study B7931005 or B7981015, including those with >30 days between the last dose in Study B7931005 or B7981015 and their first visit in Study B7981032, were enrolled.

Prior to the Protocol Amendment 4, de novo adolescent participants were required to have a clinical diagnosis of AA with no other etiology of hair loss other than androgenetic alopecia with ≥25% terminal hair loss of the scalp due to AA at both the screening and Day 1 visits. Although AA criteria were updated in the B7981032 Protocol Amendment 4 for de novo adolescents, all de novo adolescent participants were enrolled prior to approval of Amendment 4 at any sites.

De novo participants were eligible to be included in this study only if all of the inclusion and exclusion criteria applied, unless specified otherwise.

Participants originating from Study B7931005 or B7981015 were eligible to enrol in this study if the Study B7981032 inclusion and exclusion criteria applied, unless specified otherwise, and could not have any AEs meeting the B7981032 safety discontinuation criteria or have discontinued for safety-related AEs.

The key inclusion and exclusion criteria are listed in the ALLEGRO-LT (B7981032) CSR (page 24).

A19. Priority question. CS Appendix D, Figure 4. Please specify how many patients in each arm completed the 24-week placebo-controlled treatment period, and how many discontinued with the 24-week placebo-controlled treatment period and the treatment period in weeks 24-48 when all patients had switched to ritlecitinib. Please also clarify that the number of patients presented as completing the study for each arm refers to the number of patients completing the study at week 48.

In the ALLEGRO 2b/3 (B7981015) study, a total of [redacted] participants were screened and [redacted] participants were randomised to treatment. Of these, [redacted] received treatment ([redacted] participants were not treated) and [redacted] discontinued treatment, see [redacted] (this is Table 8 from ALLEGRO 2b/3 CSR page 59). In addition, the company confirms that the number of patients presented as completing the study for each arm refers to the number of patients completing the study at week 48.

The number of participants discontinuing during the placebo-controlled period was similar across treatment groups, except for 30mg, which had a higher [redacted] discontinuation rate, see [redacted] (this is Table 9 from ALLEGRO 2b/3 CSR page 60):

During the Extension Period (week 25-48), discontinuation rates ranged from [redacted] (50mg) to [redacted] (200/30mg). Overall, [redacted] participants completed treatment, see [redacted] (this is Table 10 from ALLEGRO 2b/3 CSR page 62):

[redacted] 2 [redacted]

	Ritlecitinib 200/50mg QD (N=132)	Ritlecitinib 200/30mg QD (N=130)	Ritlecitinib 50mg QD (N=130)	Ritlecitinib 30mg QD (N=132)	Ritlecitinib 10mg QD (N=63)	Placebo->Ritlecitinib 200/50mg QD (N=65)	Placebo->Ritlecitinib 50mg QD (N=66)	Placebo (N=131)	Total (N=718)
Number (%) of Participants	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Screened: [redacted]									
Screen Failure: [redacted]									
Assigned to Treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treated	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Completed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Not Treated	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysed for Efficacy									
Full Analysis Set (FAS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysed for Safety									
Safety Analysis Set (SAS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Per-Protocol Analysis Set (PPAS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

N = number of participants assigned to a treatment

[redacted] 3 [redacted]

	Ritlecitinib 200/50mg QD (N=132)	Ritlecitinib 200/30mg QD (N=130)	Ritlecitinib 50mg QD (N=130)	Ritlecitinib 30mg QD (N=132)	Ritlecitinib 10mg QD (N=63)	Placebo (N=131)	Total (N=718)
Number (%) of Participants	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Disposition Phase: LOADING (Up to Week 4)							
Discontinued	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lack of Efficacy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

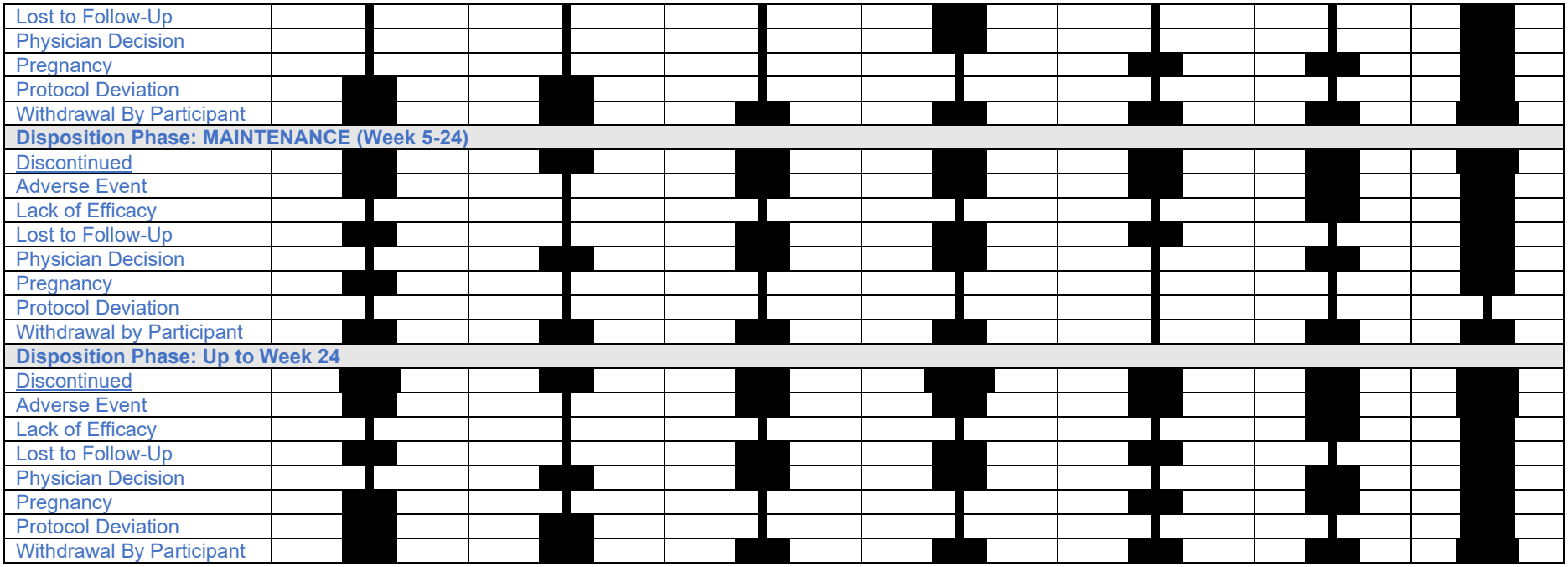


Table 4: Disposition Events Summary up to Week 48 (FAS) (Protocol B7981015)

	Ritlecitinib 200/50mg QD (N=132)	Ritlecitinib 200/30mg QD (N=130)	Ritlecitinib 50mg QD (N=130)	Ritlecitinib 30mg QD (N=132)	Ritlecitinib 10mg QD (N=63)	Placebo->Ritlecitinib 200/50mg QD (N=65)	Placebo->Ritlecitinib 50mg QD (N=66)	Total (N=718)
Number (%) of Participants	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Disposition Phase: Up to Week 24								
Discontinued								
Adverse Event								
Lack of Efficacy								
Lost to Follow-Up								

Non-Compliance With Study Drug								
Physician Decision								
Pregnancy								
Protocol Deviation								
Withdrawal By Participant								
Other								
Disposition Phase EXTENSION (Week 25-48)								
Discontinued								
Adverse Event								
Lack of Efficacy								
Lost to Follow-Up								
Non-Compliance With Study Drug								
Physician Decision								
Pregnancy								
Protocol Deviation								
Withdrawal By Participant								
Other								
Disposition Phase: Overall								
Discontinued								
Adverse Event								
Lack of Efficacy								
Lost to Follow-Up								
Non-Compliance With Study Drug								
Physician Decision								
Pregnancy								
Protocol Deviation								
Withdrawal By Participant								
Other								
Completed								

A20. CS Appendix D, Figure 4. Please clarify the protocol deviations, reasons for physician withdrawal, and what the 'other' reasons consisted of, for each arm of the ALLEGRO 2b/3 study. Please provide reasons for 'withdrawal by participant' in each case (reported by study arm) if known.

There were protocol deviations in the 200/50mg ([redacted] protocol deviation) and 200/30mg ([redacted] protocol deviation) arms of the study. The protocol deviation reported term for the [redacted]. The protocol deviation reported term for the [redacted].

Physician withdrawal occurred in the 200/30mg, 50mg, 30mg, 10mg and Placebo -> 50mg arms. The reported terms are presented:

200/30mg arm:
[redacted]
50mg arm:
[redacted]
30mg arm:
[redacted]
10mg arm:
[redacted]
Placebo -> 50mg arm
[redacted]

'Other' was listed as a reason for withdrawal from the study in the following arms: 200/50mg, 200/30mg, 50mg and 30mg. The reported terms are presented:

200/50mg arm:
[redacted]
200/30mg arm:
[redacted]
50mg arm:
[redacted]
30mg arm:
[redacted]

With drawal by participant occurred in all study arms. The reported terms are presented:

200/50mg arm:
[redacted]

[redacted]

200/30mg arm:

[redacted]

50mg arm:

[redacted]

30mg arm:

[redacted]

10mg arm:

[redacted]

Placebo -> 200/50mg arm

[redacted]

Placebo -> 50mg arm

[redacted]

A21. Please specify the protocol deviations for the ALLEGRO 2b/3, ALLEGRO-LT, ALLEGRO 2a proof of concept and ALLEGRO-2a safety studies.

The protocol deviations in each of the studies are presented in Section 14 of the respective CSRs. The tables and page numbers are referenced as follows:

ALLEGRO 2b/3:

Table 14.1.1.6 in ALLEGRO 2b/3 (B7981015) CSR (page 168).

ALLEGRO-LT:

Table 14.1.1.6 in the ALLEGRO-LT (B7981032) CSR (page 91).

ALLEGRO 2a (proof of concept): Table 14.1.1.6 ALLEGRO 2a (B7931005) CSR (page 422).

ALLEGRO 2a (safety):
Please also refer to Table 14.1.1.6 ALLEGRO 2a (B7981037) CSR (page 84).

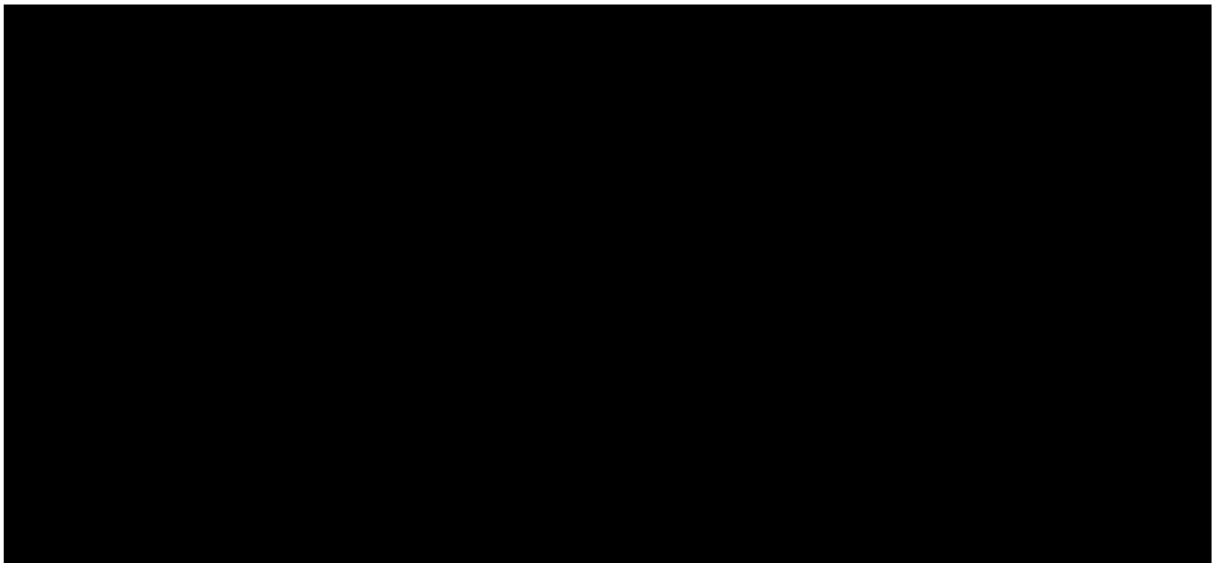
A22. Priority question. CS, Section B.2.3.2. Please specify the data cut-off timepoint for the data from the ALLEGRO-LT that has been presented in the CS.

The data cut-off date was 28th February 2022 for the interim abbreviated CSR for the ALLEGRO-LT study (B7981032). This is consistent with the data used in the submission and the model. The Interim ALLEGRO-LT CSR report can be found attached in folder Updated CSRs February 2023>B7981032-Phase 3 ALLEGRO LT-Interim Clinical Study Report.

A23. Priority question. CS, Section B.2.3.2. Please specify how many patients have rolled over from each of ALLEGRO 2b/3 and ALLEGRO 2a (proof of concept study) into ALLEGRO-LT as of the data cut-off point, and how many were treated with ritlecitinib in those studies. For patients who rolled over from ALLEGRO 2a, please clarify at which point they rolled over and, if after the 24-week double-blind phase, whether any of them received brepocitinib as a study treatment.

The study design schematic for the ALLEGRO-LT study is presented in [redacted]2.

[redacted]2 [redacted]



At the time of this interim CSR, a total of [redacted] participants were screened, of which [redacted] participants were assigned to treatment and [redacted] participants were treated. [redacted] participants were rolled over from ALLEGRO 2b/3 and ALLEGRO 2a (proof of concept) and [redacted] participants were de novo. Of the [redacted] rollover participants; [redacted] rolled over from ALLEGRO 2b/3 and [redacted] rolled over from ALLEGRO 2a (proof of concept). [redacted]5 shows the summary of prior study and treatment group for rollover participants in B7981032.

[redacted]5

	Ritlecitinib 50mg (rollover) (N=603)
Prior Study Treatment Group	**n (%)**
Study B7931005	[redacted]
PF-06651600	[redacted]
PF-06651600/Active Non-responders on PF-06651600	[redacted]
PF-06651600/Active Non-responders on PF-06651600/PF-06700841 (CO)	[redacted]
PF-06651600/Placebo (Withdrawal)	[redacted]
PF-06651600/Placebo (Withdrawal)/PF-06651600 Responders (Retreated)	[redacted]
PF-06700841	[redacted]
PF-06700841/Active Non-responders on PF-06700841	[redacted]
PF-06700841/Active Non-responders on PF-06700841/PF-06651600 (CO)	[redacted]
PF-06700841/Placebo (Withdrawal)	[redacted]
PF-06700841/Placebo (Withdrawal)/PF-06700841 Responders (Retreated)	[redacted]
Placebo	[redacted]
Placebo/Placebo Non-responders on PF-06651600	[redacted]
Placebo/Placebo Non-responders on PF-06651600/PF-06700841 (CO)	[redacted]
Placebo/Placebo Non-responders on PF-06700841	[redacted]
Placebo/Placebo Non-responders on PF-06700841/PF-06651600 (CO)	[redacted]
Study B7981015	[redacted]
Ritlecitinib 200/50 mg QD	[redacted]
Ritlecitinib 200/30 mg QD	[redacted]
Ritlecitinib 50 mg QD	[redacted]
Ritlecitinib 30 mg QD	[redacted]

Ritlecitinib 10 mg QD		[redacted]	
Placebo->Ritlecitinib 200/50 mg QD		[redacted]	
Placebo->Ritlecitinib 50 mg QD		[redacted]	

PF-06651600 = Ritlecitinib, PF-06700841 = Brepocitinib

[redacted] shows the treatments received and the treatment start and end dates for the [redacted] participants who rolled over from B7931005 to B7981032.

[redacted]

B7931005 Planned Treatment 01	B7931005 Treatment 01 Start Date	B7931005 Treatment 01 End Date	B7931005 Planned Treatment 02	B7931005 Treatment 02 Start Date	B7931005 Treatment 02 End Date	B7931005 Planned Treatment 03	B7931005 Treatment 03 Start Date	B7931005 Treatment 03 End Date	B7981032 Planned Treatment	B7981032 Treatment Start Date
[redacted]										

PF-06651600 = Ritlecitinib, PF-06700841 = Brepocitinib

A24. Priority question. CS, Table 12, page 75. Please present patient baseline characteristics for all patients who rolled over from ALLEGRO 2a and ALLEGRO 2b/3 to the ALLRGRO-LT study, and also patient disposition data for all patients in ALLEGRO-LT (separated by cohort - de novo and rolled over, including those who did and did not receive ritlecitinib in the prior studies) from enrolment until data cut-off.

The majority of participants were age ≥18 years [redacted] and the mean (SD) age was [redacted] years at B7981032 Day 1. There were [redacted] adolescents (12-17 years of age), in similar proportions of de novo [redacted] and rollover participants [redacted]. Adolescents were enrolled only in Australia, Chile, China, Japan, Mexico, Russian Federation, Republic of Korea, Taiwan, and the United States of America. There were more female [redacted] than male [redacted] participants. Similar distributions of participants by sex were observed between the de novo and rollover participants.

The median (range) duration since AA diagnosis was [redacted] years and duration since onset of the current AA episode was [redacted] years.

As entry criteria for de novo participants allowed for Baseline SALT scores of 25 or higher (whereas a SALT score of 50 or higher was required for the index studies), rollover participants generally had more extensive disease at Baseline than de novo participants:

- A total of [redacted] participants were classified as having AT/AU, based on a SALT score of 100% at Baseline; more rollover participants had AT/AU [redacted] than de novo [redacted]
- Mean (SD) Baseline SALT score was [redacted] in rollover and [redacted] in de novo participants
- Mean (SD) Baseline SALT score among participants without AT/AU was [redacted] in rollover and [redacted] in de novo participants

Full details for both the de novo and rollover groups are presented in the CSR; Demographic Characteristics – FAS (Table 5 ALLEGRO-LT CSR page 34) and Alopecia Areata History – FAS (Table 14.1.3.4.1 page 149).

A25. Priority question. CS, Figure 23, page 97. Figure 23 shows SALT ≤20 response among the de novo cohort in ALLEGRO-LT. Please provide data on SALT ≤20 response rates to month 24 for the cohort of participants who rolled over from ALLEGRO 2a and ALLEGRO 2b/3.

The proportion of both de novo and rollover participants with response based on SALT ≤20 increased over time. At Month 18, the proportion of participants with SALT ≤20 was [redacted] for the de novo group and [redacted] for the rollover group. Data is presented in [redacted] 7 (this is Table 7 from the ALLEGRO-LT CSR page 38):

[redacted] 7 [redacted]

Analysis Visit		Ritlecitinib 200/50mg (de novo) (N=449)	Ritlecitinib 50mg (roll over) (N=603)	Total (N=1052)
Month 1	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 3	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 6	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 9	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 12	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 15	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 18	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 21	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 24	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]
Month 28	N1	[redacted]	[redacted]	[redacted]
	n (%)	[redacted]	[redacted]	[redacted]
	95% CI[a]	[redacted]	[redacted]	[redacted]

N: Number of participants in FAS population; N1: Number of participants with observed data. n (%): Number of participants achieving Overall SALT score ≤ 20 (percentage based on N1).
a. Confidence Interval for percentages is based on normal approximation.

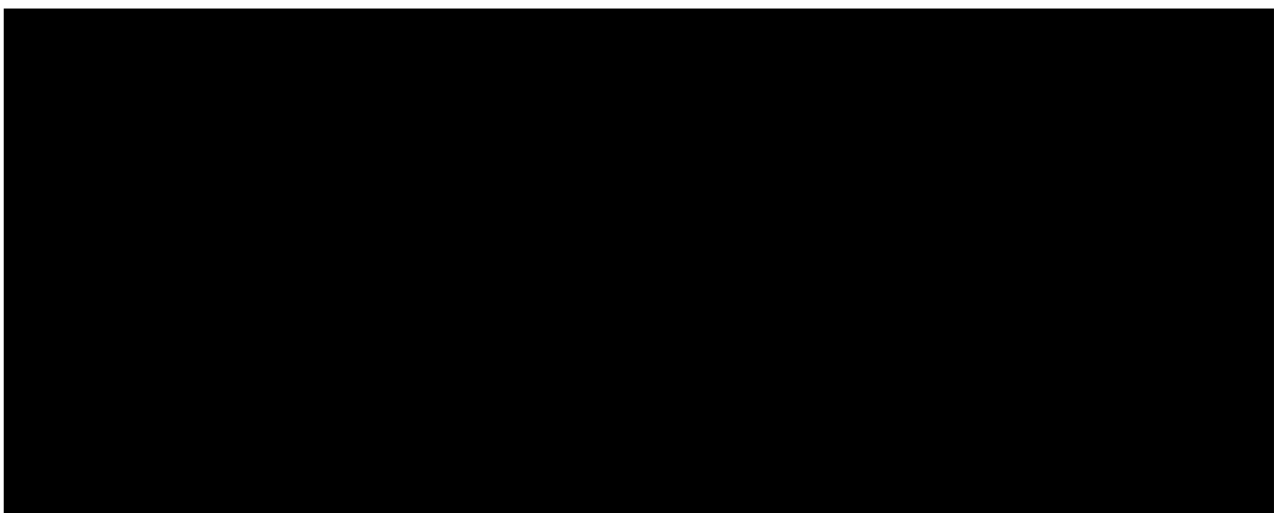
A26. Please comment on the anticipated impact of baseline differences between the ritlecitinib 50 mg arm and pooled placebo arms (reported in CS, Section B.2.3.1.2) on the reported treatment effects (response based on SALT score through week 48)

and health-related quality of life effects (EQ-5D-5L score by dimension, HADS and AAPPO scores by domain). Please provide estimates of treatment effect that are adjusted for baseline covariates.

As the study had balanced treatment arms, it is assumed that balancing baseline covariates would not materially alter the EQ-5D scores or alter any of the conclusions of the analysis of EQ-5D within the ALLEGRO 2b/3 trial data. In particular, the results of the EQ-5D-5L domain scores reported in Document B.3.4.1.1 demonstrate that the vast majority of patients report no problems across all domains assessed in the EQ-5D (Figure 31 and Figure 32 of the Company submission, and repeated in [redacted]3 and [redacted]4 for convenience). The Company has not performed the analysis adjusting for baseline covariates as there is no evidence to suggest any influence of baseline covariates interacting with the HRQoL reported by participants. In addition please see question A12 for more information on potential treatment effect modifiers.

[redacted]3

[redacted]4



A27. CS, Section B.2.10.1.3. Please provide a definition of TRAEs.

Treatment related adverse events (TRAE) are defined as any untoward medical occurrence which emerged or worsened during the treatment period that were causally related to treatment.

Section B: Clarification on cost-effectiveness data

Clinical effectiveness

B1. Priority question. Please clarify how the transition matrices per arm (columns AA to BR in the 'RAW TM data' sheet) have been derived. In particular, please clarify:

a) **for patients transitioning from the phase 2a and phase 2b/3 studies, how do the time points for the transition matrices relate to the follow-up points in ALLEGRO-LT and total time on treatment?**

b) **for the de novo cohort, how do the time points for the transition matrices relate to time points in ALLEGRO-LT? For example, does the data labelled 12 to 15 months for the transition matrices equate to the first 3 months after they started treatment in the de novo cohort, or the first 3 months after they started the maintenance dose?**

c) **how the stopping rules are applied to exclude patients from the transition matrices for de novo patients? For example, given that de novo patients must have a SALT score ≥ 25 at baseline, when is the stopping rule first applied for de novo patients?**

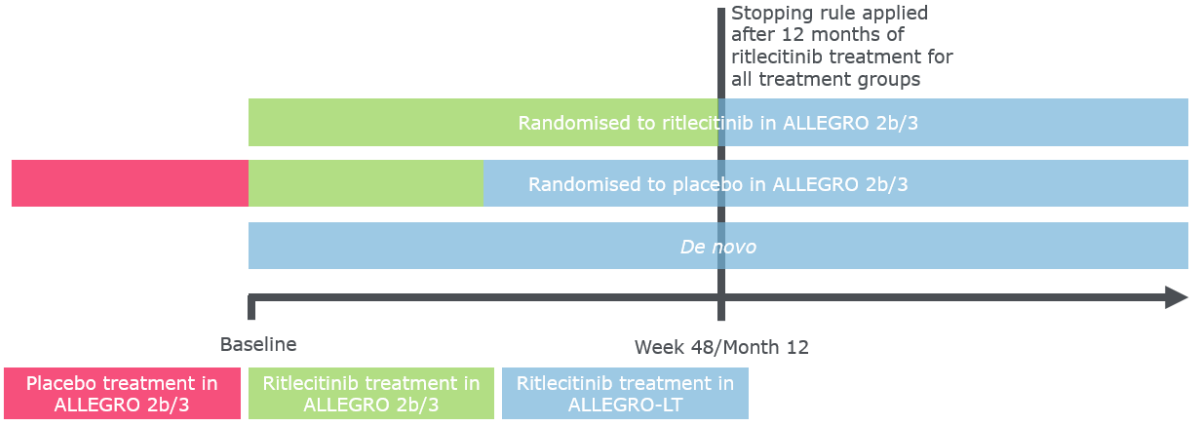
d) **how are the stopping rules applied to patients transitioning from the phase 2a and phase 2b/3 studies? Are patients only excluded from**

contributing to the transition matrices based on their SALT score at one specific time point, and if so what time point is used?

e) **why do only [redacted] de novo patients contribute to the 12-15 month transition matrix (SALT≤ 20 stopping rule for adults and adolescents; cells BD53 to BG56 of the 'RAW TM data' sheet) when there were 449 de novo patients enrolled? The number [redacted] does not appear to correlate with the number with a SALT≤ 20 at any of the timepoints in Figure 23 of the CS.**

A twelve-week block was considered as three months in the ALLEGRO-LT study protocol; therefore, Month 3 and Week 12 are interchangeable, as are Month 6 and Week 24, Month 9 and Week 36 and Month 12 and Week 48. For the ALLEGRO-LT study, patients' baseline was defined as the point of time at which treatment with ritlecitinib began. Therefore, for patients randomised to ritlecitinib in ALLEGRO 2b/3, Month 12 of exposure in the ALLEGRO-LT study corresponds to Week 48 in the ALLEGRO 2b/3 study. In pursuit of explaining how data for different patients are utilised in the cost effectiveness model (CEM) across studies for parts (a) to (d), a graphic is presented in Figure 5 below. Responses to each part of this question are explained below.

Figure 5: Patient flow from the ALLEGRO 2b/3 trial to the ALLEGRO-LT trial



a) Patients transitioning to the ALLEGRO-LT study from the ALLEGRO 2a study were not used to inform the long-term transition matrices or the short-term state membership to Week 48. This is because patients from the ALLEGRO 2a study

were not included in the initial patient distributions as all patients treated with ritlecitinib received a loading dose and there was only 24 weeks of data.[15]

Patients transitioning to the ALLEGRO-LT study from the ALLEGRO 2b/3 study were included to inform patient transitions after 48 weeks of ritlecitinib treatment (i.e., from the end of the ALLEGRO 2b/3 study period). The time points for transition matrices relate to the total time on treatment. Patients in the ALLEGRO 2b/3 study may have received either 24 or 48 weeks of treatment with ritlecitinib prior to enrolment in ALLEGRO-LT according to whether they were randomised to placebo or an active treatment arm. Therefore, health state occupancy at Month 12 of overall exposure to ritlecitinib was determined according to prior treatment:

- For patients randomised to an active treatment arm in ALLEGRO 2b/3, SALT score at Month 12 was defined as SALT score at Week 48 in the ALLEGRO 2b/3 study
- For patients randomised to placebo in ALLEGRO 2b/3, SALT score at Month 12 was defined as SALT score at Month 6 in the ALLEGRO-LT study

b) For the *de novo* cohort of ALLEGRO-LT, the data labelled as 12 to 15 months equates to 12 to 15 months after they started treatment (i.e., 12 to 15 months after entry to the ALLEGRO-LT study).

c) Patients from the ALLEGRO-LT *de novo* cohort were included in the transition matrices only if they would have passed the stopping rule at Week 48 in ALLEGRO 2b/3 (which corresponds to Month 12 in ALLEGRO-LT). This means that only *de novo* patients who had a SALT score ≤ 20 at Month 12 in the ALLEGRO-LT trial were considered for the transitions after Month 12.

d) Patients from the ALLEGRO 2a study were not included in the transition matrices.

Patients from the ALLEGRO 2b/3 study were included in transition counts if they had a SALT score ≤ 20 after 12 months of treatment according to their randomised treatment in ALLEGRO 2b/3:

- For patients randomised to an active treatment arm in ALLEGRO 2b/3, patients were included if they had a SALT score ≤20 at Month 12, which is also defined as SALT score ≤20 at Week 48 in the ALLEGRO 2b/3 study
- For patients randomised to placebo in ALLEGRO 2b/3, patients were included if they had a SALT score at ≤20 at Month 12, which is equivalent to a SALT score ≤20 at Month 6 in the ALLEGRO-LT study

Patients who did not have a SALT score ≤ 20 after 12 months of treatment were excluded from informing the transition counts. No other exclusions were applied to patients informing the inclusion of patients in the transition matrices. However, given that after every 12-week cycle in the model, patients with SALT > 20 discontinue treatment due to the stopping rule, the transition matrices for patients beginning a cycle with a SALT score > 20 do not inform anything in the model while the stopping rule is applied.

To inform the subsequent transition matrices, the SALT scores of patients after each subsequent 3 months (i.e., 12 weeks) were used to derive health state occupancy.

e) In Figure 23, there are 252 of 385 patients with a SALT score ≤20 who may have been considered for the transition matrices. However, the *de novo* group is a larger cohort than that considered in the ALLEGRO-LT transition matrices.

For the ALLEGRO-LT transition matrices, the modified *de novo* group were considered to ensure alignment with the proposed population eligible to receive ritlecitinib. The modified *de novo* group excluded:

- Participants with known androgenetic alopecia
- Participants with a screening or baseline SALT score ≤50

In the modified *de novo* cohort, there were 148 patients with a SALT score ≤20 at Month 12. Of these participants, 2 have missing SALT scores at Month 15. Therefore, 146 patients were included in the transition counts for the movement from Month 12 to Month 15.

B2. Please conduct a scenario analysis in which the transition matrices for long term (week 48+) response to ritlecitinib are calculated using data only from patients who have received the licenced dose at all points during ALLEGRO-LT (i.e., excluding de novo patients who received a loading dose during ALLEGRO-LT). Please also conduct a scenario analysis in which the long-term response is based only on those having the licensed dose at all times before and during ALLEGRO-LT (i.e., patients transitioning from the arm of the phase 2b/3 study that received 50mg from baseline).

The Company has conducted the two scenarios described above. The functionality of this within the model was implemented with a switch (Model Settings, cell E39) altering the table cells J53:M178 inclusive on sheet 'Raw TM data' to deliver either scenario selected. The results are presented in the Appendix of this document.

B3. Please clarify whether the transition matrices for long term response (week 48+) excluded data from the [redacted] patients treated in China, for whom patient-level data were not available (CS, p135). If so, please describe how many were excluded for each of the 7 groups for which transition matrices are provided. If these patients were included for the calculation of the transition matrices, please explain why the data were sufficient for that purpose but not for the purposes of estimating discontinuation.

Patients treated in China were included to inform the transition matrices for long term response. This was possible because analyses conducted to inform the long-term transition matrices were performed by the Pfizer biostatistics team. However, an external vendor performed the discontinuation analysis, with whom it was not possible to share patient level data of Chinese participants due to data protection requirements.

B4. Priority question. Please clarify how the numbers who are steady/improved vs worsened for ritlecitinib ([redacted] and [redacted] respectively) at 24 weeks are calculated

and why they do not total to the number in the ritlecitinib 50mg arm? If this is due to missing data, then please describe how those with missing data at 24 weeks were categorised in terms of the 24-week response criteria and whether they were excluded from all subsequent time points when calculating the distribution given response status at 24 weeks. Furthermore, the number of patients in the 50mg ritlecitinib arm with SALT scores at 24 and 48 weeks appears to be [redacted] based on Table 15. However, the number informing SALT scores at week 24 in the model (cell E45 of the 'Clinical' data sheet) is only [redacted]. Please clarify why these numbers differ.

The data is incomplete due to missingness. To align with the primary analysis of the ALLEGRO 2b/3 study (see Table 15 in the CS), patients missing due to COVID-19 were assumed to be missing at random and were therefore excluded from the analysis. Otherwise, it was assumed that patients with missing data may not have worsened and, therefore, are assumed steady (i.e., SALT score did not increase since baseline) and do not discontinue treatment due to the interim stopping rule.

In the ritlecitinib 50 mg treatment arm, there were 130 patients. At Week 24, [redacted] were missing due to COVID-19, therefore, there was data for only [redacted] patients. Therefore, these [redacted] patients informed the numbers for patients who are steady/improved and worsened at Week 24. Of these [redacted] patients, [redacted] were missing due to reasons other than COVID-19, and so were assumed steady. These [redacted] patients also inform the SALT scores at Week 24 in the model. In Table 15 of the submission, the sample size for the estimated response at Week 24 is not reported; it is [redacted] as for the steady/improved versus worsened analysis.

B5. Excel model, 'Clinical' sheet: According to cells E54 and E63, there appear to be [redacted] and [redacted] patients with SALT scores available at 34 and 48 weeks suggesting no further decline in data availability after 24 weeks. [redacted] of these pass the stopping rule at 24 weeks. However, the number of these patients with SALT scores available at 34 and 48 weeks in cells E120 and E129 is lower. Please explain why this is the

case and if this is due to missing data, please clarify what is assumed for those with missing data.

The reason for discrepancies in these patient counts is due to patients with COVID-19 who are excluded from response calculations in the primary analysis (i.e., where patients who are missing due to COVID-19 are not imputed).

For the data relating to the final stopping rule only:

- At Week 24, there are [redacted] missing due to COVID-19. There are [redacted] who are included in the response analysis and the stable/improved versus worsened analysis.
- At Week 34, there are [redacted] missing due to COVID-19. There are [redacted] who are included in the response analysis and the stable/improved versus worsened analysis.
- At Week 48, there are [redacted] missing due to COVID-19. There are [redacted] who are included in the response analysis and the stable/improved versus worsened analysis.

For the data relating to the interim and final stopping rule:

- As for the final stopping rule only, [redacted] patients are included. Of these, [redacted] were steady/improved and are carried forward for the response dependent on the interim stopping rule.
- Of the [redacted] patients steady/improved at Week 24, [redacted] were missing due to COVID-19 at Week 34. Therefore, 109 patients were included in the response analysis.
- Of the [redacted] patients steady/improved at Week 24, [redacted] was missing due to COVID-19 at Week 48. Therefore, 110 patients were included in the response analysis.

Data for Week 34 and 48 are independent of each other; if data for an individual is not reported at Week 34, they may still be included in the Week 48 analysis if data is available.

B6. CS, page 126: Please explain why the number of patients in Figure 28 does not match with the FAS randomised to ritlecitinib 50 mg. Also clarify why the number having '<0%' improvement does not correspond to the number identified as having 'worsened' for the purposes of the stopping rule.

The data in Figure 28, Document B includes all patients on treatments with 50 mg or 30 mg ritlecitinib once daily during the ALLEGRO 2b/3 trial and, therefore, the number of patients having '<0%' improvement does not correspond to the number identified as having 'worsened' for the purposes of the stopping rule. The number of patients treated with 50 mg ritlecitinib once daily who worsened at Week 24 was [redacted], aligned with the number of patients identified as having 'worsened' for the purposes of the stopping rule.

The number of patients in Figure 28, Document B also does not match the number of patients in the full analysis set (FAS) as it only considers patients treated with 50 mg or 30 mg ritlecitinib once daily and missing data was not included in the analysis.

An updated version of Figure 28, Document B - including patients treated with 50 mg ritlecitinib once daily only is below provided in [redacted]

[redacted]

Abbreviations: SALT, Severity of Alopecia Tool

B7. Section B.3.3.6 states that, when calculating time to discontinuation, "*Patients whose SALT score increased above SALT 20 at any time after Week 48 were excluded as these patients were assumed to automatically discontinue treatment; their removal ensured discontinuation was not double-counted.*" Please clarify whether 'excluded' refers to the censoring of data at the time their SALT score exceeded 20 rather than exclusion from the data set at all timepoints. Please confirm if applying or removing these exclusions was the change implemented in the scenario described as, "*Discontinue patients based on SALT score after 48 weeks*," in Table 61.

The Company confirms that 'excluded' refers to the exclusion from the data set at all timepoints. If a patient had a SALT score >20 after Week 48, they would then be assumed to discontinue ritlecitinib treatment in the CEM. Therefore, exclusion of these patients does not exclude any discontinuation events. If a patient had a SALT score >20, this is recorded whilst they are still on treatment and prior to any discontinuation event; should this patient go on to discontinue for reasons other than SALT score after this high SALT score, it would not be considered in the model since they would have already been assumed to discontinue due to loss of response. Since SALT score was reported whilst patients are enrolled in the study, recordings of SALT score are not taken after patients discontinue treatment and leave the study. As such, events of patients who discontinue prior to the recording of a SALT score >20 are not possible. Therefore, it is not possible to have excluded any discontinuation events that occur prior to a high SALT score that would be considered loss of response.

This is a conservative assumption as excluding the data of patients with SALT score > 20 at any timepoint after Week 48 means that the sample of data is reduced, thus slightly increasing the proportion of patients who discontinue treatment for reasons other than loss of response in comparison to the proportion of patients who would be estimated to discontinue treatment for reasons other than response if the data were censored as suggested. Overestimating the rate of discontinuation due to reasons other than response is conservative because it increases the rate of patients discontinuing with a SALT score ≤ 20, who will then regress to the SALT score 50-100 health state, associated with a lower utility value.

On the other hand, a lower discontinuation rate due to reasons other than a loss of response would mean more patients remain on treatment who have not lost response and, therefore, remain in the SALT score ≤ 20 health states, associated with a higher utility value. The mechanisms for on-treatment patients to discontinue are described in Table 33 of the Company submission. This table is copied in Table 8 below with further clarification.

Table 8: Mechanisms for on-treatment patients to discontinue

Mechanism	Approach	Patients considered
Initial stopping rule	Patients who show a worsening in their SALT score between baseline and Week 24 will discontinue.	All who initiated treatment with ritlecitinib 50 mg in ALLEGRO 2b/3
Final stopping rule	Patients who do not achieve a SALT ≤20 at Week 48 will discontinue treatment.	All who passed the initial stopping rule who were treated with ritlecitinib 50 mg in ALLEGRO 2b/3
Loss of response	If, after Week 48, patients transition to a health state with a SALT >20, they will immediately discontinue treatment due to loss of response.	Patients considered for the transition matrices who would have passed the final stopping rule after 12 months of treatment with ritlecitinib
Discontinuation for other reasons	After Week 48, patients may discontinue treatment with a SALT score of 0-20 for reasons other than loss of response.	Patients considered for the transition matrices who satisfy the following: • who would have passed the final stopping rule after Week 48/12 months of treatment with ritlecitinib • whose SALT score does not increase >20 after Week 48/12 months of treatment with ritlecitinib

The Company confirms that one of the scenarios provided in the original submission, "Discontinue patients based on SALT score after 48 weeks", reintroduces these patients for the analysis of discontinuation when the setting is toggled to no. As part of this scenario, patients whose SALT score increases to above 20 at any time after Week 48 would do not discontinue either. Therefore, patients are not excluded from the discontinuation analysis for any reason other than not passing the final stopping rule at Week 48.

B8. Section B.3.2.3.3: Please describe the justification for the trajectory of hair loss assumed upon discontinuation with particular reference to any studies where SALT

scores were measured following treatment discontinuation (for example for active responders during the withdrawal and treatment phase of ALLEGRO 2a reported by Peeva 2022. https://doi.org/10.1016/j.jaad.2021.12.008)

The available evidence from literature regarding the trajectory of hair loss after discontinuation is limited. The study highlighted by the EAG is not suitable for use, as it reports the time until patients lose 30% of hair regrown after achieving a 30% improvement in SALT score. Whilst patients must have had ≥50% hair loss at baseline, a 30% improvement in a SALT score of 50 corresponds to a SALT score of 35; a SALT score of 35 is not aligned with the stopping rule of SALT ≤20. Conversely, for patients with a higher SALT score at baseline, a response would be associated with a higher SALT score (for example, with a baseline SALT score of 90, a 30% improvement would equate to a SALT score of 63). Therefore, the definition of response in this study is not aligned with the stopping rule considered for ritlecitinib in this appraisal.[16]

There are no randomised controlled trials assessing the trajectory of hair loss after discontinuation of treatment with a JAK inhibitor, and the uncontrolled studies that have assessed the trajectory of hair loss have consisted of small sample sizes of patients and reported large ranges in time until disease relapse after treatment discontinuation.[16] For example, an open-label trial for patients with AA treated with tofacitinib assessed the durability of response after discontinuation in 20 patients and found that the median time until disease relapse after drug cessation was 8.5 weeks with a range of approximately two weeks to 25 weeks.[17] An open-label trial of patients with AA treated with oral ruxolitinib assessed the durability of response in nine patients, three of which had marked hair loss twelve weeks after treatment discontinuation and six of which reported increased shedding but without major hair loss.[18] A study assessing the durability of response of five patients to tofacitinib found that hair loss recurred after a median of two months of discontinuing tofacitinib with a range of two to seven months.[19] Finally, a study assessing the durability of response to tofacitinib in seven patients found that six patients exhibited variable hair shedding after completion of the study treatment, with two patients showing initial signs of shedding approximately 34 weeks after end of treatment, and four patients showing initial signs of shedding approximately 8 weeks after end of treatment.[20] Moreover, use of the identified studies is not possible given that they do not align with the

population eligible for treatment with ritlecitinib (i.e., one study included patients with initial hair loss covering <50% of the scalp) and the definition of response (i.e., no study defined response as attainment of a SALT score ≤20) differs across studies.

Given that the limited evidence suggests that while most patients do lose their hair after treatment discontinuation, there is some durability in response and disease relapse is not immediate, it was assumed that patients continue to have the same SALT score for one cycle after discontinuing treatment with ritlecitinib before progressively regressing to SALT 50-100. As the evidence to inform this assumption is limited, the assumption was validated by dermatologists with a specialist interest in hair disorders.[21] All of the dermatologists agreed with the assumption (3/3 clinicians, 100%).[21]

B9. Please clarify why the labels for the rows in the transition matrices in column B of the 'RAW TM data' sheet refer to SALT categories at 12 months whereas in the 'Clinical' sheet the rows are labelled differently for each subsequent transition matrix (12, 15, 18 and 21 months). If the transition matrices are being applied to the health state occupation at the previous cycle then they should relate to transitions from the time of the previous cycle. For example, the transition matrix applied to health state occupancy at 15 months to provide health state occupancy at 18 months should be one derived using the transitions observed between 15 months and 18 months not the transitions observed between 12 months and 18 months. Please clarify if these are mislabelled in the 'RAW TM data' sheet.

The Company confirms that these are mislabelled in the 'RAW TM data' sheet and the labels have been corrected in the updated version of the model.

Adverse events

B10. Priority question. Please explain why serious infections were not included in the modelling. Rare adverse events that are either high morbidity or high cost (e.g., sepsis, appendicitis, pulmonary embolism) should be

included as they can significantly impact the ICER. Please include these if providing an updated base case model.

The adverse events considered in the analysis were treatment emergent adverse events (TEAEs) occurring in greater than 5% of patients in either the placebo or ritlecitinib 50 mb arm of the FAS population in the ALLEGRO 2b/3 trial. The serious adverse events (SAEs) reported in the ALLEGRO 2b/3 trial were low; in the 50 mg treatment arm, there were [redacted] events amongst 130 patients, only one of which was considered related to treatment (0.8% of patients treated with ritlecitinib 50 mg experienced a SAE deemed related to treatment). In the placebo arm, there were [redacted] events amongst 131 patients, none of which were considered related to treatment. Based on this, the Company assumes that SAEs will have no impact on the ICER of ritlecitinib relative to best supportive care (BSC) and have not included them in the analysis.

B11. Please clarify why NHS reference costs have been used as unit costs for adverse events. Is the company expecting that the adverse events listed in Table 54, which are treatment emergent adverse events (TEAE) and not serious adverse events (SAEs) will be managed by admission rather than within primary care? Also, please provide details of the specific HRG codes that have been applied sufficient for the EAG to be able to cross check these.

The Company is expecting the TEAEs to be managed by admission. National Health Service (NHS) reference costs for regular day or night admissions were used as unit costs for TEAEs because they provide a standardised and reliable measure of the resources required to manage these events.[22] The TEAEs in B.3.5.3., Document B (Page 184) may require additional medical interventions or hospital stays, and the costs of these interventions can be estimated using NHS reference costs.

This is a conservative assumption, as if the TEAEs were managed in primary care, the associated costs would be lower. As ritlecitinib is associated with more TEAEs than BSC, this increases the cost more for ritlecitinib than it does BSC.

The adverse event costs are presented in Table 9.[22]

Table 9: Adverse event costs

Adverse event	Cost, £
Acne	627.40
Diarrhoea	231.65
Folliculitis	627.40
Headache	402.39
Nasopharyngitis	230.22
Rash	627.40
Upper respiratory tract infection	230.22
Urticaria	627.40

In the National Schedule of NHS Costs, the unit cost for acne was derived from the 'RP' sheet (currency code JD07K, cell D492).

In the National Schedule of NHS Costs, the unit cost for diarrhoea was derived from the 'RP' sheet (currency code FD10M, cell D275).

In the National Schedule of NHS Costs, the unit cost for folliculitis was derived from the 'RP' sheet (currency code JD07K, cell D492).

In the National Schedule of NHS Costs, the unit cost for a headache was derived from the 'RP' sheet (currency code AA31E, cell D43).

In the National Schedule of NHS Costs, the unit cost for nasopharyngitis was derived from the 'RP' sheet (currency code WH07G, cell D930).

In the National Schedule of NHS Costs, the unit cost for a rash was derived from the 'RP' sheet (currency code JD07K, cell D492).

In the National Schedule of NHS Costs, the unit cost for an upper respiratory tract infection was derived from the 'RP' sheet (currency code WH07G, cell D930).

In the National Schedule of NHS Costs, the unit cost for urticaria was derived from the 'RP' sheet (currency code JD07K, cell D492).

For convenience, an extract of the NHS reference costs excel file has been provided.

A scenario where it is assumed that each TEAE is managed in primary care by a GP appointment is provided. A GP appointment is assumed to cost £39.23, taken from the PSSRU.[23,24]

B12. SmPC states that the median duration of urticaria as an adverse event was 7 days. Please adjust your AE in the model to use 7 days instead of 28 days. Please also reconsider whether an assumed duration of 28 days has clinical face validity for the remaining AEs.

The Company has updated the median duration of urticaria as an adverse event to [redacted] days in the model to align with the median duration of urticaria in the summary of product characteristics (SmPC; Appendix C).

It was conservatively assumed that the duration of each disutility due to adverse events is 28 days. Dermatologists with a specialist interest in hair disorders verified this assumption (3/3 clinicians, 100%).[21] For example, one dermatologist stated that "[redacted]".[21] Moreover, as the frequency of TEAEs is more with ritlecitinib than BSC, this assumption is conservative.

B13. Model, Adverse Events, cells H35:I42: Please clarify why the alpha and beta parameters for the used Beta distribution do not sum up to the sample size experiencing and not experiencing the AE respectively. Please correct within an updated PSA.

The alpha and beta parameters used for the Beta distribution in the model are not based upon the sample size experiencing and not experiencing adverse events (AEs). The alpha and beta parameters are based upon a formula utilising the mean and standard error. These parameters are then used to run probabilistic sensitivity analysis (PSA) results.

In line with the suggestion, we have updated the parameters for alpha and beta for AEs in the way described. The updated PSA is provided in the Appendix.

B14. CS, page 134: Please clarify why the risk ratio for AEs (BSC versus ritlecitinib) was calculated from AE rates in general, and not TEAEs with an incidence rate of

≥5% at week 24. Also, it appears that no uncertainty has been included within the PSA for this risk ratio. Please correct within an updated PSA.

The values have been updated to reflect the TEAEs occurring in ≥5% of patients for both ritlecitinib and BSC, obtained from the clinical study report (CSR; Table 14.3.1.2.1.2). Other TEAEs which occurred in more ≥5% of patients at any ritlecitinib dose at Week 48 (folliculitis, rash and urticaria) were also considered to align with the adverse events included in the model.

[redacted]

	[redacted]	[redacted]
Acne	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Folliculitis	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]
Urticaria	[redacted]	[redacted]
Total number of events	[redacted]	[redacted]

Abbreviations: TEAE, treatment-emergent adverse event

The updated risk ratio value is [redacted] to reflect the relative likelihood of adverse events which occur in ≥5% of patients with placebo relative to ritlecitinib 50 mg at 24 weeks. Uncertainty for the risk ratio has now been included in the Company's updated model (Adverse Event, cells H48:J48) and is included in the PSA.

Quality of life

B15. The caregiver disutilities are the same in the main TTO analysis and the TTO sensitivity analysis (SA) in appendix H, Table 17, but they differ in the 'Utility Weights

Calcs' sheet (cell D44 and D57) of the model. Please clarify if this is a typo in Appendix H, Table 17 or an error in the model.

The Company confirms that the utility weights in the sensitivity analysis alongside the full sample analysis in Appendix H, Table 17 are correct. The 'Utility Weights Calcs' sheet (cell D44 and D57) has been corrected in the updated version of the model to align with Appendix H.

B16. Please provide EQ-5D scores averaged across all ritlecitinib arms, and across all groups (placebo and ritlecitinib arms) as additional columns in Table 39. This will reduce the uncertainty associated with small numbers and allow estimates of EQ-5D according to SALT score. Please consider providing a regression to estimate EQ-5D according to SALT scores or explain why this is not appropriate.

A summary of the EQ-5D scores averaged across all ritlecitinib arms and across all groups is provided below in [redacted]. Please note, both all treatment arms and all ritlecitinib treatment arms, include 10mg results. Ritlecitinib 10 mg QD was included in the study exclusively to support the estimation of the exposure response.

[redacted]

Cohort	[redacted]	[redacted]
Week 24 SALT score ≤20		
Participants with SALT ≤20 Response, n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
Median (Min, Max)	[redacted]	[redacted]
Week 24 SALT score >20		
Participants with SALT >20 Response, n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
Median (Min, Max)	[redacted]	[redacted]
Week 48 SALT score ≤20		
Participants with SALT ≤20 Response, n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
Median (Min, Max)	[redacted]	[redacted]
Week 48 SALT score >20		
Participants with SALT ≤20 Response, n	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
Median (Min, Max)	[redacted]	[redacted]

Abbreviations: CI, confidence interval; Max, maximum; Min, minimum; mg, milligrams; SALT, Severity of Alopecia Tool; SD, standard deviation
**Includes 10mg arm results. Ritlecitinib 10 mg QD was included in the study exclusively to support the estimation of the exposure response.*

In line with the results from the individual treatment arms, there remains limited differences in EQ-5D score according to whether patients had achieved a SALT score ≤20 or not. In line with the reasoning given in the submission, the EQ-5D from the ALLEGRO 2b/3 study is not appropriate to parametrise the health-related quality of life (HRQoL) of individuals with AA due to the inherent insensitivities caused by a lack of content validity for assessing the HRQoL in patients with AA and a high ceiling effect. Therefore, EQ-5D tariff scores according to SALT score are not considered appropriate to apply to the model.

B17. Please clarify whether any quality of life measures were included in the ALLEGRO-LT study. For any quality of life measures included in ALLEGRO-LT, please summarise outcomes for these measures using methods similar to those used to present quality of life outcomes from ALLEGRO 2b/3.

All quality-of-life measures included within the ALLEGRO 2b/3 were also included in the ALLEGRO-LT trial and can be found in the interim CSR in Table 1 page 19, Study Objectives and Endpoints. However, as the study is ongoing, interim results are only available for primary and secondary endpoints. Among the quality-of-life measures, only interim HADS, PGIC and AAPPO results (descriptive summary only) are available at this time (ALLEGRO-LT interim CSR page 46-52). We have included the AAPPO results below.

AAPPO – Improvement in Hair Loss in Scalp, Eyebrows, Eyelashes, and Body, Items 1-4 (ALLEGRO-LT interim CSR Tables page 154-173)

Patient-reported improvement (defined as achieving a score of 0 [no hair loss] or 1 [little hair loss]) on each of the AAPPO hair loss items 1-4 (scalp, eyebrows, eyelash, body hair) was assessed in participants with a Baseline score of 2-4 (indicating moderate-complete hair loss)

The proportion of both de novo and rollover participants reporting an improvement on each of these 4 AAPPO items increased from Month 1 to Month 18.

AAPPO Emotional Symptoms and Activity Limitations – Change from Baseline
(ALLEGRO-LT interim CSR Tables page 174-181)

In the AAPPO, the Emotional Symptoms sub score is defined as mean of items 5-8, and the Activity Limitations sub score is defined as mean of items 9-11.

In de novo participants

- Mean baseline scores were consistent with Emotional Symptoms that occurred “rarely” or “sometimes” and Activity Limitations that were “not at all” or “a little.” At Month 18, the mean (SD) changes from baseline in Emotional Symptoms scores and Activity Limitations scores were [redacted] and [redacted], respectively.

In rollover participants:

- Mean baseline scores were consistent with Emotional Symptoms that occurred “rarely” or “sometimes” and Activity Limitations that were “not at all” or “a little.” At Month 18, the mean (SD) changes from baseline in Emotional Symptoms scores and Activity Limitations scores were [redacted] and [redacted], respectively.

B18. On page 153, the CS states, “*In the ALLEGRO 2b/3 study, improvement in AAPPO score was measured for patients with AAPPO score ≥2 at baseline for each of the hair loss items.*” Was AAPPO measured in all patients? If so, then please also report mean AAPPO scores for all randomised patients. Please also provide figures equivalent to Figure 37 for outcome of proportion worsening for those with baseline scores <2.

AAPPO scores were measured in all patients. Please see [redacted] below reporting mean AAPPO scores for all randomised patients. We showed the proportion of patients improvement from baseline in AAPPO hair loss items amongst participants with a score ≥2. The improvement of AAPPO was the trend whereas Figure 37 shows the change in baseline.

[redacted]12[redacted]

Baseline AAPPO	Scalp	Eyebrows	Eyelashes	Body
0				
1				
2				
3				
4				

Worsening graphs for patients with baseline AAPPO of 0 or 1.

Please also find below in [redacted] to [redacted] equivalent to Figure 37 in CS for the outcome of proportion worsening for those with baseline scores <2. Note that because we're starting with smaller numbers, error bars are much larger than the improvement graphs. For example, scalp has very small numbers and, thus, 95% CI's are quite large. Therefore, the axis needed to be adjusted to fit the error bars on some of the graphs. Error bars and 95% CI's calculated using normal approximation

[redacted]

8

9

[redacted] 10

B19. CS, page 158 states "*The effect size estimate suggests that this is a large effect (d=0.79)*" Please clarify what the d-statistic here refers to. Does it relate to the differences in AAPPO emotional scores for SALT <20 versus SALT >20? If so, then what time point does it relate to? To avoid selective reporting, please also provide these statistics for other dimensions of AAPPO and for both 24 and 48 week timepoints.

The above-mentioned d-statistic [redacted] refers to Cohen's D effect size,[25] comparing SALT 0-10 versus SALT 100 groups for the AAPPO Emotional symptoms score. Data across all time points were pooled for this analysis which provided a large number of observations.[26]

The d-statistic is a measure of effect size. This is a statistical method for understanding the magnitude of differences between sub-groups in terms of the different PRO measures. Because sub-groups are fixed (defined by SALT) the d-

statistic is showing how sensitive each measure is to differences between groups. Essentially it is assessing the differences between SALT 0-10 and SALT 100 groups.

Data across all available timepoints was pooled for this analysis (including at 24 & 48 weeks). We do not believe it's appropriate to present the data in terms of weeks 24 and 48 separately.

B20. CS page 158 states, "*The AAPPO Activity limitations score was associated with a moderate effect size (d=0.48) between SALT 0-10 and SALT 100.*" Please clarify what time point this relates to. To avoid selective reporting, please also provide equivalent d-statistics for both 24 week and 48 week timepoints and for the other dimensions of AAPPO.

The AAPPO Activity limitations score was associated with a moderate effect size [redacted] between all groups by SALT score (SALT 0-10; SALT 11-20; SALT 21-49; SALT 50-74; SALT 75-99; SALT 100; SALT 50-100). Data across all time points were pooled for this analysis (i.e., from Day 1 to week 48) which provided a large number of observations.[26] As described for the previous response, the d-statistic incorporates the differences observed across all time points. We don't think this is selective reporting. This has been clarified in the latest version of the manuscript.[27]

B21. Priority question. Appendix H, p73 to 75. The caregiver vignette is explicitly described as a caregiver of an adolescent patient with AA (and a SALT score of 50 to 100). Please clarify whether the caregiver disutility derived from this vignette is applied to both adult and adolescent patients and if so what is the rationale for assuming the same caregiver disutility for adult patients. Please also provide a scenario in which the caregiver disutility is applied only for patients aged under 18.

The Company confirms that the caregiver disutility was applied to both adults and adolescents in the model. Although the caregiver disutility derived from the vignette study was calculated for adolescents, alopecia can negatively impact the caregivers for both adolescents and adults, as discussed in Section B.1.3.2.4, Document B. Dermatologists with a specialist interest in hair disorders verified the applicability of caregiver disutility for the entire population (3/3 clinicians, 100%).[21]

A scenario has been included in the updated version of the model in which the caregiver disutility is only applied to adolescents, as requested. This has been completed by calculating the duration of caregiver disutility for adolescents to identify the total quality-adjusted life years (QALYs) lost due and multiplied by the percentage of patients in the model who are adolescent in cells P23:W24 on sheets 'Ritlecitinib 50mg Final Calcs', 'Ritlecitinib 50mg I+F Calcs' and 'BSC Calcs'.

B22. Appendix H: Why were participants in the interviews to inform the vignette required to "*have experience of any of the following treatments for AA: contact immunotherapy, oral corticosteroids, oral immunosuppressant therapies, or express an interest in receiving systemic treatment*". Describe if there may be bias introduced from the potential to exclude people who are less likely to actively seek treatment and who might have different views of the impact of their AA on their life.

The study specifically selected people who reported an interest in receiving systemic therapy in order to make the data most applicable or representative for people who would consider using an immune modifying therapy to treat their alopecia. At the study outset, we assumed that some people with AA would be willing to live with the condition and would not be willing to take any systemic treatment. We also assumed that a proportion of people with AA find the condition to have such a significant effect on their HRQL that they would be willing to try different treatments including systemic treatments. Therefore, if we wish to understand the burden of AA at different levels of SALT score we think it is most appropriate to explore that burden in people who would be interested in taking a treatment for it. If we included people who would never take a systemic treatment then the results would be unfairly biased against those who are most likely to pursue systemic treatment.

B23. Appendix H, section H2.2.4 and H2.2.5: How many patients and carers were interviewed in the first and second rounds of interviews during the vignette development?

In round 1, 3 adults, 3 adolescents and 5 carers were interviewed. In round 2, 5 different adults and 5 different caregivers were interviewed.

B24. Appendix H, Table 14: Only 5% of the vignette TTO valuation sample (5/120) were reported as being from England and only 39% from the UK overall. Given that

one of the recruitment criteria was being a UK resident (section H2.2.7.1), please clarify if these data in Table 14 are typos or whether non-UK residents were included in the sample. If the latter please summarise their nationalities.

Apologies, these were typos. [redacted] of the sample were from England, followed by [redacted] from Scotland and [redacted] from Wales. We have provided an updated report and a copy of the table can be found in [redacted] below. Further information on the Vignette study methodology, targeted literature search and vignettes can be found in Appendix H.

Table 13: Sample demographic characteristics

Characteristic	UK sample (N = 120)	UK 2011 census data[1]
Age Mean (SD)	[redacted]	[redacted]
Range	[redacted]	
Gender		
Male	[redacted]	[redacted]
Female	[redacted]	[redacted]
Ethnicity		
White	[redacted]	[redacted]
Mixed or multiple ethnicity	[redacted]	
Asian or Asian British	[redacted]	
Black, African, Caribbean or Black British	[redacted]	
Other ethnic group	[redacted]	
Prefer not to answer	[redacted]	
Country		
England	[redacted]	
Wales	[redacted]	
Scotland	[redacted]	
Northern Ireland	[redacted]	
Employment		
Employed full-time	[redacted]	
Employed part-time	[redacted]	
Self employed	[redacted]	
Stay at home or full-time carer	[redacted]	
Retired	[redacted]	
Seeking work/unemployed	[redacted]	
Long term sick leave	[redacted]	

Student	[redacted]	
Other	[redacted]	
Lives with long-term condition		
Yes	[redacted]	
No	[redacted]	
Prefer not to answer	[redacted]	

B25. Appendix H, Table 14: The majority of the vignette valuation sample were female. Our discussions with clinical experts suggest that preferences for avoiding states of hair loss may be dependent on gender. Please clarify if any subgroup analysis was conducted to explore whether preferences differed by gender.

Apologies, this was a typo. The sample was in fact [redacted] male and [redacted] female (see updated table above).

B26. The EAG note that the disutility for Grade 3/4 alopecia as reported in oncology appraisals (e.g., TA627 committee papers, Table 44 of Celgene submission, provides a values of -0.045) is small compared to the difference between SALT<10 and SALT >=50 based on the vignettes [redacted] Please comment on the face validity of the utility scores generated by the vignette study for AA in comparison to the estimates applied for alopecia in other NICE TAs and those identified from the literature (CS Table 45).
[https://www.nice.org.uk/guidance/ta627/documents/committee-papers]

Utility values estimated in the vignette study for AA are within a similar range to those reported in other skin conditions, which we would consider a more appropriate comparison, such as psoriasis, hidradenitis suppurativa, atopic dermatitis and venous leg ulcers,[28] with similar levels of differentiation by severity / treatment response (e.g., most vs. least severe hidradenitis suppurativa: 0.35 vs. 0.80 for the least severe).[29,30] This is consistent with the patient lived experiences and significant burden of AA as described by patient advisory groups and dermatologists with a special interest in hair disorders.

With regards to TA627, table 44, there is a marked difference between the disutility reported for hair loss in the lymphoma submission and for alopecia areata found in the vignette performed to support this appraisal. The experience of hair loss as a

side effect of cancer treatment is a very different context. We would argue that patients in this example are coming from an entirely different starting point i.e. their HRQoL is already very low (floor effect). TA627 cites a value for hair loss from a vignette study which examined the impact of advanced/metastatic NSCLC and associated treatment toxicities.[31] The reference study proposes two vignettes associated hair loss as a result of treatment toxicities, those stable and those responding with NSCLC. The vignette describes a “life threatening illness” in which “you worry about dying and how your loved ones will cope". The NSCLC patients have experienced hair loss in the context of a very serious disease which will reduce their life expectancy as a result of a treatment. This has the potential to extend their life versus the impact associated with hair loss as a result of chemotherapy. Furthermore, hair loss associated with chemotherapy is an expected and reversible side effect whereas hair loss associated with AA is a life-long chronic condition.

This comparison is inconsistent with the etiology of AA and with our own research findings. Specifically, the qualitative research on the conceptual model of alopecia areata as described by patient advisory groups, as well as the learnings from AA patients through the development of our own disease specific vignettes. We would argue that people would be far more willing to experience alopecia from a treatment that may save their life. We expect it to be highly unlikely that the impact of hair loss on HRQoL was as great as the impact of the disease-related symptoms of NSCLC for patients with NSCLC. As patients with NSCLC have a much poorer state of HRQoL, the additional burden of alopecia is unlikely to worsen their HRQoL a great extent due to a floor effect on their HRQoL. Likewise, as patients with AA have improved general physical health compared to patients with NSCLC, the HRQoL impact of AA is greater to the general population than the impact of hair loss on patients with NSCLC. This is consistent with findings in our research. Alopecia areata can be a lifelong chronic condition that affects multiple domains of HRQoL including mental wellbeing and psychosocial functioning over the course of someone’s life. This is cumulative impact which can result in a substantial burden for patients with alopecia areata.

Regarding the literature presented in the CS, Table 45 (B.3.4) outlines what we believe is a thorough analysis of the HRQoL data in AA alongside the analysis of

HRQoL studies (CS, B.3.4.3). These studies did not include SALT scores to grade patients. It is possible that grading severity based on clinician judgment produced bias. Clinicians may consider the wider impacts of AA, especially psychosocial burden when rating patient severity. Participants may have been rated as severe for reasons other than just hair loss.

We argue specifically that the content validity of literature reporting generic measures of HRQoL such as EQ-5D are not appropriate in AA. Evidence supporting these statements can be found in CS B.3.4 and summarised below.

- A substantial body of evidence exists that describes the burden that AA can have on patients' HRQoL (CS, B.1.3.2.3). This is supported by qualitative insight from patient advisory groups and clinical experts (dermatolgists) with an interest in hair loss disorders alongside what we have learned through our vignette study (CS, B.3.4.4.2 and B.3.4.4.4)

- Generic measures of HRQoL capture some, but not all AA patient relevant domains such as social functioning, personal relationships, and appearance (CS B.1.3.2).

- Generic preference-based measures that quantify the HRQoL of patients with AA are not able to overcome the complexity of correctly estimating the burden of AA. They are not specific or sensitive enough to comprehensively capture the full patient experience as reported in the literature alongside PAG and dermatologists with a special interest in hair disorder feedback (CS B.3.4.4.2).

- If generic measures such as EQ-5D scores do not describe the burden of the condition, then this risks inequitable decision making for people with AA. First, the true burden of the condition in terms of its impact on HRQoL will not be understood if that is measured using EQ-5D. Second, the EQ-5D scores will not reflect the gains in HRQL that people may experience if they receive an effective treatment.

- Analysis of AAPPO suggest that as more disease relevant HRQoL domains are incorporated, sensitivity to changes in HRQoL improves, compared with

EQ-5D and SF-36. This suggests that disease specific multi-domain approaches in AA are appropriate potentially critical in capturing the true burden of the condition (CS. B.3.4.1.5).

- There is a clear need for health state utilities which adequately describe the impacts that AA has on the HRQoL.

[redacted]

- We argue based on our findings that the current literature utilising generic measures of HRQoL including EQ-5D, SF-36 cannot overcome the potential limitations in content validity. Therefore, we propose that a vignette approach which seeks to capture the full burden and lived experience of patients with alopecia areata is a more appropriate approach to utility generation in AA (CS B.3.4.1.6).

- Therefore, vignettes developed through direct patient and clinician input, as well as available literature and clinical data, valued by members of the UK general population, are the most valid and equitable source of health state utilities for cost-effectiveness analyses in AA (CS B.3.4.4.3 and B.3.4.4.4). As we have already discussed these utility values are consistent with other similar skin disorders.

B27. CS states on p 171, that "*there is no mapping from DLQI that is able to generate utilities.*" However, mapping studies do exist to map from the DLQI to EQ-5D. These can be identified from the HERC database. (https://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies). Please clarify why these cannot be used to estimate EQ-5D values from the DLQI scores obtained in the trials.

The Company acknowledges that mapping studies do exist to map from the DLQI to EQ-5D. However, the DLQI questionnaire is designed to measure how much a *skin problem* has affected a patient's life, not hair loss problems or alopecia areata. The DLQI refers to 'skin' or 'skin condition' in most of the items, which may contribute to

some insensitivity to AA. Moreover, four dermatologists with a specialist interest in hair disorders (4/8 clinicians, 50%) stated that DLQI is limited; with some domains remaining irrelevant, and as it does not capture the full emotional impact that AA has on patients (i.e., lack of content validity), an alternative tool is required to capture the full impact of AA on HRQoL (CS B.3.4.4.2 (page 173)). Therefore, existing DLQI to EQ-5D mapping algorithms were not considered appropriate and fit-for-purpose. We would also note that we would expect any mapping study attempting to link a measure to EQ-5D in an AA patient population would be greatly limited by the insensitivity of the EQ-5D in AA.

B28. Appendix H, Table 1; Why did the review of quality of life studies exclude validated AA specific HRQoL tools such as AA-QLI, AAQ and AASIS (see Rencz 2016 review, ref 101 of doc B)? Especially when two of these (AASIS and AA-QLI) are endorsed for use in a position statement by the European Academy of Dermatology and Venereology Task Force on Quality of Life and Patient Oriented Outcomes **(**https://onlinelibrary.wiley.com/doi/10.1111/jdv.17370**).**

The review of quality of life studies was focused on identifying studies that reported utilities or any HRQoL scores from instruments that could be mapped to utilities or EQ-5D.

The list of HRQoL scales that can be mapped to EQ-5D/utilities was identified from HERC database of mapping studies, Version 8.0. (Last updated: 14th October 2020). DOI: https://doi.org/10.5287/bodleian:5Rm5D1zke available at: http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies.

None of the AA-specific instruments were included in this published list.

In addition, our targeted searches did not identify any studies which aimed to map scores from AA-specific scales (such as AA-QLI, AAQ, or AASIS) to EQ-5D scores or utilities. For this reason, studies that reported HRQoL scores that could not be mapped to EQ-5D scores/utilities were not included in this SLR.

However, HRQoL outcomes reported in clinical trials using Alopecia Areata Symptom Impact Scale (AASIS) scale, Alopecia areata patient priority outcomes (AAPPO) scale, or other HRQoL tools were captured in the SLR.

B29. Appendix H: Several studies identified in the systematic review by Rencz *et al.* (ref 101 of doc B) were either not identified or were excluded in the systematic review of quality of life studies. Please comment on the reason for excluding these studies in particular for the following:

- de Hollanda 2014 (https://dx.doi.org/10.4103/0974-7753.136748) - which reports SF-36

This study was incorrectly rejected in the SLR. The study is now included and the PRISMA and Appendix H has been updated to reflect this change.

- Fabbrocini 2013 (https://dx.doi.org/10.1111/j.1468-3083.2012.04629.x) - which reports AA-QLI and DLQI

This study was incorrectly rejected in the SLR. The study is now included and the PRISMA and Appendix H has been updated to reflect this change.

- Endo 2012 (https://dx.doi.org/10.1684/ejd.2012.1752) - which reports SF8 and AAQ

This study was incorrectly rejected in the SLR. The study is now included and the PRISMA and Appendix H has been updated to reflect this change.

- Ozturkcan 2006 (http://dx.doi.org/10.1111/j.1365-4632.2006.02881.x) - which reports DLQI

This study was incorrectly rejected in the SLR. The study is now included and the PRISMA and Appendix H has been updated to reflect this change.

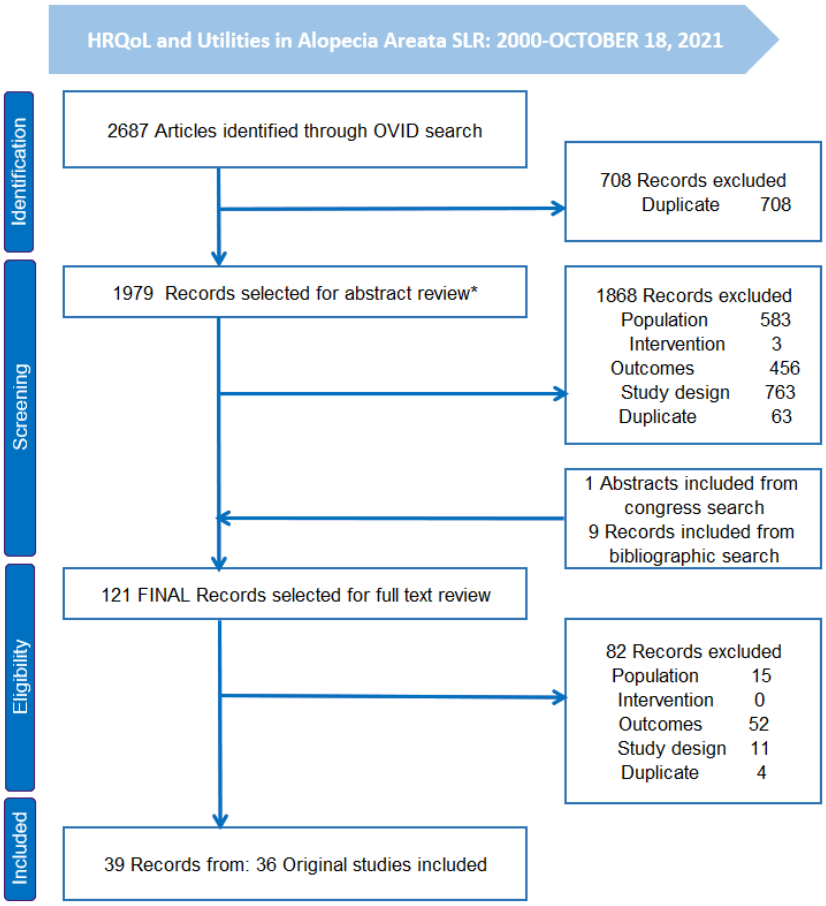
- Willemsen 2010 (https://doi.org/10.1016/j.jaad.2009.06.029) - which reports SF-36

This study was published as a “Letter to the editor” and it was excluded at the first (Title/abstract) stage of the SLR for the reason “Study design”, in accordance to the PICOS criteria.

None of the studies highlighted above provide any additional information to what has already been covered in the company submission. Please see below the updated

PRISMA for the HRQoL SLR (Figure 11). A full copy of the updated appendix H is attached for reference.

Figure 11: Updated PRISMA for the HRQoL SLR



B30. Appendix H: Health-related quality-of-life studies, section H1.3.2 Search string Table 2 shows that the search strategy was conducted in October 2021. Please confirm that the search has been updated and provide the strategy for the updated search.

The search string has not been updated since October 2021. As targeted searches were performed for the Vignette study recent literature was scanned for updates while this study were undertaken and the paucity of data in AA we do not feel any meaningful information would be found from an updated SLR search. Please refer to Appendix H section on the Vignette Study for further information on the targeted search.

Resource use

B31. Please clarify why [redacted] are not included in the model given that these are recommended in the [redacted]

It was assumed that [redacted] would be conducted by patients themselves and by clinicians during routine appointments. Routine appointments for patients treated with ritlecitinib are assumed to occur at least every six months. Therefore, it was assumed that [redacted] would not require any additional routine tests or incur additional costs and were not included in the model.

B32. Priority question. The model assumes that patients who reach a SALT score above 20 at any time after the first 48 weeks will stop treatment with stopping rule being applied every model cycle (12 weeks). Please clarify how the resource use associated with assessing response every 12 weeks is captured in the economic model.

Resource use in the model includes a variety of inputs per 12 weeks where assessment of response would occur. These inputs include an [redacted] appointment with a dermatology nurse in outpatient setting, an [redacted] appointment with a dermatology nurse in a primary care setting, an [redacted] dermatology-related General Practitioner (GP) visit and an [redacted] dermatology outpatient visit for patients with SALT 11-20. This equals to [redacted] appointments per year (i.e., [redacted]) for patients within this category. It is assumed within the model that response would be assessed during this routine monitoring, thus triggering the stopping rule if appropriate. It is also assumed that should a patient begin to worsen on treatment, they would seek an appointment to discuss their treatment.

B33. We cannot identify an NHS reference cost of £171.93 for a consultant led Dermatology outpatient visit (service code 330) or an NHS reference cost of £221.52 for a Clinical Psychology outpatient appointment (service code 656) in the 2020-21 NHS reference cost source cited. Please check your source and provide more specific information, such as the currency code. For example, a non-admitted face-to-face follow-up attendance for a Dermatology outpatient visit which is consultant led has a currency code of WF01A and a unit cost of £179.78. Providing an extract

of the NHS reference costs excel file may be the easiest way to share this information.

In the National Schedule of NHS Costs, the unit cost of a consultant-led dermatology outpatient visit was derived from the 'Total Outpatient Attendance' (service code 330, cell G94). The corresponding unit cost is £171.93.[22]

The unit cost for a Clinical Psychology outpatient appointment was also derived from the 'Total Outpatient Attendance' (service code 656, cell D127). The corresponding unit cost is £221.52.[22]

For convenience, an extract of the NHS reference costs excel file has been provided.

B34. The NHS reference costs for outpatient visits (e.g. Dermatology and Clinical Psychology) are the cost per visit not the cost per hour. Please clarify why these unit costs have been further adjusted for duration of appointment within the model. In addition, the PSSRU unit cost for a GP appointment is £39.23 per appointment lasting on average 9.22 minutes as stated in Table 53 and therefore does not need adjusting for duration within the model. Please provide an updated base case model in which these unit costs are applied per appointment.

The Company agrees this was an omission and has corrected the net costs for psychological support consultation, dermatology nurse (outpatient setting), dermatology-related nurse visit (primary care setting), dermatology-related GP visits and dermatology outpatient visits as shown in Table 14. The net costs have been revised to not adjust for time when a cost per visit is referenced or when the reference has already adjusted the cost to the correct duration (e.g. GP appointment). An updated base case model has been provided in the Appendix.

Table 14: AA management unit costs

Event	Cost, £	Duration (minutes)	Net cost, £	Reference
Wigs	[redacted]	N/A	[redacted]	Expert opinion – average NHS spends on a synthetic wig.[24]
Wig service for fitting/collection	123.00	[redacted]	[redacted]	PSSRU unit costs report 2021 - Estimated cost of a wig service with a consultant dermatologist (assumed to be a medical hospital based consultant), 10 minutes.[23]

Psychological support consultation	221.52	N/A	221.52	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. Total Outpatient Attendance. Service code 656 - Clinical Psychology.[22]
Dermatology nurse (outpatient setting)	149.15	N/A	149.15	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. Currency code WF01A (Non-Admitted Face-to-Face Attendance, Follow-up).[22]
Dermatology-related nurse visit (primary care setting)	44.00	[redacted]	[redacted]	PSSRU unit costs report 2021 - Nurse (GP practice), 15 minute appointment (duration based on expert opinion).[24 23]
Dermatology-related GP visits	39.23	N/A	39.23	PSSRU unit costs report 2021 - GP unit cost, 9.22 minute appointment (duration based on expert opinion).[23,24]
Dermatologist outpatient visit	171.93	N/A	171.93	National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. Total Outpatient Attendance. Service code 330 - Dermatology (consultant led).[22]

Abbreviations: AA, alopecia areata; GP, general practitioner; NHS, National Health Service; PSSRU, Personal Social Services Research Unit

B35. There appears to be an error made in calculating the costs for the health states (E49:E57 in 'Health State Costs' sheet) in that the unit costs for some activities appear to have been adjusted for duration in cells E36:K36 and then further adjusted by multiplying by the duration in minutes (cells E37:K36) and then dividing by 60. This applies to wig collection service, dermatology nurse outpatient appointments, and primary care nurse visits. Please correct in an updated base case.

The Company can confirm that there was an error of adjusting twice for the duration of resource use. This has now been amended in the model in cells F36, H36 and I36 with the only adjustment taking place in cells E37:K36.

B36. Please clarify why the NHS reference cost for a non-consultant led outpatient appointment was not used to capture an outpatient visit to a dermatology nurse (e.g., £149.15 for a Non-Admitted Face-to-Face Attendance, Follow-up) instead of the cost per hour for a hospital-based band 6 nurse. Please also clarify why a hospital-based consultant dermatologist is required for a wig fitting/ collection appointment.

The Company agrees that this variable was incorrectly costed and has updated the unit cost for an outpatient visit to a dermatology nurse (£149.15). The value in the model is now based on the National Schedule of NHS costs for a non-consultant led outpatient appointment (currency code WF01A).

A hospital-based consultant dermatologist was assumed the most suitable resource for a wig fitting/collection appointment due to the necessity for a resource in this context. To assess the impact of this assumption, a scenario which excluded this cost has been provided in the Appendix.

B37. Please clarify how the data presented in Table 51 and 52 have been calculated from data presented in Appendix I, with reference to the specific tables or figures within Appendix I on which the estimates are based.

The Company has updated the resource use applied in the economic model (previously Table 51 and Table 52 (CS B.3.5.2. (page 180)). This is summarised in [redacted] and [redacted].

[redacted] 15 [redacted]

Event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wigs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wig service for fitting/collection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Psychological support consultation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology nurse (outpatient setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related nurse visit (primary care setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related GP visits	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatologist outpatient visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; SALT, Severity of Alopecia Tool

[redacted] 16 [redacted]

Event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wigs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wig service for fitting/collection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Psychological support consultation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology nurse (outpatient setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related nurse visit (primary care setting)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatology-related GP visits	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dermatologist outpatient visit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; SALT, Severity of Alopecia Tool

As detailed on Appendix I, page 12, a health care resource use (HCRU) Delphi panel was assembled to develop a comprehensive understanding of health care resource utilisation associated with the management of AA in the UK.[24]

In terms of how the data presented in ***[redacted]15 (previously Table 51 (CS B.3.5.2. (page 180)) has been calculated; Figure 9, Appendix I, details the clinicians responses to: 'for patients receiving JAK inhibitors please indicate your level of agreement for the average number of appointments per year'. If more than three clinicians disagreed with the estimates in the second-round survey (which were attained from the first-round survey), then that particular appointment was considered for exclusion from the model.

As depicted in Figure 9, Appendix I, more than three clinicians disagreed with the average frequency of dermatology-related nurse visit (primary care setting). However, the disagreements may have been because the numbers presented were averages. The primary care clinician indicated that the number of dermatology-related nurse visits in a primary care setting were within the ranges suggested by secondary care clinicians in the first-round survey (apart from for patients with SALT 100), therefore, the averages were still used in the model

As evident in Figure 9, Appendix I, it was estimated that if a patient had SALT ≤10, then they would visit a dermatology nurse (outpatient setting - NHS) [redacted] per year, or [redacted] times every 12 weeks. If a patient had SALT 50-99, they were estimated to visit a dermatology nurse (outpatient setting - NHS) [redacted] times per year or [redacted] times per 12 weeks.

In terms of how the data presented in ***[redacted]16 (previously Table 52 (CS B.3.5.2. (page 181)) has been calculated; Figure 5, Appendix I, details the clinicians responses to: 'for patients receiving no pharmacological therapy please indicate your level of agreement for the average number of appointments per year' (second-round survey). As mentioned above, if more than three clinicians disagreed with the estimates in the second-round survey (which were attained from the first-round survey), then that particular appointment was considered for exclusion from the model. As depicted in Figure 5, Appendix I, more than three clinicians did not disagree with the average frequency, therefore, no appointments were excluded.

As evident in Figure 5, Appendix I, it was estimated that if a patient had SALT ≤10, then they would have a dermatology-related GP visit [redacted] per year, or [redacted] times every 12 weeks. If a patient had SALT 21-49, they were estimated to have [redacted] visits per year, or [redacted] times per 12 weeks.

In order to calculate the ritlecitinib and BSC treatment resource utilisation, the above methodology was employed for the majority of appointments outlined in Appendix I.

Data for the resource utilisation associated with wigs was attained by averaging the responses in the first-round survey from the following question: 'What proportion of the following patients (between 0-100%) get the maximum number of wigs available from the NHS per year?' (Figure 27, Appendix I). The averages from all SALT categories are presented below (***[redacted]17). As the NHS funds up to 2 wigs per patient with AA per year, it was assumed that the maximum number of wigs that are provided was 2 per year (CS B.3.5.2. (page 180)). Therefore, to attain the value of [redacted] per 12 weeks for wigs under the SALT 100 category in [redacted], the following calculation was conducted: [redacted]. This methodology was implemented with the relevant proportion of patients who receive the maximum number of wigs for all remaining SALT categories.

[redacted]17

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The data presented in [redacted] and [redacted] for wig service for fitting/collection is slightly different to the data in Figure 9 and Figure 5 (Appendix I), respectively. This is because this resource use was set to align with resource use of wigs, i.e. it was assumed that patients using wigs would have [redacted] wig fitting per year.

B38. Please clarify how compliance was estimated from the ALLEGRO-phase 2a/3 study. Was it estimated specifically for the 50mg treatment arm? How were the data

obtained? Was compliance measured in the ALLEGRO-LT study and if so, was it similar?

Compliance of the investigational product was monitored in the ALLEGRO Phase 2b/3 study by delegated site personnel by the accounting of unused medication returned by the participant at study visits. Compliance was estimated specifically for the 50mg treatment arm (at Week 48 it was [redacted]). Please refer to the ALLEGRO Phase 2b/3 (B7981015) CSR for compliance in all treatment arms up to Week 24 (Table 14.4.1.4.1, page 1674) and Week 48 (Table 14.4.1.4.2, page 1675).

In the ALLEGRO-LT study compliance was measured in the same way as in the Phase 2b/3 study (by delegated site personnel by the accounting of unused medication returned by the participant at study visits). Compliance was similar in the ALLEGRO-LT study to what had previously been observed in the Phase 2b/3 study. Please refer to Table 14.4.1.3 in the ALLEGRO-LT (B7981032) interim CSR (page 1649).

B39. Appendix I: Cost and healthcare resource identification, measurement and valuation, section I1.3.2 Search string Table 2 shows that the search strategy was last conducted in October 2021. Please confirm that the search has been updated and provide the strategy for the updated search.

The search string has not been updated since October 2021. During development of the HCRU Delphi panel study, targeted searches were performed and recent literature was scanned. We do not feel any meaningful information would be found from an updated SLR search particularly given the paucity of data identified in the original SLR dated October 2021. Please refer to HCRU Delphi Panel report which can be found in appendix I and in the reference pack > Pfizer data on file Understanding HCRU of AA in the UK.

B40. CS, page 131, Table 34: The EAG note that the proportion aged 12 to 18 years in the model (cells E45 of 'Model Settings' sheet) does not match the figure reported in Table 34 and the mean age across the whole cohort (CSR, Table 14.1.2.1) does

not match the mean age for those aged age ≥12 years in Table 34. Please clarify how these figures were obtained/calculated and account for any discrepancy.

Table 34, Document B does not report the proportion aged 12 to 18 years in the model. The values included in Table 34 are the percentage of patients aged ≥ 12 to < 18 years who are female, the percentage of patients aged ≥ 18 years who are female and the percentage of patients aged ≥ 12 who are female. These values are presented below in [redacted], now also including the percentage of patients aged 12 to 18 years in the model for transparency.

[redacted] 18 [redacted]

Parameter	[redacted]	Reference
Mean age - ≥12 to <18 years	[redacted]	ALLEGRO 2b/3 trial[32]
Mean age - ≥18 years	[redacted]	
Mean population age - ≥12 years	[redacted]	
Percent female - ≥12 to <18 years (%)	[redacted]	ALLEGRO 2b/3 trial[32]
Percent female - ≥18 years (%)	[redacted]	
Percent female - ≥12 years (%)	[redacted]	
Model population aged ≥12 to <18 years	[redacted]	ALLEGRO 2b/3 trial[32]

The proportion of patients who are adolescent in the model is calculated by using the number of adolescents within the ALLEGRO 2b/3 trial divided by the number of adults and adolescents combined.[32] The mean age is then calculated using the formula displayed below:

$$\left(\bar{x}_{adolescent\ age} \times \%_{adolescent}\right) + \left(\bar{x}_{adult\ age} \times (1 - \%_{adolescent})\right)$$

The discrepancy between the mean age in the CS and the one stated in the CSR, similar to the values used to calculate the proportion which are adolescent, exists because the CS value is based upon the number of patients across all treatment groups, rather than the ritlecitinib 50 mg treatment arm only. These total values are displayed below in [redacted].

[redacted] 19 [redacted]

Parameter	[redacted]	[redacted]
Number aged ≥12 to <18 years	[redacted]	[redacted]
Number ≥18 years	[redacted]	[redacted]

B41. Priority question. 'Ritlecitinib 50mg Final Cals' sheet, Cells P24:W24 (and equivalent cells on the 'Ritlecitinib 50mg I+F Calcs' and 'BSC Calcs' sheets). When calculating total discounted QALYs gained for time spent in each health state (P22:W22), the utility of the health state (P18:W18) is multiplied by the discounted years (P17:W17). Whereas, when calculating total discounted QALYs lost due to disutility (P24:W24), the number of discounted cycles (P14:W14) is multiplied by the sum of the disutilities for adverse events (P19:W19) and caregivers (P20:W20). Why has a different method been used, and does this mean that the cycle length duration is not accounted for when calculating discounted QALYs lost due to disutility? Please correct within an updated base case analysis if an error has been made.

The company agrees that the original model calculations were incorrect. 'Total QALYs due to disutility' and 'Total QALYs lost due to disutility' (rows 23:34, columns P:W) in 'Ritlecitinib 50mg Final Cals', 'Ritlecitinib 50mg I+F Calcs' and 'BSC Calcs' sheets have been amended to be calculated using life years instead of 'Total cycles in health state'.

B42. Priority question. Please clarify why the 'PSA Calcs' sheet only refers to the 'Final Results Calcs' sheet and not to the 'Int. Results Calcs' sheet when the EAG's understanding is that the latter contains the results when implementing both the interim and final stopping rules, which is the company's base case (as these align with deterministic results in Table 59 and description of base case in Table 61)? Please provide a model with functionality to extract the PSA results for the analysis implementing both interim and final stopping rules.

The Company has updated the model to include functionality within the 'PSA calcs' sheet (D14:G14) whereby the selecting either the interim and final, or only final stopping rules on 'Clinical' sheet will allow the PSA to pull results from either the interim results or final results.

B43. CS, Table 60: The PSA ICER reported in Table 60 appears to be the average of the ICERs generated by each individual PSA run rather than the average of the incremental costs divided by the average of the incremental QALYs, which would be £11,265 [redacted] instead of £11,708. Please confirm that this is an error.

The Company can confirm that this is an error and have updated this within the model. The new PSA ICER is equal to the £13,383 [redacted] found by the EAG.

B44. Excel Model, Raw TM data: For the probabilistic model, clarify why the continuity correction was applied by dividing two observations across the 4 SALT categories (hence adding 0.5) instead of just dividing an additional unit equally across all health states. Also, clarify why the correction was applied where observations were sufficient (e.g., J55:M55). Please provide a reference to support the company's preferred approach to handling low or missing events for some elements of the transition matrix.

The continuity correction was applied to all cells of the transition matrices to accommodate observations of missing data and low patient count. This correction was applied to all cells to ensure that all values were skewed consistently and equally across all transitions. The value of 0.5 is not linked to the number of observations or health states, but simply part of the continuity correction to eliminate sampling on values of 0 which can lead to errors.

B45. Excel Model, Raw TM data: For the deterministic model, the transition probabilities used imply that some transitions are impossible, for instance no-one can move from SALT 11-20 to SALT ≥50 for transitions at month 18 (cell J63). Please clarify why a continuity correction was not applied in the deterministic model, similar to that used in the probabilistic model.

A continuity correction is not needed in the deterministic model as there is no chance of sampling leading to errors throughout the model (for example, leading to #DIV/0! or similar errors which will break the model). A continuity correction is needed only on sampling zero or small numbers as dividing by zero is impossible and leads to errors within excel.

B46. In the scenario in which the final stopping rule is applied at 36 weeks instead of 48 weeks, please describe what transitions apply between 36 weeks and 48 weeks

for responders and non-responders? In particular, please clarify why cell H25 of the 'Ritlecitinib 50mg I+F Calcs' sheet refers to data in F15 which is from 36 weeks, whilst cell H29 refers to G14 which is from 48 weeks. If this an error, please correct in an updated base case.

When the stopping rule is applied at Week 36, it is assumed that all patients with a SALT score ≤20 will discontinue at that point. Patients who remain on treatment with ritlecitinib then remain within their health states for the next cycle. Patients who discontinue treatment begin the transition back to the health state with a SALT score >50.

Cell H29 of the 'Ritlecitinib 50mg I+F Calcs' sheet was identified as an error and has now been updated within the model, such that when the stopping rule is applied at Week 36, Week 48 data matches the previous cycle.

Section C: Textual clarification and additional points

C1. CS, page 24. Please clarify that the sentence "*Research and Surveillance Centre (RSC) network database between 2009-2018 (N=2,634,083) reported an overall diagnosed point prevalence of 0.58% among adults with AA*" is meant to say, "*an overall diagnosed point prevalence of AA of 0.58% among adults*".

This is correct, "*an overall diagnosed point prevalence of AA of 0.58% among adults*".

C2. Please provide CSRs for all Pfizer trials referred to in the submission or if these have already been provided, please provide a list of the file names to identify which document is the CSR for each trial.

Study	File Location>name
ALLEGRO-2a safety study (ALLEGRO B7981037)	This CSR can be found in the ACIC reference pack.
ALLEGRO-2b/3 Pivotal trial and dose ranging study (ALLEGRO B7981015)	This CSR can be found in the ACIC reference pack.

ALLEGRO-LT Long Term Study (ALLEGRO B7981032)	This CSR can be found in the ACIC reference pack.

C3. Please provide reports summarising methods and findings for all studies described in Table 1 of document B or provide the names of the files, or cross references to relevant sections of the appendices, if these reports have already been provided.

Study	Location
Clinical practice and therapeutic landscape Delphi panel (Therapeutic landscape Delphi panel)	The methods and findings of this study can be found in the ACIC reference pack.
Qualitative research in alopecia – PAG Study	The methods and findings of this study can be found in the ACIC reference pack.
Vignette study for utility estimation in Alopecia Areata	An updated report can be found in the ACIC reference pack.
Health care resource utilisation Delphi panel	The methods and findings of this study can be found in the ACIC reference pack.

Abbreviations: AA, alopecia areata; PAG, patient advisory group.

References

1. 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. 85: 379–387.
2. Guttman-Yassky E, Pavel AB, Diaz A, *et al.* Ritlecitinib and brepocitinib demonstrate significant improvement in scalp alopecia areata biomarkers. *Journal of Allergy and Clinical Immunology* 2022. 149: 1318–1328.
3. Winnette R, Banerjee A, Sikirica V, *et al.* Characterizing the relationships between patient-reported outcomes and clinician assessments of alopecia areata in a phase 2a randomized trial of ritlecitinib and brepocitinib. *Journal of the European Academy of Dermatology and Venereology* 2022. 36: 602–609.
4. King B, Zhang X, Harcha WG, *et al.* Efficacy and Safety of Ritlecitinib (PF-06651600) in Patients With Alopecia Areata and ≥50% Scalp Hair Loss: Results From the International ALLEGRO Phase 2b/3 Randomized, Double-Blind, Placebo-Controlled Study (NCT03732807). in (2021).
5. Mesinkovska N, Shapiro J & King B. 33183 Efficacy of the oral JAK3/TEC inhibitor ritlecitinib (PF-06651600) in patients with alopecia areata over 48 weeks: Results from the ALLEGRO Phase 2b/3 randomized, double-blind, placebo-controlled trial. *Journal of the American Academy of Dermatology* 2022. 87: AB54.
6. Sinclair R & Mesinkovska N. 33280 Improvement in patient-reported hair loss outcome measures in patients with alopecia areata treated with ritlecitinib: 48-week results from the ALLEGRO phase 2b/3 randomized, double-blind, placebo-

controlled trial. *Journal of the American Academy of Dermatology* 2022. 87: AB69.

7. King B, Szepietowski JC & Farrant P. 33182 Safety of the oral JAK3/TEC inhibitor ritlecitinib (PF-06651600) in patients with alopecia areata: Results from the ALLEGRO phase 2b/3, randomized, double-blind, placebo-controlled trial. *Journal of the American Academy of Dermatology* 2022. 87: AB106.
8. Guttman-Yassky E, Ito T, Jabbari A, *et al.* Clinical efficacy of the oral JAK3/TEC inhibitor ritlecitinib (PF-06651600) and patients' perception of improvement in alopecia areata: Results from the ALLEGRO Phase 2b/3 trial. in *EXPERIMENTAL DERMATOLOGY* (WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2021). 30: 20–20.
9. Hordinsky MK, Hebert A & Gooderham M. 33180 Efficacy and safety of the oral JAK3/TEC inhibitor ritlecitinib in adolescents with alopecia areata: Results from the ALLEGRO Phase 2b/3 randomized, double-blind, placebo-controlled trial. *Journal of the American Academy of Dermatology* 2022. 87: AB51.
10. NICE. 2022. Single technology appraisal and highly specialised technologies evaluation: User guide for company evidence submission template. Available at: https://www.nice.org.uk/process/pmg24 [Date accessed: November 2022].
11. Humphreys K, Blodgett JC & Roberts LW. The exclusion of people with psychiatric disorders from medical research. *Journal of Psychiatric Research* 2015. 70: 28–32.
12. ICH. 2021. General considerations for clinical studies E8(R1). Available at: https://database.ich.org/sites/default/files/ICH_E8-R1_Guideline_Step4_2021_1006.pdf.
13. Pfizer data on file. 2023. LCR analysis.

14. *Pfizer data on file. 2023. Supplementary subgroup analyses.*
15. King B, Guttman-Yassky E, Peeva E, *et al.* Safety and Efficacy of Ritlecitinib and Brepocitinib in Alopecia Areata: Results from the Crossover Open-Label Extension of the ALLEGRO Phase 2a Trial. *JID Innovations* 2022. 2: 100156.
16. Peeva E, Guttman-Yassky E, Banerjee A, *et al.* Maintenance, withdrawal, and re-treatment with ritlecitinib and brepocitinib in patients with alopecia areata in a single-blind extension of a phase 2a randomized clinical trial. *Journal of the American Academy of Dermatology* 2022. 87: 390–393.
17. Crispin MK, 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.
18. 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:
19. Park H-S, Kim M-W, Lee JS, *et al.* Oral tofacitinib monotherapy in Korean patients with refractory moderate-to-severe alopecia areata: a case series. *Journal of the American Academy of Dermatology* 2017. 77: 978–980.
20. 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.
21. Pfizer data on file. UK interviews with dermatologists with a specialist interest in hair disorders for the use of ritlecitinib for patients with AA. 2022.
22. NHS England. National Cost Collection for the NHS - National schedule of NHS costs 2020/21. at <https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/>

23. PSSRU. Unit Costs of Health & Social Care 2021. 2021. at <https://kar.kent.ac.uk/92342/25/Unit%20Costs%20Report%202021%20-%20Final%20version%20for%20publication%20%28AMENDED2%29.pdf>
24. Pfizer data on file. Understanding health care resource utilisation of alopecia areata in the UK. 2022.
25. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Statistical Power Analysis for the Behavioral Sciences [Internet]. 2013 May 13 [cited 2022 Oct 26]; Available from: https://www.taylorfrancis.com/books/mono/10.4324/9780203771587/statistical-power-analysis-behavioral-sciences-jacob-cohen.
26. Pfizer data on file. Does the EQ-5D measure the full impact of alopecia areata on patients' quality of life? 2022.
27. Pfizer data on file. Does the EQ-5D measure the full impact of alopecia areata on patients' quality of life? [updated manuscript]. 2023.
28. Yang Y, Brazier J & Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *The European Journal of Health Economics* 2015. 16: 927–939.
29. Song HJ, Park H, Park S-Y, *et al.* Estimation of Health Utilities Based on the Response to Treatment in Atopic Dermatitis: a Population-based Study. *Clinical Therapeutics* 2019. 41: 700–713.
30. Matusiak Ł, Bieniek A & Szepietowski JC. Psychophysical aspects of hidradenitis suppurativa. *Acta dermato-venereologica* 2010. 90:
31. Nafees B, Stafford M, Gavriel S, *et al.* Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008. 6: 84.

32. Pfizer. *A Phase 2b/3 randomised, double blind, placebo-controlled,dose ranging study to investigate the efficacy and safety of PF-06651600 in adult and adolescent alopecia areata (AA) subjects with 50% or greater scalp hair loss.* (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT03732807>

Appendix

[redacted]

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
BSC	[redacted]	[redacted]	[redacted]				
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	0	[redacted]	13,178

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life year

[redacted]

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	13,383

[redacted]

[redacted]

Abbreviations: QALY, quality-adjusted life year

[redacted]

[redacted]

Abbreviations: QALY, quality-adjusted life year; WTP, willingness to pay

[redacted]

[redacted]

Table 20: Scenario analyses of the base case of the model

Model setting tests	Base case assumption	Scenario assumptions	ICER of ritlecitinib relative to BSC (£)
Base case	-	-	13,179
Perspective	Payer	Societal	[redacted]
Time horizon	Lifetime	5 years	[redacted]
Model death in first 48 weeks?	No	Yes	[redacted]
Age group	≥12 years	≥18 years	[redacted]
		≥12 to <18 years	[redacted]
Stopping rule criteria	Interim+Final	Final Only	[redacted]
Final SALT score	SALT ≤ 20	SALT ≤ 10	[redacted]
Final stopping rule time point	48 weeks	36 weeks	[redacted]
BSC SALT 11-49 revert to SALT 50-100?	Yes	No	[redacted]
Allow ritlecitinib treatment discontinuers to have spontaneous remission?	Yes	No	[redacted]
Discontinue patients based on SALT score after 48 weeks	SALT ≤ 20	Do not discontinue	[redacted]
Extrapolation of LT data after 24 months	Stay in state	Last observation carried forwards	[redacted]
		Average	[redacted]
Treatment discontinuation rate curve	Exponential	Weibull	[redacted]
		Gompertz	[redacted]
		Log-logistic	[redacted]
		Lognormal	[redacted]
SALT >50 HCRU assumption	SALT 50-99	SALT 100	[redacted]
Utility weight source	TTO Analysis	TTO Analysis (SA)	[redacted]
Include carer disutility	Yes	No	[redacted]
Disutility weight source	TTO Analysis	TTO Analysis (SA)	[redacted]
Spontaneous remission probability	1.54%	10.00%	[redacted]
Patients considered in transition matrices for long-term response	*De novo* ALLEGRO-LT patients or ALLEGRO 2b/3 patients who had received maintenance doses of 30 mg or 50 mg	Long-term response based on only those receiving licensed dose during ALLEGRO-LT	[redacted]

		Long-term response based on only those receiving licensed dose during both ALLEGRO 2b/3 and ALLEGRO-LT	[redacted]
Source of AE costs	DRG codes, NHS reference costs	GP appointment only, PSSRU	[redacted]
Caregiver disutility population	≥12 years	≥12 to <18 years	[redacted]
Include wig fitting cost	Yes	No	[redacted]

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LT, long term; SA, sensitivity analysis; SALT, Severity of Alopecia Tool; TTO, time-trade off

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ritlecitinib for treating severe alopecia areata in people 12 years and over (ID4007)

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.
- Your response should not be longer than 10 pages.

About you

1.Your name	**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 £9250 from the National Lottery Community fund
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	An independent research grant from Pfizer was won by Alopecia UK in 2021. Value £55,026.28 The scope of that piece of research is a survey to explore the social and economic impact of alopecia areata (including totalis & universalis). Research being carried out in collaboration with the University of West of England. Eli Lilly provided unrestricted support as a corporate member £20,000. As did Concert Pharmaceuticals £7699. Soterios paid for PPI fee and trail recruitment adverts

the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No None
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 face to face meetings & events • 1:1 telephone support calls and emails • Facilitate PPI meetings (Public and Patient Involvement) meetings for alopecia related research • Own patient research questionnaires – findings published

Living with the condition

6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	• 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 scalp ointment which I've had before and has little chance of working', 'My kidsreally want to go swimming, but I haven't been since losing all my hair. I don't know if I am brave enough to go', 'I started with alopecia 3 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 I had to take time away from work and family', 'This morningI lost a clump of long hair that came off in my hands in the shower, yesterday one eyebrowfell 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 and even suicidal thoughts. • AUK research 2017 – clinically significant levels of anxiety in 35.5% and depression in 29% • 25% 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 and 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 licenced for alopecia and have limited access on the NHS • In AUK studies over 25% people voiced having hair loss had affected their close, intimate relationships • For men with alopecia there is 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 • AUK understand that approx. 40% people with alopecia areata have another autoimmune condition – such as lupus, thyroid conditions, 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, body hair). With no eyelashes you often suffer watering eyes 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 mucous • 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 normalso 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 told to 'wait and see' if more hair sheds or it grows back. Very few treatments 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 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)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 few clinical and patient care guidelines than other skin conditions
8. Is there an unmet need for patients with this condition?	Yes – absolutely! There is no on-label product for alopecia areata. 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 which is generated It gives people the hope that will then be able to live a 'normal' life and 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, were currently there is none. People have viewed the side effects and feel the benefits out-way 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 as Asian People in these groups are likely to benefit proportionally 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 post code 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 degree of hairloss 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 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

Ritlecitinib for moderate-to-severe alopecia areata (aged 12 and over) [ID4007]

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, [redacted] on behalf of the BAD's guideline development group, and [redacted] and [redacted] on behalf of the British Hair & Nail Society
2. Name of organisation	British Association of Dermatologists
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 ~~or No~~ A specialist in the treatment of people with this condition? Yes ~~or No~~ A specialist in the clinical evidence base for this condition or technology? Yes ~~or No~~ 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; more recently has approved ritlecitinib for children aged 12 years and above, young people and adults with moderate-to-severe 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 and ritlecitinib 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 both children, young people and adults with moderate-to-severe AA. Despite low rates of success, commonly used systemic treatments (dosed based on weight in children) for chronic AA include prednisolone, ciclosporin, methotrexate and azathioprine. Sulfasalazine is not as commonly prescribed. The treatment response to these immunosuppressants is very variable. Immunosuppression is not commonly used in children due to the high rate of relapse and associated long-term risks. Most commonly, children are treated with

	topical immunotherapy which may require frequent hospital appointments, therefore, interfering with their studies.
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 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 24 weeks as per the phase 2b/3 ALLEGRO trial https://clinicaltrials.gov/ct2/show/NCT03732807?id=NCT03732807&draw=2&rank=1 and https://www.jaad.org/article/S0190-9622(22)01285-3/fulltext.
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.)	In the last few years, there is increasing familiarity with the use of JAK inhibitors for atopic dermatitis. Dermatologists are likely to be informed on the general contraindications, pre-screening investigations and monitoring required for patients who are commenced on JAK inhibitors.
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	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 clinical trials. AA is also

substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	associated with time away from work for adults or for parents of children with AA, 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?	Ritlecitinib is particularly innovative as it will provide a potential effective therapeutic option for AA patients aged 12 and above and with a better side effects profile and level of immunosuppression than conventional systemic agents (i.e. ciclosporin, methotrexate and azathioprine). The clinical trial recruited children 12 years and older with SALT score of >50%, therefore, severe cases of alopecia areata which are often recalcitrant to treatments. 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. Children also find difficult wearing wigs as they feel they can get bullied in school. Ritlecitinib 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 psychological and 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	Until recently, there was no licensed systemic treatment for AA; the FDA approved baricitinib for the treatment of adults with severe AA in the US in June 2022, and more recently ritlecitinib for children aged 12 years and above,

particular unmet need of the patient population?	young people and adults with moderate-to-severe AA. The expert paper (Meah *et al.* https://pubmed.ncbi.nlm.nih.gov/32165196/) reported that consensus was achieved regarding preferred second-line agents for AA with the following statement: "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?	Ritlecitinib has a similar side effect profile to other JAK inhibitors used currently in the UK for atopic dermatitis. These may resolve/can be treated and may not necessarily lead to treatment discontinuation. If adverse effects occur, upper respiratory tract infections, cutaneous HSV/VZV and acne may affect the patient's QoL, However, other rare side effects reported in the clinical trials include sensorineural deafness in a very small proportion of patients. Significant untoward side effects may necessitate treatment discontinuation. From the ALLEGRO trial, two serious treatment-emergent adverse events were reported in adolescents, both in the 10- mg group: suicidal behaviour (considered not related to study drug, no change in dose) and eczema (considered related to study drug, patient discontinued). https://www.jaad.org/article/S0190-9622(22)01285-3/fulltext Data with ritlecitinib show that it does not have some of the side effects associated with JAK1 inhibitors, in particular changes in the lipid profile. The study presented by King et al. showed that 67% of patients that received ritlecitinib reported at least one adverse effect. The most common adverse events observed in the study were headache (5/48 [10%]), nasopharyngitis (5/ 48 [10%]), and upper respiratory tract infection (4/48 [8%]). Other common side effects include acne (5/48 [10%]) and nausea. There were no clinically significant differences from baseline in electrocardiogram findings, haematology tests, or vital signs, 1 patient had a decreased lymphocyte count. 2 patients discontinued ritlecitinib due to an adverse effect, among them: 1 patient (1/48 [25]) developed angioedema and the other one (1/48 [25]) an increased level of serum creatine phosphokinase. 8 patients developed mild-to-moderate herpes zoster. In the ritlecitinib 50 mg group, there was one case of pulmonary embolism. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8860347/pdf/dddt-16-363.pdf

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 know of as there is little real-world experience on the use of ritlecitinib in the UK and published data on adverse effects are from clinical trials.
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?	There is little real-world experience on the use of ritlecitinib in the UK.

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 people living in certain geographic locations are 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

24. 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. Increased suicide risk has also been noted in adolescent children https://pubmed.ncbi.nlm.nih.gov/24528416/. • There is a significant unmet, clinical need for a safe, effective and approved medication for children and young people aged 12 years and over and adults with moderate-to-severe AA. Some of the current available treatments, such as diphenylcyclopropenone, require several hospital attendances which can impact on children's schooling and place time/financial pressure on the parents. • Initial trial data to date indicate that this treatment is effective and with a relatively good safety profile. • In this study, statistically significantly higher proportions of patients treated with ritlecitinib 30 mg and 50 mg (with or without the loading dose) had 80% or more scalp hair coverage (SALT≤20) after 6 months of treatment versus placebo. The overall safety data demonstrated that ritlecitinib was well tolerated both in adult and adolescent patients.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

NICE National Institute for Health and Care Excellence

The information that you provide on this form will be used to contact you about the topic above.

Please 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.

Confidential until published

Ritlecitinib for treating severe alopecia areata in people 12 years and over. A Single Technology Appraisal

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

Authors Emma Hock, Senior Lecturer, ScHARR, University of Sheffield, Sheffield, UK

Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK

Andrew Metry, Research Associate, ScHARR, University of Sheffield, Sheffield, UK

Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Sarah Ren, Research Associate, ScHARR, University of Sheffield, Sheffield, UK

Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK

Gill Rooney, Programme Manager, ScHARR, University of Sheffield, Sheffield, UK

Correspondence Author Emma Hock, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Date completed Date completed (15/03/2023)

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135708

Declared competing interests of the authors

None of the authors has any conflicts of interest to declare.

Acknowledgements

We would like to thank Dr Victoria Akhras, Dr Anton Alexandroff and Dr Iaisha Ali for clinical advice relating to this project. We would also like to thank Paul Tappenden, ScHARR, for providing comments on the draft report and Gill Rooney, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

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

This report should be referenced as follows:

Hock E; Davis S; Metry A; Hamilton J; Ren S; Wong R; Rooney G. Ritlecitinib for treating severe alopecia areata in people 12 years and over: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2023.

Contributions of authors

Ruth Wong critiqued the company's search strategy. Emma Hock and Gill Rooney summarised and critiqued the clinical effectiveness data reported within the company's submission. Jean Hamilton and Sarah Ren critiqued the statistical aspects of the submission. Sarah Davis and Andrew Metry critiqued the health economic analysis submitted by the company. All authors were involved in drafting and commenting on the final report.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BOXES

ABBREVIATIONS

AA	Alopecia areata
AAPPO	Alopecia Areata Patient Priority Outcomes
AE	Adverse event
AEP	All-exposure pool
AIC	Akaike information criterion
AQoL-8D	Assessment of Quality of Life – 8 Dimensions
AT	Alopecia totalis
AU	Alopecia universalis
BAEP	Brainstem auditory evoked potential
BAD	British Association of Dermatologists
BIC	Bayesian information criterion
BSC	Best supportive care
CASP	Critical Appraisal Skills Programme
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
CRD	Centre for Reviews and Dissemination
CS	Company Submission
CSR	Clinical Study Report
C-SSRs	Columbia-Suicide Severity Rating Scale
DALY	Disability-adjusted life year
db	Decibels
DLQI	Dermatology Life Quality Index
DSU	Decision Support Unit
EAG	External assessment group
EAS	Efficacy analysis set
EBA	Eyebrow assessment
ELA	Eyelash assessment
EMA	European Medicines Agency
EQ-5D	EuroQol 5 dimensions
EQ-5D-5L	EuroQol 5 dimensions 5 level
EQ VAS	EuroQol Visual Analogue Scale
FAS	Full analysis set
FDA	United States Food and Drug Administration
GP	General Practitioner
HADS	Hospital Anxiety and Depression Scale

HADS-A	Hospital Anxiety and Depression Scale – Anxiety subscale
HADS-D	Hospital Anxiety and Depression Scale – Depression subscale
HR	Hazard ratio
HRQoL	Health-related quality of life
HST	Highly specialised technology
HTA	Health Technology Assessment
ICER	Incremental cost effectiveness ratio
IGA	Investigator Global Assessment (of change)
IQR	Inter-quartile range
IENFD	Intraepidermal nerve fibre density
JAK	Janus Kinase
JAK3	Janus Kinase 3
K-M	Kaplan-Meier
LCR	longitudinal concentration response
LOCF	last observation carried forward
LSM	Least squares mean
MHRA	Medicines and Healthcare products Regulatory Agency
MMF	Mycophenolate Mofetil
N/A	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NR	Not reported
ONS	Office for National Statistics
PAS	Patient Access Scheme
PGI-C	Patient's Global Impression of Change
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PUVA	Psoralen plus ultra violet A light therapy
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RR	Risk ratio
SAE	Serious adverse event
SALT	Severity of Alopecia Tool
SAS	Safety analysis set
SD	Standard deviation

SF-36	36-Item Short Form Health Survey
SF-6D	Short Form 6-Dimension health index
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
STAT	Signal transducer and activator of transcription
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TEC	Tyrosine kinase expressed in hepatocellular carcinoma
TM	transition matrix
TRAE	Treatment-related serious adverse event
TSD	Technical Support Document
TTO	Time-trade-off
UK	United Kingdom
US	United States
VBA	Visual basic application
WTP	Willingness to pay

1. EXECUTIVE SUMMARY

This External Assessment Group (EAG) report assesses ritlecitinib for treating severe alopecia areata (AA) in people aged 12 years and over. Severe AA is generally defined as having hair loss on the scalp of 50% or more. This executive summary provides a brief overview of the key issues identified by the external 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, and do not necessarily reflect the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Key issues identified by the EAG that impact on the incremental costs and quality-adjusted life years (QALYs) for ritlecitinib compared with best support care (BSC) are summarised in Table 1.

Table 1: Overview of the EAG's key issues

ID4007	Summary of issue	Report sections
Issue 1	The company has not provided a cost-effectiveness analysis for the AT/AU subgroup	3.1 & 5.3.4.1
Issue 2	ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	3.1 & 5.3.4.1
Issue 3	Assumption of no treatment waning based on limited long-term evidence	5.3.4.4
Issue 4	Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	5.3.4.5
Issue 5	Spontaneous remission applied when patients switch from ritlecitinib to BSC	5.3.4.6
Issue 6	Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	5.3.4.7
Issue 7	Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	5.3.4.10
Issue 8	Carer disutility based on a vignette for a carer of an adolescent with severe AA has been applied at all ages	5.3.4.11
Issue 9	Utility values are not age-adjusted	5.3.4.15

Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; BSC, best supportive care; EAG, External Assessment Group; EQ-5D, EuroQol 5 dimensions 5 level; ICER, incremental cost-effectiveness ratio; TTO, time trade off

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are:

- in the absence of the company providing an analysis based on the EQ-5D data from the trial, the EAG prefers to use EQ-5D data from the literature instead of those from the vignette study
- the EAG prefers to apply the utility values estimated from a vignette describing a carer of an adolescent patient with severe AA only to carers of patients aged 12 to 17 years in the model, whereas the company has applied it to all patients at all ages
- the company has assumed that spontaneous remission can occur after treatment discontinuation of ritlecitinib, but the EAG prefers to assume that any ritlecitinib patients having spontaneous remission are already included in the ritlecitinib responders
- the EAG prefers to use the average transition matrix from the second year of treatment to extrapolate beyond 2 years instead of the company's assumption of no treatment waning
- to estimate the efficacy during the second year, the EAG prefers to use data only from those on a 50 mg dose at the start of the second year of treatment, whereas the company includes data from those switching from 30 mg to 50 mg at the start of the second year
- the EAG prefers to assume a higher treatment discontinuation rate than the rate estimated by the company as the company's approach results in a mean time on treatment that is much higher than that observed when JAK inhibitors have been used long-term in other indications
- the EAG would prefer to include an age-adjustment for utilities to reflect the average decline in utilities with age in the general population but has been unable to implement this within the company's existing model structure.

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 health-related quality of life (HRQoL) for patients by reducing time spent with severe AA and increasing time spent with mild to moderate AA (main cause of QALY gain)
- reducing carer HRQoL burden associated with severe AA (small impact)
- increasing HRQoL losses associated with mild to moderate emergent adverse events (very small impact).

Overall, the technology is modelled to affect costs by:

- increasing costs for acquisition of ritlecitinib (main driver of incremental cost)
- reducing costs for managing AA by reducing time spent with severe AA and increasing time spent with mild AA and moderate AA which have lower resource use (small impact)
- increasing costs associated with mild to moderate adverse events (very small impact).
- increasing costs associated with monitoring ritlecitinib treatment (very small impact).

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

- the choice of utility values applied for different AA severities
- whether utility values are applied for carers of patients with severe AA
- whether spontaneous remission occurs after treatment discontinuation
- data applied for long-term extrapolation of treatment effect
- rate of treatment discontinuation.

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

The decision problem addressed in the company submission (CS) is broadly in line with the final NICE scope. The only key issues related to the decision problem were related to the handling of subgroups within the cost-effectiveness modelling and these are therefore described in section 1.5.

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

The key clinical evidence presented in the CS and that informs the economic analysis for ritlecitinib is from the ALLEGRO 2b/3 clinical trial. This compared several ritlecitinib dosing regimens with placebo, one of which was the anticipated dose for the marketing authorisation; 50 mg taken orally once per day. Other supporting clinical evidence comes from ALLEGRO-LT, which is a long-term open label study, and two phase 2a studies (ALLEGRO 2a proof of concept, ALLEGRO-2a safety study).

The EAG's key issue regarding the clinical effectiveness evidence is uncertainty over whether the proposed licensed dose of ritlecitinib (50 mg once daily) is effective over the long-term for patients with severe AA, including after treatment discontinuation. Further evidence from the ALLEGRO-LT study may elucidate this, although the EAG notes that, as the inclusion criteria allowed for those with milder AA (proportion of scalp hair loss ≥25%), it would be difficult to assess the long-term effects of ritlecitinib on those with severe AA (proportion of scalp hair loss ≥50%). The uncertainty regarding the long-term effectiveness of ritlecitinib in severe AA is discussed in cost-effectiveness Issues 3 and 4 and uncertainty regarding the rate of treatment discontinuation is discussed in cost-effectiveness Issue 6 (see Section 1.5).

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

The company's base case economic analysis compared ritlecitinib to BSC in people with severe AA aged 12 years and over. Subgroup analyses are also provided for adolescents (ages 12 to 17 years) and adults (aged 18 years and over). The model includes an interim treatment stopping rule at 24 weeks for those whose hair loss worsens on treatment, and a final stopping rule at 48 weeks for non-responders. Responders were defined as those with hair loss of less than 20% of their scalp (referred to as a SALT score of <20). The key inputs to the cost-effectiveness analysis are short-term effectiveness data and safety data from ALLEGRO 2b/3, long-term effectiveness data from ALLEGRO-LT, utility values from a vignette study, and resource use estimates from an expert elicitation study.

The key cost-effectiveness issues identified by the EAG which are associated with the greatest decision uncertainty are summarised in this section. Several other issues, which are discussed in section 5.3.4, are not described in detail here as they have a smaller impact on the ICER. These related to adverse events, resource use for different severities of AA and all cause-mortality in the first year.

Issue 1 The company has not provided a cost-effectiveness analysis for the AT/AU subgroup

Report section	5.3.4.1 Age and severity subgroups (Key issues 1 and 2)
Description of issue and why the EAG has identified it as important	The subgroup analysis for the ALLEGRO 2b/3 study indicates that the probability of patients achieving a treatment response is lower for the AT/AU subgroup than for other patients with severe AA. However, the company have not provided a cost-effectiveness analysis for this subgroup despite it being pre-specified in the scope as a subgroup of interest.
What alternative approach has the EAG suggested?	The EAG would prefer the company to provide a subgroup analysis for the cost-effectiveness model that is populated with data that reflect expected outcomes for the AT/AU subgroup.
What is the expected effect on the cost-effectiveness estimates?	The EAG have explored the potential impact of incorporating lower response rates for the AT/AU subgroup. Whilst the EAG were unable to update all model inputs to reflect expected outcomes in this group, this exploratory analysis suggests that the ICERs are likely to be higher for this subgroup when applying the company base case assumptions (£15,207 per QALY versus £13,179 per QALY) but lower when applying the EAG's preferred base case assumptions (£60,293 per QALY vs £66,674 per QALY).
What additional evidence or analyses might help to resolve this key issue?	The company should provide a cost-effectiveness model populated with appropriate inputs for the AT/AU subgroup as the EAG's current exploratory analysis relies on several assumptions.

Issue 2 ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents

Report section	5.3.4.1 Age and severity subgroups (Key issues 1 and 2)
Description of issue and why the EAG has identified it as important	Several model inputs differ for adolescents and adults in either the company's base case or under the EAG's preferred assumptions. Key differences include the short-term response rates, the long-term transition matrices, and all-cause mortality. In addition, the EAG prefers to apply carer disutility only during adolescent years, whilst the company's base case applies the same carer disutility for all ages (see Issue 8). The company's base case uses average baseline characteristics across the ALLEGRO 2b/3 population rather than modelling outcomes separately for adolescent patients starting treatment and adults starting treatment. However, the CS also presents subgroup results by starting age (12 to 17 years and age ≥18 years).
What alternative approach has the EAG suggested?	The EAG prefers to estimate model outcomes separately for adults and adolescents to accurately capture the expected outcomes in these groups. A weighted average can then be used to generate an accurate ICER for the whole population covered by the scope.
What is the expected effect on the cost-effectiveness estimates?	The weighted average approach has a small impact on the results in the company's base case, where it increases the ICER from £13,179 per QALY to £13,235 per QALY. However, it has a much larger impact when using the EAG's preferred assumptions, where it increases the ICER from £60,735 per QALY to £66,674 per QALY. This is because the EAG's preferred approach has different assumptions for carer disutility according to patient age (see Issue 8).
What additional evidence or analyses might help to resolve this key issue?	No additional analyses are required.

Issue 3 Limited evidence to support assumption of no treatment waning in the long-term

Report section	5.3.4.3 Clinical evidence used when implementing the interim stopping rule
Description of issue and why the EAG has identified it as important	The company's base case assumes that patients who have responded to treatment at 48 weeks and who have remained on treatment for a further year will have a stable SALT score going forwards until they discontinue treatment. The EAG considers that this assumption is difficult to verify given the limited duration of follow-up available across both the ALLEGRO 2b/3 and ALLEGRO-LT studies.
What alternative approach has the EAG suggested?	The EAG prefers to use the average transition matrices from the second year of treatment to estimate long-term outcomes in patients remaining on treatment beyond 2 years. The EAG also explores the impact of using the last transition matrix, representing months 21 to 24 of treatment, repeatedly going forward (the company refers to this as a last observation carried forward [LOCF] approach).
What is the expected effect on the cost-effectiveness estimates?	The impact of using the average transition matrices is to increase the company's ICER from £13,179 per QALY to £15,676 per QALY. The average transition matrices are included in the EAG's preferred base case. The impact of applying the LOCF approach as an alternative in the EAG's preferred base case is to increase the ICER from £66,674 per QALY to £77,806 per QALY.
What additional evidence or analyses might help to resolve this key issue?	Further follow-up data from ALLEGRO-LT may help identify if the company's assumption of no treatment waning is valid and will also provide further data points to inform the extrapolation when using either of the average transition matrix or LOCF approaches.

Issue 4 Long-term extrapolation based on data from patients receiving doses other than anticipated licensed dose

Report section	05.3.4.5 Generalisability of the long-term data to the 50 mg dose (Key Issue 4)
Description of issue and why the EAG has identified it as important	The company has used data from the ALLEGRO-LT study to estimate the long-term efficacy of ritlecitinib. However, some patients starting ALLEGRO-LT were previously on a 30 mg dose of ritlecitinib. Using data from these patients may capture an improvement in SALT scores in response to a dose increase rather than the expected outcomes for patients remaining on a stable 50 mg dose.
What alternative approach has the EAG suggested?	The EAG prefers to use data only from patients who have previously received a 50 mg dose of ritlecitinib to estimate the long-term transition matrices. Therefore, the EAG excluded data from patients who had switched from a 30 mg dose to a 50 mg dose at the start of ALLEGRO-LT. Given that patients were only able to contribute to the transition matrices after receiving 48 weeks of treatment, patients who had received a loading dose before switching to a 50 mg maintenance dose were considered sufficiently similar to be pooled with those who had a 50 mg dose from the start of treatment.
What is the expected effect on the cost-effectiveness estimates?	This had a marginal impact on the ICER but increased the company's base case ICER from £13,179 per QALY to £13,294 per QALY.
What additional evidence or analyses might help to resolve this key issue?	The EAG is not aware of any additional analyses that would further resolve this issue.

Issue 5 Spontaneous remission applied when patients switch from ritlecitinib to BSC

Report section	5.3.4.6 Spontaneous remission (Key issue 5)
Description of issue and why the EAG has identified it as important	A small proportion of patients in the placebo arm of the ALLEGRO 2b/3 study reached a SALT score ≤10 at 24 weeks. The company interpreted these as being cases of spontaneous remission and assumed that a stable proportion of patients on BSC would be in spontaneous remission at any time during the economic analysis. This was implemented by having a fixed proportion of patients in the BSC arm being in spontaneous remission from 24 weeks and assuming that the same proportion of patients would experience spontaneous remission when discontinuing ritlecitinib treatment. The EAG believes that any cases of spontaneous remission in the ritlecitinib arm at 24 weeks will have been classified as treatment responders. Therefore, these cases are already accounted for in the model and no additional cases of spontaneous remission need to be accounted for in the model when patients discontinue ritlecitinib treatment.
What alternative approach has the EAG suggested?	The EAG prefers to exclude spontaneous remissions when patients discontinue ritlecitinib treatment.
What is the expected effect on the cost-effectiveness estimates?	This has a moderate impact on the ICER increasing the company's base case ICER from £13,179 per QALY to £14,578 per QALY when applied as an isolated change.
What additional evidence or analyses might help to resolve this key issue?	The EAG is not aware of any additional evidence or analyses that would further address this issue.

Issue 6 Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment

Report section	5.3.4.7 Time to discontinuation (Key Issue 6)
Description of issue and why the EAG has identified it as important	The company analysis assumes that patients who experience a SALT score >20 after 48 weeks will discontinue ritlecitinib treatment. In order to avoid double counting treatment discontinuation due to loss of response and treatment discontinuations due to other reasons, the company has estimated a time to discontinuation curve using a dataset which excluded patients who experienced a SALT score >20 after 48 weeks of treatment. However, the EAG believes that this may have biased the estimate of time to discontinuation (for reasons other than lack of response) because the company excluded patients from the analysis at all time points, rather than censoring them at the time their SALT score reached >20. In addition, the EAG believes the discontinuation rate in the company's base case analysis results in a mean time on treatment that is too high in comparison to the mean duration on treatment when JAK inhibitors are used in other indications where longer follow-up is available. A scenario is provided by the company in which patients are not excluded from the time to discontinuation analysis for having a SALT score >20 after 48 weeks, however, the duration of follow-up for this scenario is inconsistent with the other scenarios and the EAG believes it lacks face validity and may have been calculated over the wrong time frame.
What alternative approach has the EAG suggested?	The EAG has applied a discontinuation rate that is double that applied in the company's base case in order to achieve a mean time on treatment that is in the range observed when JAK inhibitors are used in other indications.
What is the expected effect on the cost-effectiveness estimates?	Applying a higher treatment discontinuation rate results in a shorter duration of time on treatment, reducing both incremental costs and QALYs gained, but overall it increases the ICER (from £13,179 per QALY to £14,217 per QALY) because the proportionate reduction in QALYs gained is greater.
What additional evidence or analyses might help to resolve this key issue?	The EAG would prefer the company to repeat their survival analysis censoring patients at the time they stop responding rather than excluding patients who stop responding at any future time point. The EAG would also like further clarification from the company as to why a longer duration of follow-up is available in one of the four scenarios presented. The analysis for this scenario should be corrected if this reflects an error in the company's approach. Advice could be sought from clinical experts regarding the likely duration of treatment for patients who continue to respond and whether the company's estimates of predicted time on treatment are plausible.

Issue 7 Utilities for severity-based health states

Report section	5.3.4.10 Utility values for patient health states (Key issue 7)
Description of issue and why the EAG has identified it as important	The company has used utility values estimated from a vignette study in the base case instead of using utility values based on EQ-5D directly measured in patients from their pivotal study (ALLEGRO 2b/3) or EQ-5D estimates by severity of AA available from the literature. The company argues that the EQ-5D is not appropriate in this case citing high baseline EQ-5D values in ALLEGRO 2b/3, which they claim demonstrate a ceiling effect, and minimal changes from baseline or differences between trial arms, which they claim demonstrates a lack of responsivity. They also claim that the EQ-5D lacks content validity in severe AA based on a systematic literature review and qualitative research with patient advocacy group members and clinical experts. The EAG believes that the high baseline EQ-5D scores and lack of statistically significant changes in EQ-5D during the ALLEGRO 2b/3 study may be due to the selective group recruited into this study and the short duration of follow-up. The EQ-5D has demonstrated construct validity and responsiveness in other skin diseases and estimates from the literature (Bewley *et al.*) suggest that it has construct validity in AA. The EAG has concerns regarding the face validity of the final vignettes in comparison to the quantitative data used to develop them and therefore believes they should be treated with caution.
What alternative approach has the EAG suggested?	The EAG acknowledges that there may be some underestimation of QALY gains when using the utility values obtained directly from ALLEGRO 2b/3, due to the selective population recruited and limited follow-up. However, the EAG would have preferred to see results based on the EQ-5D data from ALLEGRO 2b/3 as the company's base case, as this is in keeping with the NICE reference case approach, with alternative sources explored as scenario analyses. In the absence of an analysis from the company using the EQ-5D data from the trial, the EAG prefers to populate the economic model using the estimates from the literature that provide EQ-5D scores by AA severity (Bewley *et al.*) instead of data from the vignette study.
What is the expected effect on the cost-effectiveness estimates?	The impact on the ICER of using the EQ-5D estimates from the literature is large, increasing the ICER in the company's base case from £13,179 per QALY to £33,945 per QALY.
What additional evidence or analyses might help to resolve this key issue?	The company should provide a cost-effectiveness analysis that uses the EQ-5D data from the trial as a reference case scenario with scenario analyses exploring alternative sources including the literature-based estimates from Bewley et al.

Issue 8 Carer disutility based on a vignette for a carer of an adolescent with severe AA has been applied at all ages

Report section	5.3.4.11 Utility values for carer HRQoL decrements (Key issue 8)
Description of issue and why the EAG has identified it as important	The company's vignette study included a vignette for a carer of an adolescent with severe AA. Members of the population providing time-trade-off valuations were specifically asked to imagine that they were the caregiver to a family member aged 12-17 years. However, in the company's base case analysis, the disutility estimated for this vignette has been applied to all patients regardless of their age.
What alternative approach has the EAG suggested?	The company provided a scenario in response to clarification in which the caregiver disutility was restricted to adolescent patients in the years when they are aged 12 to 17 years old. The EAG prefers to use this approach in their base case analysis.
What is the expected effect on the cost-effectiveness estimates?	This has a moderate impact on the QALYs gained for adolescents as the disutility is applied for fewer years. It has a larger impact on the QALYs gained for adults where the caregiver disutility was previously being applied and is no longer being applied. This change when applied in isolation increases the company's base case ICER from £13,179 per QALY to £14,192 per QALY. The EAG also explored a scenario excluding the carer disutility altogether and this increased the ICER for the EAG's preferred base case from £66,674 per QALY to £68,960 per QALY.
What additional evidence or analyses might help to resolve this key issue?	The EAG is also not convinced that the caregiver disutility provided by the vignettes is accurate as the company assumed that caregivers of adolescents with mild to moderate AA would have no disutility. Ideally the company should provide utility scores measured in caregivers before and after a response to treatment. Alternatively, the company could provide estimates of caregiver disutility for mild and moderate AA rather than assuming that all caregiver disutility resolves when the adolescent's SALT score reaches below 50.

Issue 9 Utility values are not age-adjusted

Report section	5.3.4.15 Utilities not age-adjusted (Key issue 9)
Description of issue and why the EAG has identified it as important	The EAG notes that utility values are constant across time for patients in a particular health state and are therefore not adjusted to reflect declining utilities in the general population with increasing age.
What alternative approach has the EAG suggested?	The EAG considers that the AA specific utility weights should have been applied as multipliers to the expected utility values in the general population.
What is the expected effect on the cost-effectiveness estimates?	The EAG has been unable to implement a scenario analysis which incorporates the age-adjustment accurately and fully to their satisfaction within the company's existing model structure. However, the EAG believes that the lack of age-adjustment has led to an overestimation of the incremental QALYs gained and therefore an underestimation of the ICER.
What additional evidence or analyses might help to resolve this key issue?	The company should provide a model that includes an age-adjustment for utility values.

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

No other key issues were identified by the EAG.

1.7 Summary of EAG's preferred assumptions and resulting ICER

Table 2 summarises the results of the EAG's exploratory analyses, including the EAG's preferred base. The results in Table 2 are for the whole population covered by the decision problem (people aged 12 years and over) except where they are described as applying to a specific age subgroup (adults aged ≥18 years or adolescents aged 12 to 17 years) or to the AT/AU subgroup (aged ≥12 years). The EAG's preferred base case uses a weighted average approach to estimate the ICER for the whole population using outcomes estimated for age-specific subgroups (for adults aged ≥18 year and adolescents aged 12 to 17 years). The results in Table 2 are deterministic unless otherwise stated. Modelling errors identified and corrected by the EAG are described in Section 5.4.2.1. For further details of the exploratory and sensitivity analyses done by the EAG, see Sections 5.4.2.2. to 5.4.2.14.

It can be seen from Table 2 that for the EAG's preferred base case, estimating the ICER for the whole population using a weighted average of the outcomes for adults (aged ≥18 year) and adolescents (aged 12 to 17 years) produces a higher ICER than using average baseline starting characteristics (£66,674 per QALY versus £60,735 per QALY). In addition, the probabilistic ICER for the EAG's preferred approach is higher than the deterministic ICER (£89,888 per QALY versus £66,674 per QALY) due to the continuity corrections required to handle missing observations in the long-term transition matrices, which is only applied in the probabilistic analysis.

Table 2: Summary of results of EAG's exploratory analyses, (deterministic except where stated otherwise)

Scenario	Incremental cost	Incremental QALYs	ICER (change from company base case)
Company base case (Deterministic)	[redacted]	[redacted]	£13,179
EAG's corrected company base case: correcting implementation errors in the company's economic model [included in all subsequent rows]	[redacted]	[redacted]	£13,179 (£0)
EA1: Using pooled counts from the second year to estimate the 3-month transition matrix applied from 2 years onwards	[redacted]	[redacted]	£15,676 (+£2497)
EA2: Using only patients who were on the 50 mg dose to inform the long-term matrices	[redacted]	[redacted]	£13,294 (+£115)
EA3: Assuming no spontaneous remission in the ritlecitinib arm	[redacted]	[redacted]	£14,578 (+£1399)
EA4: Assuming double the hazard of discontinuation applied in the company's base case	[redacted]	[redacted]	£14,217 (+£1038)
EA5: Allowing mortality in the first 48 weeks of the model	[redacted]	[redacted]	£13,139 (-£40)
EA6: Using the utility values reported by Bewley *et al.*	[redacted]	[redacted]	£33,945 (+£20,766)
EA7: Carer disutility applies only during adolescent years	[redacted]	[redacted]	£14,192 (+£1013)
EA8: Assuming the same psychological support for ritlecitinib as for BSC	[redacted]	[redacted]	£13,170 (-£9)
EA9: Assuming TEAEs are managed in primary care	[redacted]	[redacted]	£12,976 (-£203)
EAG's preferred base case applying analyses EA1 to EA9 - deterministic	[redacted]	[redacted]	£60,735 (+£47,556)
EAG's preferred base case applying analyses EA1 to EA9 - deterministic (adults)	[redacted]	[redacted]	£69,246 (+£56,067)
EAG's preferred base case applying analyses EA1 to EA9 - deterministic (adolescents)	[redacted]	[redacted]	£55,349 (+£42,170)
EAG's preferred base case applying analyses EA1 to EA9 - deterministic (weighted average of the two subgroups)	[redacted]	[redacted]	£66,674 (+£53,495)
EAG's preferred base case applying analyses EA1 to EA9 - probabilistic (weighted average of the two subgroups)	[redacted]	[redacted]	£82,152 (+£68,973)

2 BACKGROUND

This report provides a review of the evidence submitted by the company (Pfizer) in support of ritlecitinib for treating patients with alopecia areata. It includes evidence presented within the company's submission (CS) received on the 11th January 2023[1] and the responses to clarification questions provided by the company on the 17th February 2023.[2]

This chapter presents a brief summary and critique of the company's description of the disease and the current treatment pathway for alopecia areata in England.

2.1 Critique of company's description of underlying health problem

Alopecia areata (AA) is a chronic inflammatory autoimmune condition that affects the hair follicles and leads to a sudden onset of hair loss, with no scarring or permanent damage to the hair follicles.[3] Any hair-bearing skin can be affected, such as the scalp, beard, and, less commonly, eyebrows, eyelashes, body and limbs, usually presenting as small, round or oval patches of baldness.[3] In rare cases, the whole scalp (alopecia totalis [AT]) or the entire body and scalp (alopecia universalis [AU]) may be affected.[3] Nail changes (e.g., pitting, thinning, thickening, nail loss) may occur in 10 to 50% of people with alopecia areata, usually in those with more severe disease.[4]

Alopecia areata is diagnosed based on pattern of hair loss, 'exclamation mark' hairs (short, broken hairs that taper proximally) and a positive pull test.[4] The main symptom of alopecia areata is hair loss; however, patients may also experience paraesthesia with mild to moderate pruritus, tenderness, burning sensation or pain prior to the hair loss.[4] The exact cause of the condition is unknown, although contributory factors may include genetics, autoimmune reactions, stressful life events, and neurogenic changes.[3] Males and females are affected equally, and alopecia areata can present in people of any age, with a higher incidence among children and young adults.[4, 5]

Alopecia areata can have a psychosocial impact on those affected, including altered body image, reduced self-esteem, social withdrawal and increased risk of anxiety and depression, which in some cases can affect occupational activity.[3, 6] Prevalence is estimated at 0.58% of adults in the UK,[5] and the prognosis is unpredictable, although worse prognosis is associated with a large surface area, long disease duration and associated nail abnormalities.[7] Evidence suggests that between 50% and 80% of people with milder alopecia areata (characterised by limited patches of hair loss of less than a year's duration) can experience spontaneous remission within one year,[3, 8, 9] although those with more extensive hair loss rarely experience spontaneous or sustained remission.[5] Most patients in remission will experience further episodes of alopecia areata and it is estimated that between 5% and 30% of patients with patchy hair loss will eventually progress to alopecia totalis.[8, 10]

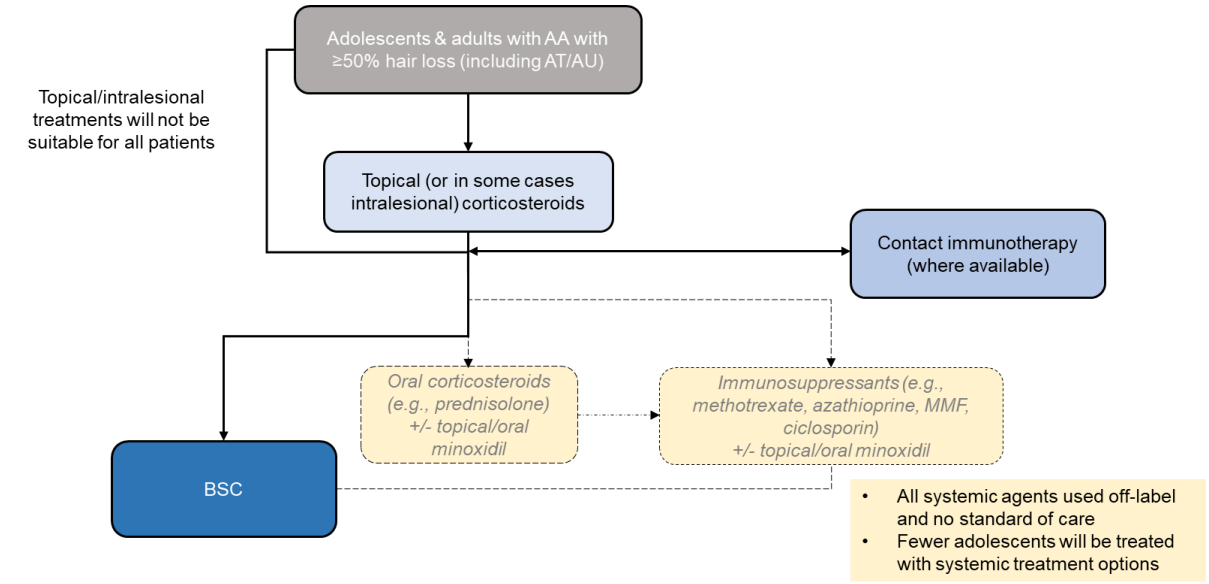
Strategies for clinical management of alopecia areata in England are based on the severity of hair loss, and begin in primary care. If there is evidence of regrowth or if there is less than 50% hair loss, watchful waiting is advised, alongside advice about sun protection and cosmetic options.[3] For more severe alopecia areata with no regrowth and more than 50% hair loss, the main primary care treatment option is topical corticosteroids.[3] If the condition does not respond to topical corticosteroids, the patient can be referred to specialist dermatology management, where treatments can include unlicensed topical immunotherapy and off-label intralesional corticosteroids, oral corticosteroids, psoralen plus ultra violet A light therapy (PUVA), and immunosuppressive drugs such as oral ciclosporin or methotrexate.[3] Currently licensed, unlicensed and off-label treatment options may induce hair growth, but none alter the course of the disease or offer a cure.

2.2 Critique of company's overview of current service provision

The CS provides a comprehensive overview of service provision.[1] The CS states correctly that no licensed systemic treatment is available for adults and adolescents with severe AA. Therapeutic options consist of topical corticosteroids as a first-line treatment, followed by contact immunotherapy where available, followed by off-label oral corticosteroids or immunosuppressants, in conjunction with or followed by best supportive care (BSC), which typically consists of non-pharmacological management largely consisting of prosthetics to mask hair loss, such as wigs, false eyelashes, and semi-permanent make-up (see Figure 1). The CS describes oral corticosteroids and immunosuppressant therapies as the most commonly used off-label systemic treatment options, with patients often started on oral corticosteroids in the short term with an immunosuppressant treatment started as patients are tapered off oral corticosteroids.

Clinical advice received by the EAG indicated broad agreement with the clinical pathway outlined in the CS (Figure 1), although one clinician reported using more aggressive combination immunosuppressants early in the disease course and reported beneficial results. All clinicians consulted by the EAG agreed with prompt treatment for a greater chance of success and none were keen to engage in watchful waiting prior to commencing treatment, although clinicians advised that some patients would opt for no pharmacological treatment. Clinicians reported using contact immunotherapy (diphencyprone), oral steroids and systemic immunosuppressants including methotrexate, ciclosporin and azathioprine, and two clinicians had used janus kinase (JAK) inhibitors. Despite the reporting of a wide variety of off-label pharmacological treatments, clinicians had different preferences and used treatments in different ways (e.g., one clinician reported using combination immunosuppressants). None of the clinicians advising the EAG reported using minoxidil or mycophenolate mofetil (MMF). The EAG's clinical advisors reported dissatisfaction with current treatment options and a desire for an effective systemic treatment option.

Figure 1: AA pharmacological treatment pathway (reproduced from CS, Figure 8)*

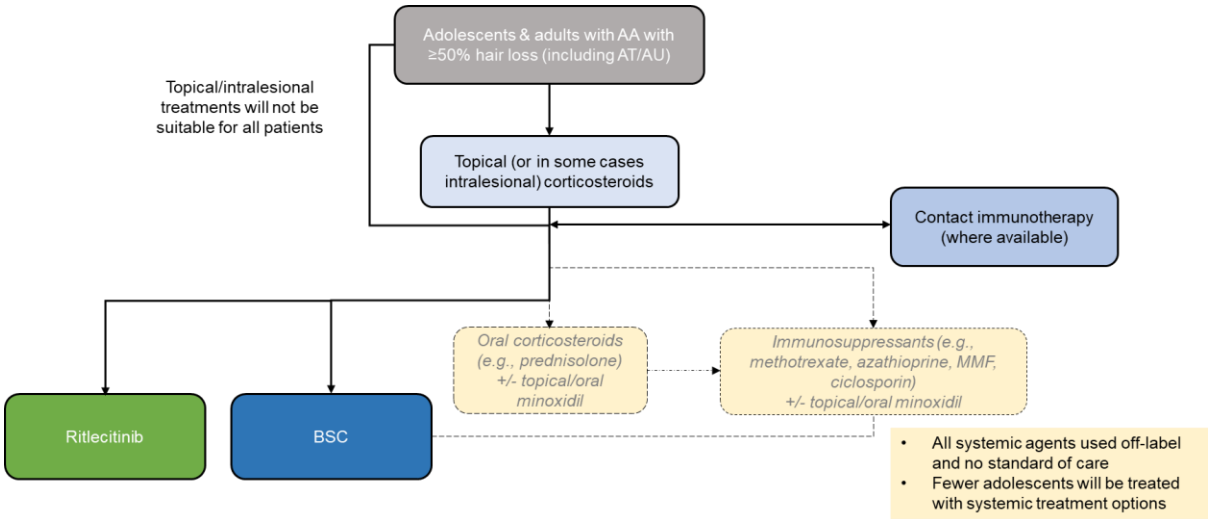


AA - alopecia areata; AT - alopecia totalis; AU - alopecia universalis; BSC - best supportive care; MMF - mycophenolate mofetil
**75% of dermatologists with a specialist interest in hair disorders agreed with an earlier version of the diagram and then following three further interviews the diagram was optimised and finalised.*

2.3 Critique of company's proposed positioning of ritlecitinib in the treatment pathway

The company's proposed positioning of ritlecitinib is shown in Figure 2. Ritlecitinib is proposed as a systemic treatment option for severe AA, either after topical corticosteroids and contact immunotherapy or when these topic treatments are not suitable or available. Clinical advice received by the EAG suggests that this positioning is aligned with how clinicians would want to use ritlecitinib in clinical practice.

Figure 2: AA pharmacological treatment pathway with ritlecitinib (reproduced from CS, Figure 10)



AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; BSC, best supportive care; MMF, mycophenolate mofetil

3 CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

3.1 Population

The population addressed in the CS is people aged 12 years and over with severe AA. This is consistent with the population specified in the NICE final scope and with the anticipated marketing authorisation. In general, severe AA is defined in the CS in terms of the Severity of Alopecia Tool (SALT) score, which is a measure of the percentage of hair loss on the scalp (and thus a higher score indicates greater severity of AA). In the CS, the company defines severe AA as patients with ≥50% of scalp hair loss (CS, p10), equivalent to a SALT score of ≥50, which includes those with AT or AU. The population enrolled in the pivotal clinical study (ALLEGRO 2b/3) and the population reflected in the company's base case economic analysis were both consistent with the population specified in the NICE final scope.

The NICE final scope specifies that subgroups based on severity and type of AA (e.g., AT / AU) will be considered if evidence allows. The NICE final scope does not specify that subgroups based on age will be considered. The CS provides clinical subgroup analyses for patients with AT/AU and by age (12-17 years, ≥18 years, 18-44 years, 45-64 years and ≥65 years) for the key clinical outcomes (SALT≤20 and SALT≤10) from the pivotal comparative study (ALLEGRO 2b/3). The company's base case cost-effectiveness analysis is for the whole population aged 12 years and over. The CS also presents subgroup cost-effectiveness results for adults (aged over 18 years) and adolescents (aged 12 to 18 years). The EAG notes that the company's base case model reflects the overall population (aged 12 years and over) using average baseline characteristics rather than taking a weighted average of the outcomes expected for each age subgroup (adolescents and adults); this issue is discussed further in Section 5.3.4.1. In the cost-effectiveness section, no subgroup analyses are provided for patients with AT/AU. The EAG would have preferred to see a cost-effectiveness analysis for this subgroup. This issue is further discussed in Section 5.3.4.1.

3.2 Intervention

The intervention is ritlecitinib which is an orally bioavailable small molecule that inhibits JAK3 and the TEC kinase family with very high selectivity over the other three JAK isoforms, as well as over the broader human kinome. In AA, there is over-activation of the JAK/STAT pathway and this results in damage to hair follicles. Ritlecitinib works by inhibiting the JAK/STAT signalling pathways (see CS, Table 3 for further details of the mechanism of action).

The dosing regimen assumed in the CS is 50 mg of ritlecitinib taken orally once daily. The EAG notes that whilst several other dosing regimens have been studied in clinical trials, the 50 mg once daily dose, which is the focus on the submission, is in keeping with the anticipated marketing authorisation. The

key efficacy data presented in the submission are based on the 50 mg arm of the ALLEGRO 2b/3 study. However, data from other study arms of ALLEGRO 2b/3 and other studies (ALLEGRO 2a proof of concept, ALLEGRO-LT and ALLEGRO-2a safety study), which used different ritlecitinib dosing regimens, have contributed to the efficacy and safety data presented in the CS. Data from the ALLEGRO-LT study, including data from groups which received different dosing regimens, have also contributed to estimating long-term outcomes in the cost-effectiveness analysis.

The list price is £[redacted] per pack of 30 capsules of ritlecitinib 50mg giving an annual cost of £[redacted] for the anticipated dose of 50mg per day. The economic analysis includes a patient access scheme (PAS) discount which reduces the pack price to £[redacted] and the annual cost to £[redacted].

The company's base case economic analysis assumes that patients whose SALT score worsens during the first 24 weeks will stop treatment, patients who do not achieve a SALT score ≤20 at 48 weeks will discontinue treatment, and patients who have a SALT score >20 after 48 weeks will also discontinue (CS, Table 33).

3.3 Comparators

The comparator in the CS is BSC, which the company considers to be non-pharmacological management of severe AA, mainly consisting of prosthetics and cosmetics to mask hair loss. The NICE final scope specified the comparator as "*established clinical management without ritlecitinib*". Although the scope describes a range of treatments used for severe AA, the company argues that none of these are suitable comparators. The CS describes how contact immunotherapy would be typically used prior to systemic treatments, but it is not widely available in the UK (see Section 2.2). The company reports that the clinical experts they interviewed (n=3) all agreed that the most relevant comparator for ritlecitinib in the UK is BSC with no pharmacological treatment. The CS states that BSC in the UK aligns with the placebo arm of the key comparative study (ALLEGRO 2b/3) in which placebo patients were permitted to use non-pharmacological management such as wigs.

The EAG accepts that there is significant variation in current NHS clinical practice and many of the treatments currently used are either off-label/unlicensed for severe AA or are only available at a limited number of sites (e.g., contact immunotherapy). Based on this, the EAG accepts that BSC is the only comparator consistently available within current NHS clinical practice.

3.4 Outcomes

Clinical outcomes listed in the final NICE scope[11] include:

- Severity of AA

- Percentage of area affected by hair loss
- Adverse effects of treatment
- Health-related quality of life (HRQoL).

All outcomes defined in the final NICE scope[11] were included in the CS,[1] although percentage of area affected by hair loss was not reported at each timepoint, and was instead reported at baseline and then as either proportions of people meeting thresholds or in terms of change from baseline.

The economic analysis estimates the incremental costs and incremental quality-adjusted life-years (QALYs) over a lifetime horizon (discounted at 3.5% per annum) to provide an incremental cost-effectiveness ratio (ICER) for ritlecitinib versus BSC. Costs are assessed from an NHS and Personal Social Services (PSS) perspective in the base case, although a scenario including societal costs is also provided. QALYs are those accrued by patients but in the company's base case analysis the QALYs are adjusted for the disutility experienced by carers.

3.5 Special considerations

The CS highlights that AA is more common in some groups (females, people of Asian ethnicity, people from deprived economic backgrounds and people diagnosed with Down Syndrome) and that the increased prevalence in those from deprived backgrounds is problematic given that non-pharmacological management often involves significant out of pocket expenses for patients. The British Association of Dermatologists (BAD)[12] commented that restriction of treatment to tertiary centres may lead to geographic inequalities and for this reason, in the case of a positive recommendation, it would support ritlecitinib being offered across secondary care dermatology centres. They noted that this would also address current geographic inequalities in access to wig provision. The BAD also noted that beard hair loss may have particular significance for people of certain faiths and that hair loss during adolescence can have a profound psychological impact.[12] The EAG also notes that ritlecitinib is not recommended in people who are [redacted]. Therefore, the availability of ritlecitinib would not widen the therapeutic options available to these groups for which there are already fewer treatment options.

3.6 Other relevant factors

The CS states that ritlecitinib does not meet the criteria for severity weighting.

.

Table 3: The decision problem (reproduced from CS, Table 3 with minor amendments and comments from the EAG)

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale if different from NICE scope	EAG comments
Population	People aged 12 years and over with severe AA	As per scope.	N/A
Intervention	Ritlecitinib	Ritlecitinib 50mg once daily dose. This is the dose proposed for registration for the treatment of severe AA.	Whilst the main efficacy and safety data presented in the CS relate to the 50mg arm of the ALLEGRO 2b/3 study, some of the data presented in the CS and contributing the economic model are based on patients who received a different dosing regimen either in a different arm of ALLEGRO 2b/3 study or in another study.
Comparator(s)	Established clinical management without ritlecitinib.	The comparator is BSC as BSC is established clinical management	Systemic immunosuppressants and contact immunotherapy are used as off-label treatments in some patients but there is considerable variation in current NHS practice. Therefore, the EAG accepts that BSC is the only comparator consistently available within current NHS clinical practice.
Outcomes	The outcome measures to be considered include: • severity of alopecia areata • percentage of area affected by hair loss • adverse effects of treatment • health-related quality of life	As per scope.	As per scope, although percentage of area affected by hair loss was not reported at each timepoint, and was instead reported at baseline and then either proportions of people meeting thresholds or in terms of change from baseline.
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	As per scope. Consideration has also been given to indirect costs associated with absenteeism & presenteeism which are included as a scenario in the economic model.	The company's base case analysis uses an NHS and PSS perspective which is consistent with the NICE reference case. The EAG considers that the indirect costs included in the company's scenario analysis are outside of the NICE reference case and this scenario should not be considered relevant to decision making.

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale if different from NICE scope	EAG comments
	differences in costs or outcomes between the technologies being compared. Costs will be considered from a NHS and PSS perspective.		
Subgroups to be considered	• If evidence allows, subgroups based on severity and type of alopecia areata (e.g., alopecia totalis (AT) and alopecia universalis (AU)) will be considered.	The differences between each ritlecitinib group and placebo in the proportion of response based on SALT ≤20 at Week 24 were consistent across most pre-specified subgroups (age, BMI, weight, gender, race, region, severity of disease, duration since diagnosis, duration of current episode, prior pharmacological treatment for AA) for all doses. Furthermore, no subgroups have been considered in the economic analyses as N values are too small to draw any conclusions beyond consistency. Outcomes in the trial population split by age (adolescent and adult populations), previous treatment and race are provided in Appendix E.	The company has presented key clinical outcomes for the AT/AU subgroup but no economic analysis is provided for this subgroup. The company has provided subgroup analysis by age (12 to 17 years and ≥18 years) for both clinical and cost-effectiveness outcomes but their base case analysis is for all patients aged 12 years and over and uses average baseline characteristics. The EAG considers that the cost-effectiveness is likely to differ for adults and adolescents and therefore would support estimating a combined ICER using a weighted average of outcomes across these age subgroups.
Special considerations including issues related to equity or equality	No consideration highlighted in the final scope.	There are inequalities in the characteristics of people diagnosed with AA (females, people of Asian ethnicity, people from deprived economic backgrounds and people diagnosed with Down Syndrome are more likely to be diagnosed with AA.) Research suggests there are significant out of pocket expenses associated with an individual patient's management of AA. Considering that research also shows that AA follows an inverse social gradient and therefore, may pose an issue related to equity.	The EAG agrees that AA is more common in some groups and that the impact of hair loss may vary between individuals depending on their characteristics. The EAG notes that the submission from the British Association of Dermatologists[12] highlighted that restricting provision to tertiary centres may lead to geographic inequalities and for this reason it would be preferable for treatment to be available at all secondary care dermatology services if a positive recommendation were made.

Abbreviations: AA, alopecia areata; AT, alopecia totalis; AU, alopecia universalis; BSC, best supportive care; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PSS, personal social services

4 CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the clinical effectiveness evidence contained within the CS[1] for ritlecitinib for the treatment of severe AA in people 12 years and over. Section 4.1 provides a critique of the company's systematic review of clinical and safety evidence. Section 4.2 provides a summary of the clinical effectiveness and safety results, together with a critique of the included studies. Sections 4.3 to 4.5 of the template (relating to indirect comparisons and additional work undertaken by the EAG) are not applicable. Section 4.6 provides the conclusions of the clinical effectiveness section.

4.1 Critique of the methods of review(s)

The company undertook a systematic literature review (SLR) to identify all clinical evidence regarding the efficacy and safety of ritlecitinib and comparator treatments in patients with AA. The methods for the company's SLR of clinical evidence are detailed in CS Appendix D.[13]

4.1.1 Searches

The company performed one clinical effectiveness search to identify all clinical effectiveness and safety studies of ritlecitinib or comparator treatments of patients with alopecia areata.

The company searched several electronic bibliographic databases in October 2021 (original review) followed by an update in September 2022 (Appendix D.1 Identification and selection of relevant studies): MEDLINE [via Embase.com], EMBASE [via Embase.com], Cochrane Database of Systematic Reviews [via Wiley], Cochrane Central Register of Controlled Trials [via Wiley], Database of Abstracts of Reviews of Effects [via CRD], Health Technology Assessment database [via CRD] and NHS Economic Evaluation Database [via CRD]. The company also identified further studies through reference tracking of the bibliographies of systematic reviews and meta-analysis.

The company searched several key conference abstract and society presentation websites from 2020 to 2022: American Academy of Dermatology, Academy of Managed Care Pharmacy, European Association of Dermatology and Venereology, the Professional Society for Health Economics and Outcomes Research, and Society of Investigative Dermatology, British Association of Dermatologists, British Society of Investigative Dermatology, World Congress for Hair Research, American Hair research Society, and European Hair Research Society.

The company searched one clinical trials registry (clinicaltrials.gov) in November 2021 and September 2022 but not the WHO International Clinical Trials Registry Platform (ICTRP). Banno, Tsujimoto & Kataoka (2020)[14] compared the coverage of these two trials registries and CENTRAL and concluded

that clinicaltrials.gov and ICTRP sources should be searched together with CENTRAL to identify unpublished trials.

The EAG found no errors in the search strategies and considers that the search is comprehensive. However, it was unclear to the EAG how the company identified the comparator treatments, as the list of intervention and comparator terms was extensive and covered specialist treatments such as intralesional corticosteroids, oral corticosteroids, topical immunotherapy, Psoralen plus ultraviolet A light therapy, and immunosuppressive drugs.

4.1.2 Inclusion criteria

The inclusion criteria are generally consistent with the final NICE scope,[11] with three main inconsistencies: (1) the company's SLR is broader in terms of population, with no age limit, and mentions AT and AU as well as some other subtypes of AA, with no mention of severity, whereas the final NICE scope refers to people with severe AA aged 12 years and over; (2) the company's SLR inclusion criteria are broader in terms of interventions, listing ritlecitinib, immunomodulators, systemic non-corticosteroids, oral corticosteroids, intralesional corticosteroids, topical corticosteroids, topical non-steroids, contact immunotherapy, platelet-rich plasma, other JAK inhibitors, biologics, treatments in development and other pharmacological treatments, with no non-pharmacological comparators listed, whereas the final NICE scope[11] only refers to ritlecitinib as an intervention and established clinical management without ritlecitinib as the comparator; and (3) the final NICE scope[11] specifies established clinical management without ritlecitinib as the comparator of interest, whereas the company's SLR does not specify any non-pharmacological comparator (e.g., established clinical management or BSC) in Table 1 of CS Appendix D.[13] The company's clarification response (questions A1 and A2)[2] clarified that non-pharmacological clinical management was not considered as a comparator in the SLR, however, placebo was considered as a comparator if the intervention was a comparator of interest and that "*non-pharmacological treatment may have been used alongside investigational treatment and it would be an add on to pharmacological treatment rather than an alternative to it. For example, in the ALLEGRO 2b/3 study, patients were able to continue using non-pharmacological clinical management such as wigs alongside the investigational treatments (ritlecitinib or placebo)*". Whilst these inconsistencies differ from the decision problem set out in the final NICE scope, the EAG does not consider them to be problematic, as they would broaden rather than narrow the scope of the review, meaning that the relevant studies would still have been identified.

4.1.3 Critique of study selection

CS Appendix D[13] states that two reviewers independently screened titles and abstracts of each record and then full texts, with any discrepancies resolved by a third reviewer. The EAG considers this to be an appropriate and high-quality reviewing method. Figure 1 in Appendix D of the CS[13] states that

screening was technology-aided; however, the company's clarification response (question A3)[2] clarifies that while this process was planned to be technology-aided, technology was not required during screening when the review was undertaken. A list of studies excluded at the full text stage is provided in CS Appendix D, Table 8.[13] The EAG has screened the titles of the full texts excluded by the company and agrees that nothing of potential relevance has been excluded. Neither the EAG nor their clinical advisors are aware of any additional relevant studies within the scope of this appraisal. The SLR did not explicitly state the number of included studies for which ritlecitinib was the intervention. The company's clarification response[2] (question A4) clarifies that two of the included studies included ritlecitinib as the intervention; these were the ALLEGRO 2a proof of concept study and the ALLEGRO 2b/3 study.

4.1.4 Critique of data extraction

CS Appendix D[13] states that two reviewers independently extracted data into a pre-defined template with these extractions checked and reconciled by a third independent reviewer. The EAG considers this to be an appropriate and high-quality reviewing method. Figure 1 in Appendix D of the CS[13] states that data extraction was technology-aided; however, the company's clarification response (question A3)[2] clarifies that while this process was planned to be technology-aided, technology was not required during data extraction when the review was undertaken. Data fields extracted are provided in CS Appendix D, Table 2, and the EAG is satisfied that the fields extracted are comprehensive.

4.1.5 Critique of quality assessment

The quality of the ALLEGRO 2b/3 study was assessed using a checklist based on the Centre for Reviews and Dissemination (CRD) criteria for assessing the risk of bias in randomised controlled trials (RCTs),[15] which is based on the Cochrane Risk of Bias tool,[16] which is widely regarded as a robust tool for assessing bias in RCTs. The company's clarification response (question A5)[2] clarifies that a single reviewer performed the quality assessment of the ALLEGRO 2b/3 and ALLEGRO-LT studies. The EAG notes that a more robust method for assessing quality would have been for two reviewers to perform quality assessment, ideally independently.[15] The quality of the ALLEGRO-LT study was assessed using a checklist adapted from the Critical Appraisal Skills Programme (CASP) Cohort Study Checklist.[17] The EAG notes that only seven of the twelve questions in the CASP checklist have been used in the checklist presented in CS Appendix D,[13] Table 11 and applied to the ALLEGRO-LT study. However, these questions appear to be the most appropriate and relevant questions for the appraisal of this study, and the company's clarification response[2] (question A6) states that these questions were omitted from the outline suggested in the NICE methods guide and thus were not addressed in the quality assessment of the ALLEGRO-LT study.

No attempt has been made to integrate the assessment of study quality into the findings reported in the CS, or to consider the overall impact of the quality of the included studies on the results.

Quality assessment of the four ALLEGRO studies (ALLEGRO 2b/3, ALLEGRO-LT, ALLEGRO 2a proof of concept study and ALLEGRO-2a safety study) as undertaken by the company (where undertaken) and the EAG, is presented in Section 4.2.3.

4.1.6 Critique of evidence synthesis

The CS does not include any formal evidence synthesis. The EAG notes that while the company identified relevant evidence using systematic, high-quality procedures, no results from the included studies were reported (aside from the ritlecitinib studies) and no attempt was made to synthesise the evidence identified (e.g., to examine the efficacy of each type of pharmacological treatment), narratively or by presenting the study findings in tables to allow for comparison. Therefore, the EAG considers the SLR to be incomplete.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

4.2.1 Studies included in/excluded from the submission

The CS[1] includes two studies that examine the efficacy and safety of ritlecitinib for the treatment of AA: ALLEGRO 2b/3,[18] a pivotal RCT; and ALLEGRO-LT,[19] an open-label extension of the ALLEGRO 2b/3 study. These studies are reported in the CS as the most relevant to the decision problem.[1] Two further studies of ritlecitinib for the treatment of AA are also mentioned in the CS: ALLEGRO 2a proof of concept study[20] (which reports safety and efficacy data but was excluded from the CS due to the dosage not matching the proposed licensed dose); and ALLEGRO-2a safety study[21] (which was not reported on in the CS as this study presents safety data only). The study characteristics of these studies are presented in Table 4.

Table 4: Characteristics of the ALLEGRO studies

Study	Design	Population	Interventions	Comparator	Primary outcome
ALLEGRO 2b/3	RCT	Adolescents and adults aged ≥12 years with clinical diagnosis of AA, ≥50% scalp hair loss (SALT ≥50) including AT and AU, no regrowth within 6 months, current episode ≤10 years.	Ritlecitinib 200 mg once daily for 4 weeks followed by 50 mg once daily (n=132) Ritlecitinib 200 mg once daily for 4 weeks followed by 30 mg once daily (n=130) Ritlecitinib 50 mg once daily (n=130) Ritlecitinib 30 mg once daily (n=132) Ritlecitinib 10 mg once daily (n=63)	Placebo ritlecitinib to match 200 mg daily (4 tablets) for 4 weeks followed by 50 mg (1 tablet) daily (n=65) Placebo ritlecitinib to match 50 mg (1 tablet) daily (n=66)	(1) SALT ≤20 at week 24 (study primary outcome and FDA primary outcome); (2) SALT ≤10 at week 24 (EMA primary outcome)
ALLEGRO-LT (ongoing)	Single-arm	(1) Those who previously participated in ALLEGRO 2a proof of concept study and ALLEGRO 2b/3 (if ≥30 days since participation, requirement for ≥25% scalp hair loss); (2) *De novo* patients aged ≤12 years with AA, and ≥25% scalp hair loss, including AT and AU	Ritlecitinib 200 mg once daily for 4 weeks followed by 50 mg once daily (*de novo* participants) Ritlecitinib 50 mg once daily (roll-over participants from ALLEGRO 2a/ALLEGRO 2b/3)	N/a (single-arm)	Incidence of AEs, SAEs, and AEs leading to discontinuation.
ALLEGRO 2a proof of concept study	RCT	Adults ≥18 years old with severe AA (≥50% scalp hair loss, including AT and AU), no regrowth within 6 months, current episode ≤7 years.	Ritlecitinib 200 mg once daily for 4 weeks followed by 50 mg once daily (n=48)	Placebo ritlecitinib or placebo brepocitinib to match dosing regime (n=47)	Change from baseline in SALT score at week 24.

Study	Design	Population	Interventions	Comparator	Primary outcome
			Brepocitinib 60 mg once daily for 4 weeks followed by 30 mg once daily (n=47)		
ALLEGRO-2a safety study (ongoing)	RCT	Adults aged 18 to ≤50 years, with ≥25% scalp hair loss.	Ritlecitinib 200 mg once daily for 4 weeks followed by 50 mg once daily until month 60	Placebo ritlecitinib to match 200 mg once daily (4 tablets) for 4 weeks followed by 50 mg (1 tablet) once daily to month 24 (followed by ritlecitinib 50 mg once daily to month 60)	Auditory function, assessed via the BAEP at a stimulus intensity of 80 decibels at Month 9.

AA - alopecia areata; AT - alopecia totalis; AU - alopecia universalis; EMA - European Medicines Agency; FDA - United States Food and Drug Administration; N - number; N/a - not applicable; RCT - randomised controlled trial; SAE - serious adverse event; SALT - Severity of Alopecia Tool.

ALLEGRO 2b/3 is a pivotal double-blind, randomised, placebo-controlled, dose-ranging Phase 2b/3 clinical trial. The study record registered on Clinicaltrials.gov (NCT03732807)[22] states that patients were enrolled into the ALLEGRO 2b/3 study at 155 investigational sites across 18 countries: The US (46 sites), Argentina (2 sites), Australia (8 sites), Canada (12 sites), Chile (4 sites), China (11 sites), Columbia (3 sites), Czechia (5 sites), Germany (6 sites), Hungary (4 sites), Japan (6 sites), Republic of Korea (2 sites), Mexico (2 sites), Poland (8 sites), Russian Federation (9 sites), Spain (10 sites), Taiwan (7 sites), and the UK (10 sites; 8 of which were in England).

ALLEGRO-LT is an ongoing Phase 3, multi-centre, single-arm, open-label extension study. Patients were enrolled into the study in two cohorts: (1) those who rolled over from the ALLEGRO 2a proof of concept study and ALLEGRO 2b/3; and (2) *de novo* adult and adolescent patients with AA recruited into the study. The study record registered on Clinicaltrials.gov (NCT04006457)[23] states that patients were enrolled into the ALLEGRO-LT study at 148 investigational sites across 17 countries: the US (49 sites), Argentina (2 sites), Australia (8 sites), Canada (12 sites), Chile (4 sites), China (8 sites), Columbia (3 sites) Czechia (4 sites), Germany (6 sites), Japan (7 sites), Republic of Korea (2 sites), Mexico (2 sites), Poland (12 sites), Russian Federation (9 sites), Spain (7 sites), Taiwan (6 sites), and the UK (7 sites; 6 of which were in England).

The ALLEGRO 2a proof of concept study is a double-blind, randomised, placebo-controlled, parallel-group, multi-centre Phase 2a clinical trial. The study record registered on Clinicaltrials.gov (NCT02974868)[24] states that patients were enrolled into the ALLEGRO 2a proof of concept study at 55 investigational sites across three countries: the US (39 sites), Australia (10 sites), and Canada (6 sites).

The ALLEGRO-2a safety study is an ongoing global double-blind, randomised, placebo-controlled Phase 2a clinical trial. The study record registered on Clinicaltrials.gov (NCT04517864)[25] states that patients were enrolled into the ALLEGRO-2a safety study at 40 investigational sites across four countries: the US (24 sites), Australia (3 sites), Canada (4 sites), and Poland (9 sites).

In summary, ALLEGRO 2b/3 and ALLEGRO-LT form the main clinical effectiveness and safety evidence base in this appraisal, and thus the EAG's appraisal of the CS focuses on these two key studies. However, for completeness, and because these studies contribute safety evidence to the pooled safety data used in the cost-effectiveness model, the EAG has also examined evidence from the ALLEGRO 2a proof of concept study and the ALLEGRO-2a safety study. The evidence relating to the clinical effectiveness of ritlecitinib addresses the final NICE scope.[11]

4.2.1.1 Patients

Eligibility criteria for the four ALLEGRO studies are presented in Table 5. Clinical advisors to the EAG have confirmed that the eligibility criteria for these studies is reasonable and representative of the patients seen in UK routine clinical practice. One advisor highlighted that patients tend to be seen at an earlier stage in the clinic than those who would be included in these studies (who were required to have no evidence of regrowth within 6 months at screening and baseline visits), and another commented that the baseline SALT score would typically be lower in clinic, closer to 60. The EAG notes that patients with depression and suicide ideation were excluded from the ALLEGRO 2b/3 study. Given the psychological impact of AA on patients (as documented in the CS[1] and Section 2.1), it is possible that the participants of the ALLEGRO 2b/3 study were not representative of AA patients overall. The company's clarification response[2] (question A9) clarifies that the decision to exclude patients with depression and suicide ideation from the ALLEGRO 2b/3 study was taken on ethical grounds, to protect potentially vulnerable patients from harm, and that five patients were documented as 'screen failures' based on this criterion, further adding that "*Investigators may not have approached patients that did not meet these criteria but there isn't any empiric evidence of this*". Therefore, it seems unlikely that large numbers of relevant patients were excluded based on this criterion, although the EAG notes that this is uncertain as no details were provided as to how patients were identified and recruited into the study.

Table 5: Key inclusion criteria of the ALLEGRO studies (adapted from CS, Table 8 and CS Appendix D, and ALLEGRO 2a proof of concept study CSR, Section 9.3.2)

Criteria	ALLEGRO 2b/3	ALLEGRO-LT	ALLEGRO 2a proof of concept	ALLEGRO-2a safety study
Inclusion criteria	• Male or female aged ≥12 years • Have a clinical diagnosis of AA with no other aetiology of hair loss (e.g., telogen effluvium, androgenetic alopecia). • ≥50% hair loss of the scalp (measured by SALT), including AT and AU, without evidence of terminal hair regrowth within 6 months at both the screening and baseline visits. • Current episode of hair loss ≤10 years. • If receiving permitted concomitant medications for any reason other than AA, subjects should have been on a stable regimen, which was defined as not starting a new drug or changing dosage within 7 days or 5 half-lives (whichever is longer) prior to Day 1. Subject must have been willing to stay on a stable regimen during the duration of the study.	• Patients ≥ 12 years • Diagnosis of AA with ≥ 25% scalp hair loss due to AA, including AT or AU (*de novo* patients and those rolled over from ALLEGRO 2b/3 or ALLEGRO 2a with last dose >30 days previously) • No evidence of terminal hair regrowth within 6 months at both screening and baseline visits (this applies to *de novo* patients only) • Maximum duration of current episode of hair loss ≤ 10 years (this applies to *de novo* patients only)	• Adults aged ≥18 years • Severe AA (≥50% scalp hair loss, including AT and AU) • No hair regrowth within 6 months of the screening and baseline visits • Current episode of hair loss ≤7 years in duration	• Adults aged 18 to ≤50 years • Diagnosis of AA, including AT and AU • Scalp hair loss ≥25% due to AA • Normal hearing and normal brainstem evoked potentials • Normal neurological exam; can have a stable unilateral median neuropathy or ulnar neuropathy • Signed informed consent • Stable regimen for other medications before and during the study
Exclusion criteria	• Participation in other studies involving investigational drug(s) within 8 weeks or within 5 half-lives (if known), whichever is longer, prior to study entry and/or during study participation	For *de novo* patients and those rolled over from ALLEGRO 2b/3 and ALLEGRO 2a with last dose >30 days previously: • Hearing loss with progression over the previous 5 years, or	• Any psychiatric condition including recent or active suicidal ideation or behaviour • Other types of alopecia • Other scalp disease that may impact AA assessment	• Other significant medical conditions • Occupational or recreational noise exposure • History of peripheral neuropathy or first degree

Criteria	ALLEGRO 2b/3	ALLEGRO-LT	ALLEGRO 2a proof of concept	ALLEGRO-2a safety study
	• Other types of alopecia • Other scalp disease that may impact AA assessment • Active systemic diseases that may cause hair loss • Any psychiatric condition including recent or active suicidal ideation or behaviour that meets any of the following criteria: • Suicidal ideation associated with actual intent and a method or plan in the past year: "Yes" answers on items 4 or 5 of the C-SSRS • For subjects who had previous history of suicidal behaviours in the past >1 year to <5 years: "Yes" answer (for events that occurred in the past 5 years) to any of the suicidal behaviour items of the C-SSRS or any lifetime history of serious or recurrent suicidal behaviour, a risk assessment must be performed, and documented, by a qualified mental health professional to assess whether it is safe for the subject to participate in the trial • Clinically significant depression as indicated by the PHQ-8 total score ≥15 • The presence of any current major psychiatric disorder that	sudden hearing loss, or middle or inner ear disease • History of current malignancies (except for adequately treated or excised non metastatic basal cell or squamous cell cancer of the skin or cervical carcinoma in situ) • History of a single episode of disseminated herpes zoster or disseminated herpes simplex, or a history of more than one episode of localized, dermatomal herpes zoster • Infection requiring hospitalisation, or parenteral antimicrobial therapy within 6 months prior to Day 1 For all patients: • Previous use of any JAK inhibitors ≤12 weeks prior to the screening visit	• Current or recent history of clinically significant severe, progressive, or uncontrolled renal, hepatic, haematological, gastrointestinal, metabolic, endocrine, pulmonary, cardiovascular, psychiatric, immunologic/rheumatologic or neurologic disease; or had any other severe acute or chronic medical or psychiatric condition or laboratory abnormality that could have increased the risk associated with study participation or IP administration, or interfere with the interpretation of study results; or in the opinion of the investigator or Pfizer (or designee), the participant was inappropriate for entry into this study, or unwilling/unable to comply with study procedures and lifestyle requirements • Current or recent history of clinically significant severe, progressive, or uncontrolled hearing loss or auditory disease • History of herpes zoster • Active acute or chronic skin infection within 4 weeks prior to Day 1 or superficial skin infections within 2 weeks prior to Day 1	relative with a hereditary peripheral neuropathy • HbA1c ≥7.5% at Screening • Recurrent or disseminated Herpes Zoster • Active or chronic infection; or infection requiring hospitalization or intravenous antimicrobials within 6 months • Active or latent (insufficiently treated) Hepatitis • Active or latent (insufficiently treated) tuberculosis • Concomitant medications associated with peripheral neurologic or hearing loss • Protocol specific laboratory abnormalities

Criteria	ALLEGRO 2b/3	ALLEGRO-LT	ALLEGRO 2a proof of concept	ALLEGRO-2a safety study
	is not explicitly permitted in the inclusion/exclusion criteria • Have hearing loss with progression over the previous 5 years, or sudden hearing loss, or middle or inner ear disease • Have previous use of any JAK inhibitor for use in any disease indication or any non-B-cell selective lymphocyte-depleting agent. • No active tuberculosis infection, confirmed by chest X-Ray • Any of the following abnormalities in clinical laboratory tests at screening, as assessed by the study-specific laboratory and confirmed by a single repeat, if deemed necessary: • Absolute neutrophil count <1.2 x 109/L (<1200/mm3); • Haemoglobin <11.0 g/dL or haematocrit <33%; • Platelet count <150 × 109/L or <150,000/mm3; • Absolute lymphocyte count of <0.8 x 109 /L (<800/mm3); • Estimated glomerular filtration rate <60 ml/min/1.73 m2 based on the age appropriate calculation; • Enzymes aspartate transaminase or alanine		• Known immunodeficiency disorder or first-degree relative with hereditary immunodeficiency • Malignancy or history of malignancies with the exception of adequately treated or excised non-metastatic basal cell or squamous cell cancer of the skin or cervical carcinoma in situ • Required treatment with prohibited concomitant medication(s) • Infected with Mycobacterium tuberculosis • Any of the following abnormalities in clinical laboratory tests at screening, as assessed by the study-specific laboratory and confirmed by a single repeat, if deemed necessary: • absolute neutrophil count of <2.5 × 109/L (<2500/mm3); • haemoglobin <10.0 g/dL or haematocrit <30%; • platelet count below the lower limit of normal (LLN) at screening; • absolute lymphocyte counts of <0.8 × 109 /L (<800/mm3);	

Criteria	ALLEGRO 2b/3	ALLEGRO-LT	ALLEGRO 2a proof of concept	ALLEGRO-2a safety study
	transaminase values >2 × upper limit of normal (ULN); • Total bilirubin >1.5 × ULN; subjects with Gilbert's syndrome would be eligible for this study provided the direct bilirubin is ≤ULN; • Have any clinically significant laboratory abnormality that that could affect interpretation of study data.		• serum creatinine >ULN or estimated glomerular filtration • rate <80 ml/min/1.73m2 based on the age appropriate calculation; • enzymes aspartate aminotransferase or alanine aminotransferase values >2 × ULN; • total bilirubin ≥1.5 ×ULN; participants with a history of Gilbert's syndrome may have had a direct bilirubin measured and were eligible for this study provided the direct bilirubin was ≤ULN; • creatine kinase (CK) >3 ×ULN and positive urine myoglobin; • In the opinion of the investigator or sponsor (or designee), had any uncontrolled clinically significant laboratory abnormality that could affect interpretation of study data or the participant's participation in the study.	

AA - alopecia areata; AT - alopecia totalis; AU - alopecia universalis; C-SSRS - Columbia-Suicide Severity Rating Scale; JAK - Janus kinase; LLN - lower limit of normal RCT - randomised controlled trial; PHQ-8 - eight-item Patient Health Questionnaire depression scale; SAE - serious adverse event; SALT - Severity of Alopecia Tool; ULN - upper limit of normal.

The eligibility criteria for the pivotal study, ALLEGRO 2b/3, closely align with the population as defined in the final NICE scope,[11] which is defined as "*People aged 12 years and over with severe alopecia areata*". The eligibility criteria for the ALLEGRO-LT study match the population defined in the final NICE scope[11] in terms of age, although study participants are required to have ≥25% scalp hair loss, so it is possible for those who have moderate AA to participate. Given that the prognosis of patients with milder AA is more positive, it is possible that the patients enrolled in the ALLEGRO-LT study may experience a different treatment response to those with severe AA. The ALLEGRO 2a proof of concept study matches the final NICE scope in terms of the requirement for severe AA (≥50% scalp hair loss); however, the patients were required to be aged ≥18 years. It should be noted though that a full list of inclusion and exclusion criteria for the ALLEGRO 2a proof of concept study was not available to the EAG. Participants in the ALLEGRO-2a safety study were also required to be aged ≥18 (and ≤50) years, and could have less severe AA than specified in the final NICE scope[11] (≥25% scalp hair loss), therefore the ALLEGRO-2a safety study is likely to represent an older population with potentially milder AA than those who may be eligible to receive ritlecitinib in clinical practice in England.

The CSRs for the ALLEGRO 2b/3 study,[18], the ALLEGRO-LT study,[19] the ALLEGRO 2a proof of concept study,[26] and the ALLEGRO-2a safety study[21] do not report any detail on how patients were identified and recruited (in the case of ALLEGRO-LT, this relates to the *de novo* patients). The company's clarification response[2] (question A18) outlined eligibility criteria, but did not clarify how participants were identified and recruited. Therefore, the EAG cannot assess whether the process of recruitment may have introduced selection bias into these studies, nor whether a representative sample of patients with severe AA is likely to have been recruited to each.

For the ALLEGRO-LT study, the company's clarification response[2] (question A23) reveals that at the time of data cut-off [redacted] of the roll-over patients had rolled over from the ALLEGRO 2b/3 study [redacted], with [redacted] rolling over from the ALLEGRO 2a proof of concept study. Of those rolling over from ALLEGRO 2b/3, [redacted] had been in the ritlecitinib 50 mg arm. Table 6 in the company's clarification response[2] (question A23) clarifies that [redacted] of those rolling over from the ALLEGRO 2a proof of concept study had previously received brepocitinib either as first-line treatment or as cross-over treatment following ritlecitinib or placebo ritlecitinib treatment failure.

A diagram illustrating participant flow in the ALLEGRO 2b/3 study is presented in Figure 4 of CS Appendix D.[13]

[redacted]

[redacted] The company's clarification response[2] (question A19) states that [redacted] Participants in both placebo arms switched to ritlecitinib in a blinded fashion at week 24.

[redacted] The company's clarification response[2] (question A20) states that reasons for physician withdrawal included [redacted].[1] The company's clarification response[2] (question A20) clarifies that reasons for participant withdrawal included [redacted]

[redacted]

Patient disposition in the ALLEGRO-LT study as of the data cut-off point is reported in the interim CSR.[19]

[redacted]

Patient disposition within the ALLEGRO 2a proof of concept study is presented in Table 10 of the CSR.[26]

[redacted]

Patient disposition in the ALLEGRO-2a safety study as of the data cut-off point is presented in the interim CSR.[21]

[redacted]

In the ALLEGRO 2b/3 study, baseline demographic and clinical characteristics (as presented in the CS, Table 10) were comparable between arms, with the following exceptions:

- There was a higher proportion of males in the ritlecitinib 50 mg arm ([redacted]) compared with both the 200/50 mg placebo arm ([redacted]) and 50 mg placebo arm ([redacted]);
- A greater proportion of patients in the ritlecitinib 50 mg arm were Asian ([redacted]) compared with both placebo arms (200/50 mg placebo: [redacted]; 50 mg placebo: [redacted]), and a smaller proportion of patients in the ritlecitinib 50 mg arm were white ([redacted]) compared with both placebo arms (200/50 mg placebo: [redacted]; 50 mg placebo: [redacted]);
- A smaller proportion of patients in the ritlecitinib 50 mg arm had AU ([redacted]) compared with both placebo arms (200/50 mg placebo: [redacted]; 50 mg placebo: [redacted]);
- A greater proportion of patients in the ritlecitinib 50 mg arm had AT ([redacted]) compared with the pooled placebo arm ([redacted]) and the 50 mg placebo arm ([redacted]), although the 200/50 mg placebo arm was comparable ([redacted]);
- Slightly less time since diagnosis in the ritlecitinib 50 mg arm (mean years since diagnosis: [redacted], standard deviation (SD) [redacted]) than in the pooled placebo arm ([redacted], SD [redacted]) and the 50 mg placebo arm ([redacted], SD [redacted]), although the 200/50 mg placebo arm was comparable ([redacted], SD [redacted]);
- A slightly lower proportion of patients in the ritlecitinib 50 mg arm had prior intralesional steroid injections ([redacted]) and other topical anti-inflammatory treatment ([redacted]) than among both placebo arms (intralesional steroid injections: 200/50 mg placebo [redacted] and 50 mg placebo [redacted]; other topical anti-inflammatory: 200/50 mg placebo [redacted] and 50 mg placebo [redacted]);
- A higher proportion of patients in the ritlecitinib 50 mg arm had prior topical corticosteroid treatment ([redacted]) than among both placebo arms (200/50 mg placebo: [redacted]; 50 mg placebo: [redacted]).

It is possible that these differences may impact on treatment effects. Clinical advice provided to the EAG suggested that females are more greatly impacted than males in terms of psychological impact and quality of life by having AA, but not treatment effects, and thus differences in the balance of males and females between the ritlecitinib 50 mg arm and placebo arms may impact on quality of life and patient satisfaction results but not hair regrowth. Differences in the proportions of patients with AU (lower) and AT (higher) in the ritlecitinib 50 mg arm than in the placebo arms may potentially impact the treatment response outcomes, and clinical advice received by the EAG concurred that a smaller treatment response would be expected in patients with AT/AU (see Section 4.2.4.6), although the mean SALT score at baseline was comparable between the ritlecitinib 50 mg ([redacted]) and pooled placebo ([redacted]) arms. Likewise, differences in prior treatment between the ritlecitinib 50 mg

and placebo arms may potentially impact treatment mean difference in SALT score and proportion of SALT 20 and SALT 10 responders at follow-up timepoints. Clinical advice received by the EAG has confirmed that the baseline demographic and clinical characteristics of the patients enrolled in this study were comparable with patients usually seen in clinical practice in England. One clinician stated that patients tend to be seen at an earlier stage in the clinic than those included in this study (who were required to have no evidence of regrowth within 6 months at screening and baseline visits), often due to patient demand.

Patient characteristics of *de novo* patients in the ALLEGRO-LT study are presented in Table 12 of the CS.[1] Clinical advice received by the EAG suggested that the proportion of female patients is broadly similar to (or a little lower than) that seen in clinical practice. Compared with the baseline characteristics of patients in the ALLEGRO 2b/3 study, the *de novo* patients in the ALLEGRO-LT study are comparable in terms of the proportion of patients with AT/AU ([redacted]); however, the mean SALT score is lower among the ALLEGRO-LT *de novo* patient cohort ([redacted], SD [redacted]), suggesting this cohort of patients may have less severe AA overall than those enrolled in the ALLEGRO 2b/3 study. Demographic characteristics of roll-over patients at the baseline of the index study are presented in Table 5 in the interim CSR.[19] Baseline characteristics are similar to the *de novo* cohort, with [redacted] female and [redacted] white. Mean age at ALLEGRO-LT baseline [redacted] was comparable with the patients in the ALLEGRO 2b/3 study. Table 14.1.3.1 in the interim CSR[19] shows a similar if not slightly higher duration since diagnosis among roll-over [redacted] and a slightly higher duration of current episode [redacted]. A slightly higher proportion of roll-over patients had AT/AU at baseline compared with *de novo* patients [redacted] and a higher proportion had ever experienced AT/AU compared with *de novo* patients [redacted] (Table 14.1.3.4.1, interim CSR[19]).

The ALLEGRO 2a proof of concept study is not a key focus in the CS, and only provides randomised placebo-controlled evidence of the 200/50 mg dose of ritlecitinib up to 24 weeks (after which some patients switch treatment, depending on the response). Nevertheless, this study contributes safety data to analyses and therefore the EAG has examined baseline patient characteristics, as presented in Table 18 of the CSR.[26] [redacted]

[redacted]

4.2.1.2 Intervention

The doses of ritlecitinib administered in the ALLEGRO 2b/3, ALLEGRO-LT, ALLEGRO 2a proof of concept and ALLEGRO-2a safety studies are outlined in the CS[1] (Section B.2.2). In all cases, ritlecitinib is administered orally once daily. The ritlecitinib doses administered in the ALLEGRO 2b/3 study (as outlined in CS,[1] Table 7) are: 200 mg once daily loading dose for 4 weeks followed by 50 mg once daily maintenance dose (referred to as 200/50 mg dose); 200 mg once daily loading dose for 4 weeks followed by 30 mg once daily maintenance dose (referred to as 200/50 mg dose); 50 mg once daily dose, 30 mg once daily dose and 10 mg once daily dose. In the ALLEGRO-LT study, patients who had rolled over from ALLEGRO 2a or ALLEGRO 2b/3 and had previously been treated with ritlecitinib received a dose of 50 mg ritlecitinib one daily, whereas *de novo* patients and those who did not receive ritlecitinib in either previous study received a 200 mg once daily loading dose of ritlecitinib for 4 weeks followed by a 50 mg once daily maintenance dose. All patients randomised to ritlecitinib in the ALLEGRO 2a proof of concept study received a 200 mg once daily loading dose for 4 weeks followed by a 50 mg once daily maintenance dose for 20 weeks (referred to as 200/50 mg dose). In the ALLEGRO-2a safety study, participants in the ritlecitinib arm also received a 200 mg once daily loading dose for 4 weeks followed by a 50 mg once daily maintenance dose (referred to as 200/50 mg dose).

The dose of ritlecitinib administered in the 50 mg dosing arm of the ALLEGRO 2b/3 study and to patients in the ALLEGRO-LT study who had previously received ritlecitinib in the ALLEGRO 2a proof of concept study and ALLEGRO 2b/3 is consistent with the proposed licensed dose of 50 mg once daily.

The ALLEGRO-LT study is ongoing and the CS[1] does not contain details of how long patients had been on treatment at the time of the data cut-off. According to the interim CSR,[19] data were available for [redacted].

In the ALLEGRO 2b/3 study, [redacted] of the [redacted] patients had protocol deviations that led to discontinuation; [redacted] of these were in the ritlecitinib 50 mg arm.[13] The number of patients with important protocol deviations in the ALLEGRO-LT study was not reported in the CS or interim CSR. In the ALLEGRO 2a proof of concept study, [redacted] patients had at least one protocol deviation considered to [redacted] (CSR, page 422).[26] In the ALLEGRO-2a safety study as of

the data cut-off date, [redacted] participants enrolled had at least one protocol deviation considered to [redacted].[21] Further details are provided in Section 4.2.3.3.

4.2.1.3 Comparator

The comparator in the ALLEGRO 2b/3, ALLEGRO 2a proof of concept and ALLEGRO-2a safety studies was/is placebo capsules identical to the ritlecitinib capsules with the same dosage instructions.[1] The comparator in the final NICE scope[11] is established clinical management without ritlecitinib. The CS[1] confirms that patients in the placebo arms of the pivotal ALLEGRO 2b/3 study were permitted to use non-clinical management (e.g., wigs), so to this end the EAG considers the placebo arms of the ALLEGRO 2b/3 study to be consistent with the NICE scope.[11] It is unclear whether participants in the placebo arms of the ALLEGRO 2a proof of concept study and the ALLEGRO-2a safety study were permitted to use non-clinical management during the study.

The ALLEGRO-LT study adopted a single-arm design; hence, no comparator was included. No indirect comparison was undertaken between data from ALLEGRO-LT and data from those who had established clinical management (e.g., hospital registry data), which is not entirely consistent with the final NICE scope.[11] Guidance on performing clinical trials of medicines published by the European Medicines Agency (EMA) recommends that trials aiming to demonstrate/confirm efficacy are controlled, with randomised allocation to arms.[27] The ALLEGRO-LT study, however, is a long-term extension of two RCTs (the ALLEGRO 2a proof of concept study and the ALLEGRO 2b/3 study), and thus any patients rolling over from either RCT would have been randomly allocated to active treatment or placebo at the start of the initial RCT. It should be borne in mind; however, that additional *de* novo patients were recruited directly into the ALLEGRO-LT study, and therefore the open-label, uncontrolled design of ALLEGRO-LT should be taken into consideration during review of data from that study, in particular data pertaining to efficacy.

4.2.1.4 Outcomes

The key outcomes listed in the CS for the ALLEGRO 2b/3, ALLEGRO-LT, ALLEGRO 2a proof of concept and ALLEGRO-2a safety studies are summarised in

Table 6, Table 7, Table 8 and Table 9. All outcomes presented in the CS were included in the final NICE scope,[11] and all outcomes from the final NICE scope were reported on in the CS.

All efficacy outcomes in the ALLEGRO 2b/3 study and the ALLEGRO 2a proof of concept study[26] were analysed using the full analysis set (FAS), defined as all randomised patients, regardless of whether they received study intervention, analysed according to the treatment to which they were randomised.[1] Efficacy outcomes in the ALLEGRO-LT study were analysed using the FAS, defined as all participants assigned to treatment.[13, 19] Efficacy outcomes in the ALLEGRO-2a safety study were assessed using the efficacy analysis set (EAS), defined as all randomised participants who received at least one dose of study intervention, analysed according to the intervention to which they were assigned.[21] Safety outcomes in the ALLEGRO 2b/3, ALLEGRO-LT and ALLEGRO 2a proof of concept studies were analysed using the safety analysis set (SAS), defined as all patients who received at least one dose of the study drug, classified as the actual intervention received for most of the time during the study.[1, 13, 18, 19, 26] Safety outcomes in the ALLEGRO-2a safety study were assessed using the SAS, defined as all participants who received at least one dose of study intervention, classified according to the actual treatment received.[21]

Table 6: Summary of ALLEGRO 2b/3 key outcomes listed in the CS and their relationship to the final NICE scope and the company's economic model

Outcome	In NICE scope?	Used in economic model?	Defined *a priori*?
Primary outcome			
Proportion of participants with a response based on an absolute SALT ≤20 at Week 24 (trial and FDA)	Yes ("*severity of alopecia areata*")	Yes	Yes
Proportion of participants achieving an absolute SALT score ≤10 (response) at Week 24 (EMA)[a]	Yes ("*severity of alopecia areata*")	Yes	Yes (originally the main primary outcome, in protocol)
Secondary outcomes			
Response based on an absolute SALT score ≤20 up to Week 48	Yes ("*severity of alopecia areata*")	Yes	Yes
Response based on an absolute SALT score ≤10 up to Week 48	Yes ("*severity of alopecia areata*")	Yes	Yes
Change from baseline in SALT scores up to Week 48	Yes ("*percentage of area affected by hair loss*")	No	Yes
Response based on at least a 2-grade improvement from baseline or a score of 3 in EBA score up to Week 48	Yes ("*severity of alopecia areata*")	No	Yes
Response based on at least a 2-grade improvement from baseline or a score of 3 in ELA score up to Week 48	Yes ("*severity of alopecia areata*")	No	Yes
PGI-C response defined as a score of "moderately improved" or "greatly improved" up to Week 48	Yes ("*severity of alopecia areata*")	No	Yes
EQ-5D-5L dimension scores at week 24 and week 48	Yes ("*health-related quality of life*")	No	No
Mean change from baseline in EQ-5D-5L to week 24 and week 48	Yes ("*health-related quality of life*")	No	No
EQ VAS scores at week 24 and week 48	Yes ("*health-related quality of life*")	No	No
SF-36 scores at week 24 and week 48	Yes ("*health-related quality of life*")	No	Yes
Change from baseline in SF-36 scores to week 24 and week 48	Yes ("*health-related quality of life*")	No	Yes
HADS depression and HADS anxiety scores at weeks 24 and 48	Yes ("*health-related quality of life*")	No	Yes
Change from baseline in HADS depression and HADS anxiety scores to weeks 24 and 48	Yes ("*health-related quality of life*")	No	Yes
AAPPO key item and domain scores at week 24 and week 48	Yes ("*health-related quality of life*")	No	Yes
Change from baseline in AAPPO item and domain scores at week 24 and week 48	Yes ("*health-related quality of life*")	No	Yes
Incidence of SAEs	Yes ("*adverse effects of treatment*")	Yes (for those with	Not specified in protocol

Outcome	In NICE scope?	Used in economic model?	Defined *a priori*?
		>5% incidence)	
Incidence of non-serious AEs	Yes ("*adverse effects of treatment*")	Yes (for those with >5% incidence)	Not specified in protocol

AA - alopecia areata; AAPPO - Alopecia Areata Patient Priority Outcomes; AE - adverse event; AT - alopecia totalis; AU - alopecia universalis; EBA - Eyebrow Assessment; ELA - Eyelash Assessment; EQ-5D-5L - EuroQol 5 dimensions 5 level; EQ VAS - EuroQol Visual Analogue Scale; HADS - Hospital Anxiety and Depression Scale; PGI-C - Patient's Global Impression of Change; SAE - serious adverse event; SALT - Severity of Alopecia Tool; SF-36 - 36-Item Short Form Health Survey
[a] This was a secondary outcome for the trial

Table 7: Summary of ALLEGRO-LT key outcomes listed in the CS (Appendix D) and their relationship to the final NICE scope and the company's economic model

Outcome	In NICE scope?	Used in economic model?	Defined *a priori*?
Primary outcome			
Incidence of TEAEs	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of SAEs and adverse events AEs leading to discontinuation	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of clinically significant abnormalities in vital signs	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of clinically significant abnormalities in clinical laboratory values	Yes ("*adverse effects of treatment*")	No	Yes
Secondary outcomes			
Response based on an absolute SALT score ≤10 at all timepoints	Yes ("*severity of alopecia areata*")	Yes	Yes
Absolute SALT score at all time points collected	Yes ("*severity of alopecia areata*")	Yes (distributed across four states: SALT <10, SALT 11 to 20, SALT 21 to 49 and SALT 50+)	Yes
Change from baseline in SALT scores at all time points collected	Yes ("*percentage of area affected by hair loss*")	No	Yes
Response based on achieving at least 50% improvement in SALT (SALT50) from baseline at all time points collected	Yes ("*severity of alopecia areata*")	No	Yes
Response based on achieving at least 75%, improvement in SALT (SALT75) from baseline at all time points collected	Yes ("*severity of alopecia areata*")	No	Yes
Response based on achieving at least 90% improvement in SALT (SALT90) from baseline at all time points collected	Yes ("*severity of alopecia areata*")	No	Yes

Outcome	In NICE scope?	Used in economic model?	Defined *a priori*?
Response based on at least a 2-grade improvement from baseline or a score of 3 in EBA score at all time points collected	Yes ("*severity of alopecia areata*")	No	Yes
Response based on at least a 2-grade improvement from baseline or a score of 3 in ELA score at all time points collected	Yes ("*severity of alopecia areata*")	No	Yes
Change from baseline in AAPPO scale total score, hair loss domain score, and psychological and functional impact domain at all time points collected[a]	Yes ("*health-related quality of life*")	No	Yes
Change from baseline in the HADS depression and anxiety scores at all time points collected[a]	Yes ("*health-related quality of life*")	No	Yes
Change from baseline in SF-36 individual Mental Component and Physical Component Score at all time points collected[a]	Yes ("*health-related quality of life*")	No	Yes

AA - alopecia areata; AAPPO - Alopecia Areata Patient Priority Outcomes; AE - adverse event; AT - alopecia totalis; AU - alopecia universalis; EBA - Eyebrow Assessment; ELA - Eyelash Assessment; HADS - Hospital Anxiety and Depression Scale; PGI-C - Patient's Global Impression of Change; SAE - serious adverse event; SALT - Severity of Alopecia Tool; SF-36 - 36-Item Short Form Health Survey; TEAEs treatment-emergent adverse events

[a] *The results relating to this outcome were not reported in the CS*

Table 8: Summary of ALLEGRO 2a proof of concept study key outcomes listed in the CS and CSR, and their relationship to the final NICE scope and the company's economic model

Outcome	In NICE scope?	Used in economic model?	Defined *a priori*?
Primary outcome			
Change from baseline in SALT score at Week 24 [a]	Yes ("*percentage of area affected by hair loss*")	No	Yes
Secondary outcomes			
Proportion of participants achieving ≥30% improvement in SALT score (SALT 30) at Week 24 [a]	Yes ("*severity of alopecia areata*")	No	Yes
Change from baseline in IGA at all time points up to Week 24 [a]	Yes ("*percentage of area affected by hair loss*")	No	Yes
Change from baseline in SALT score at intermediate time points up to Week 24 [a]	Yes ("*percentage of area affected by hair loss*")	No	Yes
Proportion of participants achieving SALT 30 at intermediate time points up to Week 24 [a]	Yes ("*severity of alopecia areata*")	No	Yes
Proportion of participants achieving ≥50%, ≥75%, ≥90% and 100% improvement from baseline in SALT score (SALT 50, SALT 75, SALT 90 and SALT 100) at all time points up to Week 24 [a]	Yes ("*severity of alopecia areata*")	No	Yes

Percentage change from baseline in SALT score at intermediate time points up to Week 24 [a]	Yes ("*percentage of area affected by hair loss*")	No	Yes
Incidence of TEAEs up to Week 24	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of specific clinical laboratory abnormalities including but not limited to anaemia, neutropenia, thrombocytopenia, lymphopenia, changes in lipid profile, and LFTs up to Week 24 [a]	Yes ("*adverse effects of treatment*")	No	Yes

IGA - Investigator Global Assessment; LFT - Liver Function Test; SALT - Severity of Alopecia Tool; SF-36 - 36-Item Short Form Health Survey; TEAEs treatment-emergent adverse events
[a] The results relating to this outcome were not reported in the CS

Table 9: Summary of ALLEGRO-2a safety study key outcomes listed in the CS and their relationship to the final NICE scope and the company's economic model

Outcome	**In NICE scope?**	**Used in economic model?**	**Defined *a priori*?**
Primary outcome			
Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80 dB at Month 9[a]	Yes ("*adverse effects of treatment*")	No	Yes
Secondary outcomes			
Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80dB at Months 6, 18 (aka 9E), and 24 (15E) [a]	Yes ("*adverse effects of treatment*")	No	Yes
Change from baseline in axonal dystrophy in skin punch biopsies at Month 9 and Month 24 (15E) [a]	Yes ("*adverse effects of treatment*")	No	Yes
Change from baseline in IENFD in skin punch biopsies at Month 9 and Month 24 (15E) [a]	Yes ("*adverse effects of treatment*")	No	Yes
Change from baseline in amplitude of wave V on BAEP at a stimulus intensity of 80 dB at Months 6, 9, 18 (9E), and 24 (15E) [a]	Yes ("*adverse effects of treatment*")	No	Yes
Absence of wave V on BAEP at stimulus intensities ranging from 80dB to 40dB at Months 6, 9, 18 (9E) and 24 (15E) [a]	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of TEAEs, TESAEs, and adverse events AEs leading to discontinuation	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of clinically significant abnormalities in vital signs [a]	Yes ("*adverse effects of treatment*")	No	Yes
Incidence of clinically significant abnormalities in clinical laboratory values [a]	Yes ("*adverse effects of treatment*")	No	Yes
Response to ritlecitinib measured by the SALT[a]	Yes ("*severity of alopecia areata*")	No	Yes
Response to ritlecitinib measured by PGI-C [a]	Yes ("*severity of alopecia areata*")	No	Yes

9E - month 9 of active therapy extension; 15E - month 15 of active therapy extension; AE - adverse event; BAEP - brainstem auditory evoked potential; dB - decibels; IENFD - intraepidermal nerve fibre density; PGI-C - Patient's Global Impression of Change; SALT - Severity of Alopecia Tool; TEAEs treatment-emergent adverse events; TESAEs treatment-emergent serious adverse events
[a] The results relating to this outcome were not reported in the CS

Primary outcomes

The ALLEGRO 2b/3 study had two primary outcomes, one for the trial overall and the US Food and Drug Administration (FDA) and one for the EMA. The trial/FDA primary outcome was the proportion of participants with a response based on an absolute SALT score of ≤20 at Week 24, and the EMA primary outcome was the proportion of participants achieving an absolute SALT score ≤10 at Week 24.[1] Both of these primary outcomes were used in the company's health economic model. Clinical advice received by the EAG has suggested that an acceptable level of scalp hair regrowth generally depends on the patient, with clinicians generally suggesting that an outcome of ≤20% hair loss would be generally acceptable, but that ≤10% would be more desirable.

The four primary outcomes of the ALLEGRO-LT study related to the incidence of various types of adverse event (see Table 7), with efficacy outcomes listed as secondary outcomes. The primary outcome for the ALLEGRO 2a proof of concept study was change from baseline in SALT score at Week 24. The primary and key secondary outcomes of the ALLEGRO-2a safety study relate to safety, with a focus on auditory function; the primary outcome is change from baseline in I-V interwave latency on brainstem auditory evoked potential (BAEP) at a stimulus intensity of 80 decibels (dB) at Month 9.

Secondary outcomes

In the pivotal ALLEGRO 2b/3 study, all outcomes reported in Table 8 of the CS[1] as key secondary outcomes were listed in the final NICE scope[11] (see

Table 6). Of these, only response based on an absolute SALT score ≤20 up to Week 48, response based on an absolute SALT score ≤10 up to Week 48, and treatment-emergent adverse events (TEAEs) with an incidence of >5% were used in the company's health economic model. The critical appraisal of clinical effectiveness evidence in the EAG report therefore focuses on this secondary outcome from the ALLEGRO 2b/3 study. Clinical advisors to the EAG confirmed that the most important outcome to their patients is good hair regrowth, which for some patients would be full regrowth and remission.

For the ALLEGRO-LT, ALLEGRO 2a proof of concept and ALLEGRO-2a safety studies, all outcomes reported in the CS were listed in the final NICE scope;[11] SALT response in ALLEGRO-LT was the only outcome from these studies to contribute to the company's health economic model. Therefore, the critical appraisal of clinical effectiveness evidence in the EAG report focuses on hair regrowth outcomes as reported by each study, which clinical advisors to the EAG highlighted as being clinically important for patients. Pooled safety data from all four studies is also provided in Appendix F of the CS,[28] although this only lists the most commonly reported TEAEs, and does not give frequencies of overall TEAEs or frequencies/proportions of patients with at least one TEAE, TRAE, SAE, severe TEAE, or TEAE leading to discontinuation.

4.2.1.5 Study design

The ALLEGRO 2b/3 study is a pivotal double-blind, randomised, placebo-controlled, dose-ranging Phase 2b/3 clinical trial, where eligible patients (n=718) were randomised to receive 200 mg once daily for 4 weeks followed by 50 mg once daily (referred to as 200/50 mg); 200 mg once daily for 4 weeks followed by 30 mg once daily (referred to as 200/50 mg); 50 mg once daily, 30 mg once daily, 10 mg once daily, placebo ritlecitinib tablets with an identical dosing regimen to the 200/50 mg arm, or placebo ritlecitinib tablets with an identical dosing regimen to the 50 mg arm at a 2:2:2:2:1:1:1 ratio using an interactive web response system.[1] The individual and/or organisation performing randomisation was not specified in the CS.[1] Randomisation was stratified by presence/absence of AT/AU diagnosis and age (<18 or ≥18 years), with 15% and 40% enrolment targets for patients aged <18 and with AT/AU, respectively.[1] Clinical advice received by the EAG suggests that patients with AT/AU have a worse prognosis than other AA patients, and the EAG considers stratification by AT/AU status to be appropriate. The ALLEGRO 2b/3 study consisted of a 24-week double-blind treatment period, following which all patients received ritlecitinib for a further 24 weeks (patients from the placebo arm received a 200mg once daily loading dose for 4 weeks, followed by 50 mg once daily, and patients from the ritlecitinib arm received 50 mg once daily, with an additional three placebo capsules over the 4-week loading dose period, to maintain blinding).[1] Following the completion of the 48-week treatment period, participants were eligible to enrol on the ALLEGRO-LT study. Investigators, patients and the sponsor study team were blinded to the treatment assigned at randomisation throughout the duration of the study. As a double-blind, placebo-controlled RCT, the EAG considers the study design to be

rigorous. Clinical advice received by the EAG has suggested that 24 weeks is a sufficient treatment duration for hair regrowth.

The ALLEGRO-LT study is an ongoing Phase 3, multi-centre, single-arm, open-label extension study of patients with AA (n=1052,[19] as of the data cut-off date, which the company's clarification response[2] (question A22) and the interim CSR[19] report as the 28th February 2022) who were treated with ritlecitinib either at a dose of 50 mg once daily (for patients who rolled over from either the ALLEGRO 2b/3 study or the ALLEGRO 2a proof of concept study and received ritlecitinib within that study) or 200 mg once daily for 4 weeks followed by 50 mg once daily (for *de novo* patients and those who rolled over from the ALLEGRO 2b/3 study or the ALLEGRO 2a proof of concept study and had not received ritlecitinib within that study). The ALLEGRO-LT study consists of an open-label 36-month treatment period. Interim data from the cohort of *de novo* patients has been reported in the CS; the cut-off date was not provided in the CS, however a cut-off date of February 2022 was provided in a conference paper.[29] Safety data from the ALLEGRO-LT study were presented up to Month 24, although safety data specific timepoints were not specified. The EAG considers the design of the ALLEGRO-LT study to be open to potential biases such as attrition bias, natural recovery and regression to the mean (particularly in relation to efficacy),[30] due to being open-label and single-arm. However, as the primary purpose was to evaluate safety, a single-arm open-label study is an appropriate design, as the focus is on identifying potentially rare clinical events rather than ensuring accurate judgement of treatment benefit.[31]

4.2.1.6 Ongoing studies

The ALLEGRO-LT study and ALLEGRO 2a safety study are both currently ongoing. Interim data from Month 24 are available from the ALLEGRO-LT study at the time of the data cut-off date (28th February 2022[19]). Interim data for the ALLEGRO-2a safety study at the time of the cut-off date (4th January 2022) are reported in the interim CSR.[21] The CS specifies that the primary completion date and study completion date are expected to be July 2024 and January 2026, respectively.[1] Some safety data from the ALLEGRO-2a safety study have been used in the pooled safety analysis.[28] The CS specifies that the primary completion date was the 4th of January 2022 and the study completion date is expected to be the 8th of January 2026.[1]

4.2.2 Details of relevant studies not included in the submission

The EAG is confident that the ALLEGRO 2b/3, ALLEGRO-LT, ALLEGRO 2a and ALLEGRO-2a studies are the only relevant studies on the effectiveness of ritlecitinib for the treatment of AA, and that no relevant studies have been omitted from the CS. Clinical advisors to the EAG were not aware of any other studies relevant to the decision problem.

4.2.3 Summary and critique of the company's quality assessment

4.2.3.1 Critical appraisal of study quality of the ALLEGRO 2b/3 study, ALLEGRO 2a proof of concept study and ALLEGRO-2a safety study

The company provided a critical appraisal of the ALLEGRO 2b/3 study using the checklist recommended by NICE, but did not provide a critical appraisal of the ALLEGRO 2a proof of concept study nor the ALLEGRO-2a safety study (see Section 4.1.5). Table 10 presents a summary of the company's assessment of study quality (focusing on risk of bias) in the ALLEGRO 2b/3 study alongside the EAG's independent quality assessment, and the EAG's assessment of study quality of the ALLEGRO 2a proof of concept study and ALLEGRO-2a safety study.

The results of the company's and the EAG's quality assessments are largely similar, with the exception of randomisation and allocation concealment, which the EAG rated as unclear due to a lack of clarity relating to how patients were identified and recruited into the study, and how the stratified randomisation took place, including who wrote the program and who undertook randomisation. Baseline characteristics of all groups were broadly similar, although there were fewer patients in the AT and prior use of corticosteroids in the combined placebo arm. The EAG concludes that the ALLEGRO 2b/3 study is at an unclear risk of bias; the company did not provide a summary appraisal of risk of bias.

Based on the information available, the EAG judged risk of bias of the ALLEGRO 2a proof of concept study and the ALLEGRO-2a safety study to be unclear, due to the lack of information available on the conduct of the study. The main sources of bias for both studies are that it is not clear how these patients were recruited or where from, and whether treatment allocation was concealed; although the CSRs[21, 26] state that interactive response technology was used for randomisation, it is not clear who conducted the randomisation or who wrote the codes.

Table 10: Quality assessment of the ALLEGRO 2b/3 study, ALLEGRO 2a proof of concept study and ALLEGRO-2a safety study (adapted from CS Appendix D, Table 10)

	ALLEGRO 2b/3		ALLEGRO 2a proof of concept		ALLEGRO-2a safety	
Criterion	**Company assessment**	**EAG assessment**	**EAG assessment**	**EAG comments**	**EAG assessment**	**EAG comments**
Was randomisation carried out appropriately?	Yes	Not clear	Not clear	Interactive response technology was used for randomisation. However, it is unclear how people were recruited.	Not clear	Intervention Model Sequential Assignment randomisation. However, is unclear how people were recruited
Was the concealment of treatment allocation adequate?	Yes	Not clear	Not clear	States that there is *"Masking: Triple (Participant, Investigator, Outcomes Assessor)"* but does not say who conducted the randomisation or who generated the codes.	Not clear	States there is *"Masking: Triple (Participant, Investigator, Outcomes Assessor)"* but does not say who conducted the randomisation or who generated the codes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes	No significant differences between groups	Yes	They appear to be
Were the care providers, participants and outcome assessor blind to treatment allocation?	Yes	Yes	Yes	Masking: Triple (Participant, Investigator, Outcomes Assessor)	Yes	Masking: Triple (Participant, Investigator, Outcomes Assessor)
Were there any unexpected imbalances to drop-outs between groups?	No	No	No	None reported	No	Trial ongoing

	ALLEGRO 2b/3		ALLEGRO 2a proof of concept		ALLEGRO-2a safety	
Criterion	**Company assessment**	**EAG assessment**	**EAG assessment**	**EAG comments**	**EAG assessment**	**EAG comments**
Is there evidence to suggest that the authors measured more outcomes than they reported?	No	No	No	Main endpoint was change from baseline of Severity of Alopecia Tool (SALT) score at Week 24. Other endpoints also measured.	No	Primary outcome was to measure functional auditory testing via the BAEP at a stimulus intensity of 80 decibels (dB) at Month 9. Other AEs and also efficacy measured
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Yes	"all randomized participants, assigned to the randomized treatment regardless of what treatment, if any, was received" - further details of ITT analysis not reported	Yes	Trial ongoing. It appears that if a participant received one or more dose, they were included in the analysis.

Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination).

4.2.3.2 Critical appraisal of study quality of the ALLEGRO-LT study

The company provided a critical appraisal of the ALLEGRO-LT study using the checklist recommended by NICE (see Section 4.1.5). Table 11 presents a summary of the study quality of the ALLEGRO-LT study alongside the EAG's independent quality assessment. The EAG has rated the ALLEGRO-LT study as moderate in terms of study quality. The main concerns are the lack of clarity relating to how the *de novo* patients were recruited, difficulty in assessing the completeness of follow-up (due to the study being ongoing at the time of writing), and the applicability of the findings to the local population, given that the majority of participants were outside of the UK. In addition, some outcomes have been stated as being assessed (in CS Appendix D[13]), but were not reported on in the interim CSR,[19] for instance, response based on achieving at least 50% and 90% improvement in SALT (SALT50 and SALT90) (see Table 7). Interim results seem plausible and blinding of treatment received in the index study was maintained for roll-over patients. Nevertheless, there is a certain amount of ambiguity around the information relating to this study which makes the quality assessment difficult.

Table 11: Quality assessment of the ALLEGRO-LT study

Criterion	Company Assessment	EAG Assessment
Did the study address a clearly focussed issue?	N/A*	Yes
Was the cohort recruited in an acceptable way?	Yes	Not clear. (No information given on recruitment).
Was the exposure accurately measured to minimise bias?	Yes	Not clear
Was the outcome accurately measured to minimise bias?	Yes	Not clear
Have the authors identified all important confounding factors?	Not clear	Not clear
Have the authors taken account of the confounding factors in the design and/or analysis?	Not clear	Not clear
Was the follow-up of patients complete?	Not clear	No – Ongoing study
Was the follow up of patients long enough?	N/A*	No – Ongoing study
What are the results of this study?	N/A*	Ongoing study. The interim results look credible.
How precise (for example, in terms of confidence intervals and p values) are the results?	No, descriptive analyses are performed only	No
Do you believe the results?	N/A*	Not clear – Ongoing study. The interim

		results appear to be credible.
Can the results be applied to the local population?	N/A*	Not clear – there were patients included from the UK (4 locations), but the majority of the sample was from other countries.
Do the results of this study fit with other available evidence?	N/A*	Not clear - Ongoing study
What are the implications of this study for practice?	N/A*	Not clear

Adapted from Critical Appraisal Skills Programme (CASP): Making sense of evidence 12 questions to help you make sense of a cohort study

* Company did not answer these questions

4.2.3.3 Protocol deviations

In the ALLEGRO 2b/3 study, protocol deviations that led to discontinuation were reported for [redacted] of the [redacted] patients (CS Appendix D, page 61). [redacted] of these patients were in the ritlecitinib 50 mg arm ([redacted] in the ritlecitinib 200/50 mg arm and [redacted] in the ritlecitinib 200/30 mg arm). No further details on these protocol deviations were provided. In Table 14.1.1.6 in the ALLEGRO 2b/3 CSR,[18] [redacted] in the ritlecitinib 200/50 mg arm did not have a clinical diagnosis of AA with no other aetiology of hair loss, and [redacted] in the ritlecitinib 200/50 mg arm, [redacted] in the ritlecitinib 50 mg arm, and [redacted] in the ritlecitinib 10 mg arm did not meet the inclusion criterion of ≥50% scalp hair loss or had evidence of regrowth within 6 months. It is unclear how many (if any) of these patients were excluded from analyses.

The number of patients with important protocol deviations in the ALLEGRO-LT study was not reported in the CS or interim CSR. The ALLEGRO-LT interim CSR reports that as of the data cut-off date, the most frequent protocol deviations were "*procedure was not done*" [redacted], "*baseline viral screen sample collected in error*" [redacted], and "*procedure was performed in error*" [redacted].[19] In Table 14.1.1.6 in the interim CSR, [redacted], all in the *de novo* population [redacted] did not meet the protocol-specific criteria for AA.[19]

In the ALLEGRO 2a proof of concept study, protocol deviations considered to be [redacted] were reported for [redacted] patients. Numbers and proportions of patients with protocol deviations were similar across the ritlecitinib 200/50 mg ([redacted]) and placebo ([redacted]) treatment arms. [redacted] patients in the 200/50 mg ritlecitinib arm and [redacted] patients in the placebo arm did not meet the criterion of having moderate to severe AA.

In the ALLEGRO-2a safety study, as of the data cut-off date, [redacted] participants enrolled had at least one protocol deviation considered to [redacted].[21] In Table 14.1.1.6 in the interim CSR,[21] [redacted] in the ritlecitinib 200/50 mg arm did not meet the inclusion criterion of having AA or had <25% hair loss due to AA by SALT, or hair loss was not assessable.

4.2.4 Summary and critique of results

4.2.4.1 SALT response

Table 12 summarises results relating to SALT response outcomes for the ALLEGRO 2b/3, ALLEGRO-LT and ALLEGRO 2a proof of concept studies. On the SALT, a lower score indicates a lower disease severity (more favourable outcome). The ALLEGRO-2a safety study results relating to the majority of the SALT response outcomes were not available to the EAG.

Table 12: Summary of results relating to SALT response outcomes in the ALLEGRO 2b/3 study, (FAS) ALLEGRO-LT study and ALLEGRO 2a proof of concept study (FAS) (adapted from CS, Table 16 and Figures 15, 23 and 24, ALLEGRO 2b/3 CSR, Tables 14.2.2.5.1.1, 14.2.2.4.1, 14.2.2.4.2, 14.2.2.5.2.1 and 14.2.3.1, ALLEGRO-LT CSR Tables 6, 7 and 8, and ALLEGRO 2a CSR, Tables 30, 31, 32 and 33)

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
SALT score ≤20 up to Week 24[c]											
Participants with SALT ≤20 Response, n	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Estimated response rate, n (%)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Difference from placebo (95% CI)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	NR	-
p-value	-		<0.001 [d]	<0.001 [d]	**<0.001 [d]**	<0.001 [d]	0.963 [d]	-	-	NR	-
SALT score ≤20 up to Week 48 [e]											
Estimated response rate, n/N (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SALT score ≤10 up to Week 24											
Participants with SALT ≤10 Response, n	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Estimated response rate, n (%)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
Difference from placebo (95% CI)								-	-	NR	-
p-value	-		<0.001	<0.001	**<0.001**	0.003	-			NR	-
SALT score ≤10 up to Week 48[e]											
Estimated response rate, n/N (%)										NR	NR
LS mean change from baseline in SALT score to Week 24											
LS mean change (95% CI)								NR	NR		
LS mean difference from placebo (95% CI)								-	-		
p-value								-	-		
LS mean change from baseline in SALT score to Week 48 [e]											
LS mean change (95% CI)								NR	NR		

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
≥30% improvement from baseline in SALT score to Week 24											
n/N (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR		
95% CI	NR	NR	NR	NR	**NR**	NR	NR	NR	NR		
Difference from placebo, % (95% CI)	-	-	NR	NR	**NR**	NR	NR	NR	NR		
p-value	-	-	NR	NR	**NR**	NR	NR	NR	NR		
≥50% improvement from baseline in SALT score to Week 24											
n/N (%)								NR	NR		
95% CI								NR	NR		
Difference from placebo, % (95% CI)								-	-		
p-value								-	-		

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
≥50% improvement from baseline in SALT score to Week 48[e]											
n/N (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR
≥75% improvement from baseline in SALT score to Week 24											
n/N (%)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference from placebo, % (95% CI)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	[redacted]	[redacted]
p-value	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	[redacted]	[redacted]
≥75% improvement from baseline in SALT score to Week 48[e]											
n/N (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
≥90% improvement from baseline in SALT score to Week 24											
n/N (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]
95% CI	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]
Difference from placebo, % (95% CI)	-	-	NR	NR	**NR**	NR	NR	-	-	[redacted]	[redacted]
p-value	-	-	NR	NR	**NR**	NR	NR	-	-	[redacted]	[redacted]

Note: Denominators refer to the number of participants with valid data at the analysis visit (non-response for missing due to reasons unrelated to COVID-19). Data in bold refers to the anticipated licensed dose of ritlecitinib (50 mg once daily).

CI - confidence intervals; FAS - full analysis set; LS - least squared; NR - not reported; SALT - Severity of Alopecia Tool.

[a] *Combined placebo arms for 24-week double-blind period*

[b] *Interim analysis*

[c] *A generalised linear mixed effect model without imputation using observed data up to Week 24 was used as the imputation model, assuming missing at random under a Bayesian framework with non-informative/weakly informative prior densities and Markov chain Monte Carlo. For a participant with missing SALT score at Week 24 due to any reason, imputation was done based on predictive Bernoulli distribution with a probability equal to the probability under MAR calculated using the sampled parameters. MN method was used for the calculation of 95% Cis and p-values for testing the difference in the proportion of response between each treatment group and placebo.[c] Based on the Cochran-Mantel-Haenszel test adjusting for stratification factor (PFIC type)*

[d] *Confidence interval and p-value are calculated using Miettinen and Nurminen method. Missing data due to COVID-19 was excluded from this analysis, whereas participants with missing data due to other reasons were considered as non-responders*

[e] *Patients on placebo switched to active treatment at week 24*

[f] *Calculated using Mixed Model Repeated Measure (MMRM) containing fixed factors of treatment, week, baseline, treatment by week and treatment by baseline interaction and a random effect for participant. Unstructured matrix was used to model the covariance structure.*

[g] *90% CI.*

SALT score ≤20

The proportion of patients with a SALT score of ≤20 at Week 24 was significantly greater in the ritlecitinib arms than in the combined placebo arm, in the ALLEGRO 2b/3 study (see Table 12 and

Figure 3). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had a SALT score of ≤20, compared with [redacted] in the placebo arms combined, which represented a statistically significantly difference of [redacted] (p<0.001). By Week 48, the proportion of patients with a SALT score of ≤20 in the ritlecitinib 50 mg arm had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib (200/50 mg) at 24 weeks. The company's clarification response[2] (question A11) gives some insight into the [redacted] patients from the placebo arm who achieved a SALT score of ≤20 at Week 24: "[redacted] *had a baseline SALT* [redacted]*, AA duration since first diagnosis of* [redacted] *years and duration of current episode was* [redacted]*. The* [redacted] *had a baseline SALT score of* [redacted]*, AA duration since first diagnosis of* [redacted] *years and duration of current episode was* [redacted] *years. For this participant, the duration of current episode is longer than the duration since first diagnosis because the formal diagnosis came after the start of the current episode.*" In the latter case, the EAG does not regard it possible for [redacted] to have met the inclusion criterion of no regrowth within 6 months (as they would not have had a sufficient duration within which to demonstrate this). In addition, clinical advisors to the EAG have stated that regrowth is very rare in untreated patients who have had AA for > 6 months, and this participant has had AA for [redacted] months. The EAG note that the proportions of patients achieving a SALT score of ≤20 at Week 24 for each study arm as presented in the CS, Table 16, differ slightly from the equivalent figures relating to the FAS, as presented in the CSR,[18] Table 14.2.1.5.

In the ALLEGRO-LT study, [redacted] of *de novo* participants had a SALT score of ≤20 at Week 24 (Month 6) (with a dose of 200 mg once daily for four weeks, followed by 50 mg once daily subsequently) in an interim analysis, and by Week 48 (Month 12), the proportion was [redacted] (see Table 12 and Figure 4). A similar proportion ([redacted]) of roll-over participants had a SALT score of ≤20 at Week 24 (Month 6), and by Week 48 (Month 12), the proportion was [redacted].[19] This outcome was not assessed in the ALLEGRO 2a proof of concept study.

Figure 3: Response based on SALT ≤20 up to Week 48 (FAS) in the ALLEGRO 2b/3 study (reproduced from CS, Figure 13)

FAS - full analysis set; mg - milligram; Pbo - placebo; SALT - Severity of Alopecia Tool
Treatment group listed as loading dose (if applicable)/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Participants in the Pbo→200/50 mg group received 24 weeks of placebo, and then switched to 200/50 mg ritlecitinib, while in the Pbo→50 mg group participants received 24 weeks of placebo and then switched to 50 mg.

Figure 4: Response based on SALT ≤20 up to Month 24 (interim analysis, *de novo* cohort) in the ALLEGRO-LT study (reproduced from CS, Figure 23)

CI - confidence interval; QD - once daily; SALT - Severity of Alopecia Tool.
Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤20. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.
Source: Sinclair R, et al. (EADV 2022).[29]

SALT score ≤10

The proportion of patients with a SALT score of ≤10 at Week 24 was significantly greater in the ritlecitinib arms than in the combined placebo arm, in the ALLEGRO 2b/3 study (see Table 12 and

Figure 5). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had a SALT score of ≤10, compared with [redacted] in the placebo arms combined, which represented a statistically significantly difference of [redacted] (p<0.001). By Week 48, the proportion of patients in the ritlecitinib 50 mg arm with a SALT score of ≤10 had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib (200/50 mg) at 24 weeks. As for the outcome SALT score ≤20, the company's clarification response[2] (question A11) reports that of the [redacted] patients from the placebo arm who achieved a SALT score of ≤10 at Week 24, [redacted] had a baseline SALT [redacted], AA duration since first diagnosis of [redacted] years and duration of current episode of [redacted], whereas the [redacted] had a baseline SALT score of [redacted], AA duration since first diagnosis of [redacted] years and duration of current episode of [redacted] years. In the latter case, the EAG does not regard it possible for [redacted] to have met the inclusion criterion of no regrowth within 6 months (as this participant has had AA for [redacted] months), and clinical advisors to the EAG have stated that regrowth is very rare in untreated patients who have had AA for > 6 months. The EAG note that the proportions of patients achieving a SALT score of ≤10 at Week 24 for each study arm as presented in the CS, Table 16, differ slightly from the equivalent figures relating to the FAS, as presented in the CSR,[18] Table 14.2.1.5.

In the ALLEGRO-LT study, [redacted] of *de novo* participants had a SALT score of ≤10 at Week 24 (Month 6) (with a dose of 200 mg once daily for four weeks, followed by 50 mg once daily subsequently) in an interim analysis, and by Week 48 (Month 12), the proportion was [redacted] (see Table 12 and Figure 6). A similar proportion ([redacted]) of roll-over participants had a SALT score of ≤10 at Week 24 (Month 6), and by Week 48 (Month 12), the proportion was [redacted].[19] This outcome was not assessed in the ALLEGRO 2a proof of concept study.

Figure 5: Response based on SALT ≤10 up to Week 48 (FAS) in the ALLEGRO 2b/3 study (reproduced from CS, Figure 14)

FAS - full analysis set; mg - milligram; Pbo - placebo; SALT - Severity of Alopecia Tool
Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomisation

Figure 6: Response based on SALT ≤10 up to Month 24 (interim analysis, *de novo* cohort) in the ALLEGRO-LT study (reproduced from CS, Figure 24)

CI - confidence interval; QD - once daily; SALT - Severity of Alopecia Tool.
Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤10. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.
Source: Sinclair R, et al. (EADV 2022).[29]

Change from baseline in SALT score

The least squares mean (LSM) change (improvement) from baseline in SALT score at Week 24 was significantly greater in each of the ritlecitinib arms (with the exception of the ritlecitinib 10 mg arm) than in the combined placebo arm, in the ALLEGRO 2b/3 study (see Table 12 and Figure 7). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, the LSM change from baseline in SALT score was [redacted], compared with [redacted] in the placebo arms combined, which represented a statistically significantly difference of [redacted] (p<0.001). By Week 48, the LSM change from baseline in SALT score was [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib at 24 weeks.

In the ALLEGRO 2a proof of concept study, the LSM change (improvement) from baseline in SALT score at Week 24 was significantly greater ([redacted]) in the ritlecitinib 200/50 mg arm [redacted] than in the placebo arm [redacted].

In the ALLEGRO-2a safety study, the LSM change from baseline (improvement) in SALT score at Month 6 was greater in the ritlecitinib 200/50 mg arm ([redacted]) than in the placebo arm ([redacted]) (see CSR, Table 7).[21]

Figure 7: Least squared means of absolute change from baseline in SALT score for initial active groups up to Week 48 (FAS) in the ALLEGRO 2b/3 study (reproduced from CS, Figure 15)

[redacted]

Treatment group listed as loading dose [if applicable]/maintenance dose. Participants were randomised to one of seven groups, 200/50 mg, 200/30 mg, 50 mg, 30 mg, 10 mg, Pbo→200/50 mg and Pbo→50 mg. Please refer to Figure 13 footnote for information on randomization.
Abbreviations: FAS, full analysis set; mg, milligram; Pbo, placebo; SALT, Severity of Alopecia Tool

Improvement in SALT score

Improvement of ≥30% in SALT score from baseline to Week 24 was not reported on in the ALLEGRO 2b/3 study nor in the ALLEGRO-LT study. In the ALLEGRO 2a proof of concept study, the proportion of patients with ≥30% improvement in SALT score from baseline to Week 24 was significantly greater ([redacted]) in the ritlecitinib 200/50 mg arm ([redacted]) compared with the placebo arm ([redacted]).

The proportion of patients with ≥50% improvement in SALT score from baseline to Week 24 was significantly greater in the ritlecitinib arms than in the placebo arms, with the exception of the 10 mg dose, in the ALLEGRO 2b/3 study (see Table 12). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had ≥50% improvement in SALT score, compared with [redacted] in the placebo 50 mg arm. By Week 48, the proportion of patients with ≥50% improvement in SALT score had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib at 24 weeks. This outcome was not assessed in the ALLEGRO-LT study. In the ALLEGRO 2a proof of concept study, the proportion of patients with ≥50% improvement in SALT score from baseline to Week 24 was significantly greater ([redacted]) in the ritlecitinib 200/50 mg arm ([redacted]) compared with the placebo arm ([redacted]).

The proportion of patients with ≥75% improvement in SALT score from baseline to Week 24 was significantly greater in the ritlecitinib arms than in the placebo arms, with the exception of the 10 mg dose, in the ALLEGRO 2b/3 study (see Table 12). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had ≥75% improvement in SALT score, compared with [redacted] in the combined placebo arms, which represented a statistically significantly difference ([redacted]). By Week 48, the proportion of patients with ≥75% improvement in SALT score had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib at 24 weeks. In the ALLEGRO-LT study, the proportion of patients with ≥75% improvement in SALT score from baseline to Week 24 (Month 6) was greater than those in the ALLEGRO 2b/3 study in both the *de novo* ([redacted]) and roll-over ([redacted]) patients. This may potentially be because the *de novo* cohort included those with both moderate and severe AA, and the roll-over patients had received up to 48 months of prior ritlecitinib.[19] By Week 48 (Month 12), the proportion of patients with ≥75% improvement in SALT score had reached [redacted] in the *de novo* cohort and [redacted] in the roll-over cohort.[19] In the ALLEGRO 2a proof of concept study, the proportion of patients with ≥75% improvement in SALT score from

baseline to Week 24 was significantly greater ([redacted]) in the ritlecitinib 200/50 mg arm ([redacted]) compared with the placebo arm ([redacted]).

Improvement of ≥90% in SALT score from baseline to Week 24 was not reported on in the ALLEGRO 2b/3 study nor in the ALLEGRO-LT study. In the ALLEGRO 2a proof of concept study, the proportion of patients with ≥90% improvement in SALT score from baseline to Week 24 was significantly greater ([redacted]) in the ritlecitinib 200/50 mg arm ([redacted]) compared with the placebo arm ([redacted]).

4.2.4.2 Eyebrow and eyelash assessment

Table 13 summarises results relating to eyebrow assessment (EBA) and eyelash assessment (ELA) scores for the ALLEGRO 2b/3, ALLEGRO-LT and ALLEGRO 2a proof of concept studies.

Table 13: Summary of results relating to EBA and ELA response outcomes in the ALLEGRO 2b/3 study (FAS), ALLEGRO-LT study and ALLEGRO 2a proof of concept study (FAS) (adapted from CS, Table 18, ALLEGRO-LT CSR, Tables 10 and 11 and ALLEGRO 2a CSR, Table 14.2.5.2.1)

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo n=131**		**Ritlecitinib**					**Ritlecitinib[a]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
EBA response (≥2-grade improvement from baseline/score of 3) at Week 24											
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Difference from placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	NR	NR
p-value	-	-	<0.001	<0.001	**<0.01**	0.005	0.368	-	-	NR	NR
EBA response (≥2-grade improvement from baseline/score of 3) at Week 48											
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EBA response (≥1-grade improvement from baseline) at Week 24											
Estimated response rate (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]
Difference from placebo (95% CI)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	-
p-value	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	-

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept	
Outcome	**Placebo n=131**		**Ritlecitinib**					**Ritlecitinib[a]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**		
ELA response (≥2-grade improvement from baseline/score of 3) at Week 24											
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Difference from placebo (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-	-	NR	NR
p-value	-	-	<0.001	0.001	**<0.001**	<0.01	0.946	-	-	NR	NR
ELA response (≥2-grade improvement from baseline/score of 3) at Week 48											
Estimated response rate (%)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR

Note: Denominators refer to the number of participants with valid data at the analysis visit (non-response for missing due to reasons unrelated to COVID-19). Data in bold refers to the anticipated licensed dose of ritlecitinib (50 mg once daily).

CI - confidence intervals; EBA - eyebrow assessment; ELA - eyelash assessment; FAS - full analysis set; NR - not reported.

[a] *Interim analysis*

[b] *a generalized linear mixed effect model without imputation using observed data up to Week 24 was used as the imputation model. Estimation of model parameters was performed assuming MAR under a Bayesian framework with non-informative/weakly informative prior densities and Markov chain Monte Carlo. For a participant with missing SALT score at Week 24 due to COVID-19 related reasons, imputation was done based on predictive Bernoulli distribution with a probability equal to the probability under MAR calculated using the sampled parameters. Participants with missing SALT score at Week 24 due to reasons other than COVID-19 were considered non-responders. A single complete imputed data set for Week 24 was then analysed using the Miettinen method as the analysis model.*

The proportion of patients with an EBA response (≥2-grade improvement from baseline/score of 3) at Week 24 was greater in the ritlecitinib arms than in the placebo arms, with the exception of the 10 mg dose, in the ALLEGRO 2b/3 study (see Table 13). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had an EBA response, compared with [redacted] in the placebo 50 mg arm and [redacted] in the placebo 200/50 mg arm. By Week 48, the proportion of patients with EBA response in the ritlecitinib 50 mg arm had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib at 24 weeks. In the ALLEGRO-LT study, the proportion of patients with an EBA response (≥2-grade improvement from baseline/score of 3) at Week 24 (Month 6) was greater than those in the ALLEGRO 2b/3 study in both the *de novo* ([redacted]) and roll-over ([redacted]) patients. This may potentially be because the *de novo* cohort included those with both moderate and severe AA, and the roll-over patients had received up to 48 months of prior ritlecitinib.[19] By Week 48 (Month 12), the proportion of patients with an EBA response had reached [redacted] in the *de novo* cohort and [redacted] in the roll-over cohort.[19] In the ALLEGRO 2a proof of concept study, EBA response was categorised differently to the definition used in the ALLEGRO 2b/3 and ALLEGRO-LT studies, and thus the results are not directly comparable. The proportion of patients with an EBA response (≥1-grade improvement from baseline) at Week 24 was significantly greater ([redacted]) in the ritlecitinib 200/50 mg arm ([redacted]) compared with the placebo arm ([redacted]).

The proportion of patients with an ELA response (≥2-grade improvement from baseline/score of 3) at Week 24 was greater in the ritlecitinib arms than in the placebo arms, with the exception of the 10 mg dose, in the ALLEGRO 2b/3 study (see Table 13). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had an ELA response, compared with [redacted] in the placebo 50 mg arm and [redacted] in the placebo 200/50 mg arm. By Week 48, the proportion of patients with ELA response in the ritlecitinib 50 mg arm had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib at 24 weeks. In the ALLEGRO-LT study, the proportion of patients with an ELA response (≥2-grade improvement from baseline/score of 3) at Week 24 (Month 6) was greater than those in the ALLEGRO 2b/3 study in both the *de novo* ([redacted]) and roll-over ([redacted]) patients. This may potentially be because the *de novo* cohort included those with both moderate and severe AA, and the roll-over patients had received up to 48 months of prior ritlecitinib.[19] By Week 48 (Month 12), the proportion of patients with an ELA response had reached [redacted] in the *de novo* cohort and [redacted] in the roll-over cohort.[19] This outcome was not assessed in the ALLEGRO 2a proof of concept study.

4.2.4.3 Clinician and patient global assessment of change

Table 14 summarises results relating to clinician and patient global assessment of change relating to hair regrowth, for the ALLEGRO 2b/3, ALLEGRO-LT, ALLEGRO 2a proof of concept and ALLEGRO-2a safety studies.

Table 14: Summary of results relating to clinician and patient global assessment of change in the ALLEGRO 2b/3 study (FAS), ALLEGRO-LT study (FAS), ALLEGRO 2a proof of concept study (FAS) and ALLEGRO-2a safety study (EAS) (adapted from CS, Table 17, ALLEGRO-LT CSR, Table 12, ALLEGRO 2a proof of concept study CSR, Table 14.2.3.1, and ALLEGRO-2a safety study CSR, Table 8)

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept		ALLEGRO-2a safety study	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritleci-tinib 200/50 mg n=48**	**Placebo (comb-ined) n=47**	**Ritleci-tinib 200/50 mg n=36**	**Placebo n=35**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**				
PGI-C response, Week 24													
Participants with PGI-C response,[c] n										NR	NR		
Estimated response rate (%)		[d]	[d]	[d]	[d]	[d]	[d]			NR	NR		
Difference from placebo (95% CI)			[d]	[d]	[d]	[d]	[d]	-	-	NR	NR		-
p-value	-		<0.001	<0.001	**<0.001**	<0.001	-	-	-	NR	NR		-
PGI-C response, Week 48													
n/N (%)										NR	NR	NR	NR
95% CI[f]										NR	NR	NR	NR
IGA, Week 24													
No change or further loss, n (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR			NR	NR
1-24% regrowth, n (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR			NR	NR

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept		ALLEGRO-2a safety study	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b]**		**Ritleci-tinib 200/50 mg n=48**	**Placebo (comb-ined) n=47**	**Ritleci-tinib 200/50 mg n=36**	**Placebo n=35**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=132**	**200/30 mg n=130**	**50 mg n=130**	**30 mg n=132**	**10 mg n=63**	***De novo* cohort 200/50 mg n=449**	**Roll-over cohort 50 mg N=603**				
25-49% regrowth, n (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]	NR	NR
50-74% regrowth, n (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]	NR	NR
75-99% regrowth, n (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]	NR	NR
100% regrowth, n (%)	NR	NR	NR	NR	**NR**	NR	NR	NR	NR	[redacted]	[redacted]	NR	NR

Note: Denominators refer to the number of participants with valid data at the analysis visit (non-response for missing due to reasons unrelated to COVID-19). Data in bold refers to the anticipated licensed dose of ritlecitinib (50 mg once daily).

CI - confidence intervals; EAS - efficacy analysis set; FAS - full analysis set; IGA - Investigator Global Assessment (of change); NR - not reported; PGI-C - patient's global impression of change.

[a] *Combined placebo arms for 24-week double-blind period*

[b] *Interim analysis*

[c] *PGI response is defined as a score of 'moderately improved' or 'greatly improved' relative to baseline*

[d] *a generalized linear mixed effect model without imputation using observed data up to Week 24 was used as the imputation model. Estimation of model parameters was performed assuming MAR under a Bayesian framework with non-informative/weakly informative prior densities and Markov chain Monte Carlo. For a participant with missing SALT score at Week 24 due to COVID-19 related reasons, imputation was done based on predictive Bernoulli distribution with a probability equal to the probability under MAR calculated using the sampled parameters. Participants with missing SALT score at Week 24 due to reasons other than COVID-19 were considered non-responders. A single complete imputed data set for Week 24 was then analysed using the Miettinen method as the analysis model.*

[e] *95% confidence interval is calculated based on Chan and Zhang's exact method*

[f] *Interval is calculated based on normal approximation*

The proportion of patients with a patient's global impression of change (PGI-C) response (defined as a score of 'moderately improved' or 'greatly improved', relative to baseline) at Week 24 was significantly greater in the ritlecitinib arms than in the placebo arms, with the exception of the 10 mg dose, in the ALLEGRO 2b/3 study (see Table 14). In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had a PGI-C response, compared with [redacted] in the combined placebo arms, which represented a statistically significantly difference of [redacted] ([redacted]). By Week 48, the proportion of patients with a PGI-C response in the ritlecitinib 50 mg arm had reached [redacted], although comparison with a no-treatment control condition was not possible due to patients in the placebo arms switching to ritlecitinib at 24 weeks. In the ALLEGRO-LT study, the proportion of patients with a PGI-C response at Week 24 (Month 6) was greater than those in the ALLEGRO 2b/3 study in both the *de novo* ([redacted]) and roll-over ([redacted]) patients. This may potentially be because the *de novo* cohort included those with both moderate and severe AA, and the roll-over patients had received up to 48 months of prior ritlecitinib, or it could be because the study was open-label, and patients were aware that they were taking active treatment.[19] By Week 48 (Month 12), the proportion of patients with a PGI-C response had reached [redacted] in the *de novo* cohort and [redacted] in the roll-over cohort.[19] The ALLEGRO 2a proof of concept study did not assess PGI-C response. In the ALLEGRO-2a safety study, the proportion of patients with a PGI-C response at Week 24 (Month 6) was greater in the ritlecitinib 200/50 mg arm ([redacted]) compared with the placebo arm ([redacted]), although it is not reported whether this difference was statistically significant.

Investigator Global Assessment (IGA) of change was only reported for the ALLEGRO 2a proof of concept study (see Table 14). The proportion of patients with an IGA rating of no change or further loss at Week 24 was greater in the combined placebo arm ([redacted]) than the ritlecitinib 200/50 mg arm ([redacted]), whereas the proportion of patients with an IGA rating of 25-49% regrowth ([redacted]), 50-74% regrowth ([redacted]), 75-99% regrowth ([redacted]) and 100% regrowth ([redacted]) at Week 24 was greater than for the combined placebo arm ([redacted]), and the proportion of patients with an IGA rating of 1-24% regrowth was greater for the combined placebo arm ([redacted]) than for the ritlecitinib 200/50 mg arm ([redacted]).

4.2.4.4 Health-related quality of life

Table 15 summarises results relating to HRQoL, for the ALLEGRO 2b/3 and ALLEGRO-LT studies.

Table 15: Summary of results relating to HRQoL in the ALLEGRO 2b/3 study (FAS) and ALLEGRO-LT study (FAS) (adapted from CS, Tables 41, 42, 43, 44, ALLEGRO 2b/3 CSR, Tables 14.2.3.3.1.1, 14.2.3.3.2.1 and 14.2.2.8.1.2, and ALLEGRO-LT CSR, Table 14.2.1.5.1)

Outcome	ALLEGRO 2b/3							ALLEGRO-LT	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b]	
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=132	200/30 mg n=130	50 mg n=130	30 mg n=132	10 mg n=63	*De novo* cohort 200/50 mg n=449	Roll-over cohort 50 mg N=603
EQ-5D-5L Index Value, Adults, baseline									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EQ-5D-5L Index Value, Adults, Week 24									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EQ-5D-5L Index Value, Adults, Week 48									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EQ-5D-5L VAS, Adults, baseline									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR

Outcome	ALLEGRO 2b/3							ALLEGRO-LT	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b]	
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=132	200/30 mg n=130	50 mg n=130	30 mg n=132	10 mg n=63	*De novo* cohort 200/50 mg n=449	Roll-over cohort 50 mg N=603
EQ-5D-5L VAS, Adults, Week 24									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EQ-5D-5L VAS, Adults, Week 48									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EQ-5D-5L VAS, Adolescents, baseline									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
EQ-5D-5L VAS, Adolescents, Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)									
EQ-5D-5L VAS, Adolescents, Week 48									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SF-36 physical component summary scores, baseline									

Outcome	ALLEGRO 2b/3							ALLEGRO-LT	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b]	
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=132	200/30 mg n=130	50 mg n=130	30 mg n=132	10 mg n=63	*De novo* cohort 200/50 mg n=449	Roll-over cohort 50 mg N=603
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SF-36 physical component summary scores, Week 24									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SF-36 physical component summary scores, Week 48									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SF-36 mental component summary scores, baseline									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SF-36 mental component summary scores, Week 24									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
SF-36 mental component summary scores, Week 48									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR

Outcome	ALLEGRO 2b/3							ALLEGRO-LT	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b]	
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=132	200/30 mg n=130	50 mg n=130	30 mg n=132	10 mg n=63	*De novo* cohort 200/50 mg n=449	Roll-over cohort 50 mg N=603
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR
HADS-D absolute scores, baseline									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS-D absolute scores, Week 24									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS-D absolute scores, Week 48									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS-A absolute scores, baseline									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HADS-A absolute scores, Week 24									
Mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (IQR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Outcome	ALLEGRO 2b/3							ALLEGRO-LT	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b]	
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=132	200/30 mg n=130	50 mg n=130	30 mg n=132	10 mg n=63	*De novo* cohort 200/50 mg n=449	Roll-over cohort 50 mg N=603
HADS-A absolute scores, Week 48									
Mean (SD)									
Median (IQR)									
AAPPO Emotional Symptoms scores, baseline									
Mean (SD)									
Median (IQR)									
AAPPO Emotional Symptoms scores, Week 24									
Mean (SD)									
Median (IQR)									
AAPPO Emotional Symptoms scores, Week 48									
Mean (SD)	NR		NR	NR	**NR**	NR	NR		
Median (IQR)	NR		NR	NR	**NR**	NR	NR		
AAPPO Activity Limitations scores, baseline									
Mean (SD)									
Median (IQR)									
AAPPO Activity Limitations scores, Week 24									

Outcome	ALLEGRO 2b/3							ALLEGRO-LT	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b]	
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=132	200/30 mg n=130	50 mg n=130	30 mg n=132	10 mg n=63	*De novo* cohort 200/50 mg n=449	Roll-over cohort 50 mg N=603
Mean (SD)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median (IQR)	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AAPPO Activity Limitations scores, Week 48									
Mean (SD)	NR	NR	NR	NR	**NR**	NR	NR	[redacted]	[redacted]
Median (IQR)	NR	NR	NR	NR	**NR**	NR	NR	[redacted]	[redacted]

Note: Based on number of adult patients with observed data. Data in bold refers to the anticipated licensed dose of ritlecitinib (50 mg once daily).
AAPPO - Alopecia Areata Patient Priority Outcomes; CI - confidence intervals; EQ-5D-5L - EuroQol 5-Dimensions-5 levels; FAS - full analysis set; HADS-A - Hospital Anxiety and Depression Scale – Anxiety subscale; HADS-D - Hospital Anxiety and Depression Scale – Depression subscale; IGA - Investigator Global Assessment (of change); IQR - inter-quartile range; NR - not reported; PGI-C - patient's global impression of change; SD - standard deviation; SF-36 - Short Form 36 items; VAS - visual analogue scale.
[a] *Combined placebo arms for 24-week double-blind period*
[a] *Combined placebo arms for 24-week double-blind period*
[b] *Interim analysis*
[c] *Median (range)*

The CS[1] reports very little change in HRQoL scores across a range of measures from baseline to Week 24 and Week 48 (e.g., see CS, Table 39) and no clinically meaningful changes, with baseline HRQoL scores generally indicating that study participants had good HRQoL at the start of the study. This is reflected in the HRQoL scores presented in Table 15, which are drawn from the CS[1] and the ALLEGRO 2/3[18] and ALLEGRO-LT[19] CSRs. The greatest difference was seen on the emotional symptoms subscale of the AAPPO, where the CS reports the largest difference between groups across time, and an effect size estimate of *d*=0.79 (large effect). Patients with depression and suicide ideation were excluded from the ALLEGRO 2b/3 study, which may at least partially explain the apparent ceiling effect. Table 10 of the CS[1] indicates that at baseline, patients in the ritlecitinib 50 mg arm and pooled placebo arm entered the ALLEGRO 2b/3 study at a mean of [redacted] years and [redacted] years, respectively, after their initial diagnosis and their current episode of AA had lasted a mean of [redacted] and [redacted] years, respectively. Clinical advice received by the EAG suggests that patients with AA tend to be more distressed at the onset of the condition and after a longer duration may have developed some acceptance for the condition. Clinicians also advised that patients who were receiving active treatment may also be feeling hopeful.

4.2.4.5 Safety and tolerability

The CS[1] defined a treatment-emergent adverse event (TEAE) as "*any untoward medical occurrence which emerged or worsened during the treatment period but these were not necessarily causally related to treatment unless classified as treatment-related*" (page 103). A serious adverse event (SAE) was defined as "*any untoward medical occurrence at any dose that results in death, is life-threatening (immediate risk of death), requires inpatient hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability/incapacity (substantial disruption of the ability to conduct normal life functions), results in congenital anomaly/birth defect, and is an important medical event based on investigator's judgment*" (page 103). The intensity of adverse events (AEs) was graded as mild, moderate or severe based on how much it interferes with a patient's usual function: "*A mild AE does not interfere with patient's usual function whereas a moderate AE interferes to some extent. A severe AE is one that interferes significantly with a patient's usual function although it is not necessarily an SAE. For example, a headache may be severe (interferes significantly with the patient's usual function) but would not be classified as serious unless it met one of the criteria for SAEs.*" (page 103). The company's clarification response[2] (question A27) defines a treatment-related adverse event (TRAE) as "*any untoward medical occurrence which emerged or worsened during the treatment period that were causally related to treatment*". No detail was given as to who made this decision or how causality was determined.

Safety analyses for the ALLEGRO 2b/3 study were conducted using the SAS, defined as all patients who received at least one dose of study intervention, classified as the actual intervention received for

most of the time during the study.[18] Safety analyses for the ALLEGRO-LT study were conducted using the SAS, defined as all patients who took at least one dose of study intervention.[19] Safety analyses for the ALLEGRO 2a proof of concept study were conducted using the SAS, defined as all patients who received at least one dose of study intervention, classified according to the actual treatment received.[26] Safety analyses for the ALLEGRO-2a safety study were conducted using the SAS, defined as all patients who received at least one dose of study intervention, classified according to the actual treatment received.[21] The proportions of patients experiencing TEAEs, SAEs, severe TEAEs, TRAEs, and severe and serious TRAEs were generally comparable across study arms, with no consistent differences between ritlecitinib and placebo arms (see Table 16). Generally, ritlecitinib appeared to be well tolerated. In the ALLEGRO 2b/3 study, [redacted] severe TEAEs were reported in the ritlecitinib 50 mg arm to Week 48: [redacted].[1] [redacted] serious TEAEs were reported in the ritlecitinib 50 mg arm up to Week 48, [redacted] of which [redacted] considered to be treatment-related ([redacted]); the other was [redacted].[1] All SAEs are listed in the CS, Appendix F, Table 2.[28]

Audiological adverse events

In the ALLEGRO 2b/3 study, [redacted] participants met the criteria for sensorineural hearing loss (not consistent with central hearing disorder and no participants discontinued the study as a result).[1] Audiologic events were identified during routine audiologic testing throughout the study, and [redacted] TEAEs relating to hearing loss were spontaneously reported by participants.[28] [redacted] of the [redacted] participants with audiological TEAEs was in the ritlecitinib 50 mg treatment arm, and [redacted]; the others were in the 200/50 mg arm ([redacted]), the 200/30 mg arm ([redacted]), and the 30 mg arm ([redacted]).[28] In the ALLEGRO-LT study, [redacted] participants had TEAEs associated with hearing loss considered to be sensorineural hearing loss by an external Neurosafety Event Adjudication Committee ([redacted]); [redacted] audiological TEAEs were identified through protocol specified audiologic testing and [redacted] reported by a participant.[19] Audiologic TEAEs were not reported in the ALLEGRO 2a proof of concept study (with the exception of one participant in the placebo arm),[26] nor in the ALLEGRO-2a safety study.[21]

Serious infections

In the ALLEGRO 2b/3 study, [redacted] serious infections occurred in [redacted] patients ([redacted]) treated with ritlecitinib, none of whom were in the ritlecitinib 50 mg treatment arm.[1] In the ALLEGRO-LT study, [redacted] participants experienced a serious infection: [redacted]

[redacted].[19] Serious infections were not reported in the ALLEGRO 2a proof of concept study,[26] nor in the ALLEGRO-2a safety study.[21]

Table 16: Summary of adverse events in the ALLEGRO 2b/3 study (SAS), ALLEGRO-LT study (SAS), ALLEGRO 2a proof of concept study (SAS) and ALLEGRO-2a safety study (SAS) (adapted from CS, Tables 25, 27 and 29, ALLEGRO 2b/3 CSR, Tables 20, 21, 25 and 26, ALLEGRO-LT CSR, Tables 18 and 19, ALLEGRO 2a proof of concept study CSR, Tables 54 and 55, and ALLEGRO-2a safety study CSR, Tables 10 and 11)

	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept		ALLEGRO-2a safety study	
Outcome	**Placebo[a] n=131**		**Ritlecitinib**					**Ritlecitinib[b,c]**		**Ritlecitinib 200/50 mg n=48**	**Placebo (combined) n=47**	**Ritlecitinib[b] 200/50 mg n=36**	**Placebo[b] n=35**
	Placebo 200/50 mg n=65	**Placebo 50 mg n=66**	**200/50 mg n=131**	**200/30 mg n=129**	**50 mg n=130**	**30 mg n=132**	**10 mg n=62**	***De novo* cohort 200/50 mg n=447**	**Roll-over cohort 50 mg N=603**				
TEAEs up to Week 24, n (%)													
Number of TEAEs								NR	NR				
Participants with TEAE								NR	NR				
Participants with SAE								NR	NR				
Participants with severe AE								NR	NR				
Number of TRAEs	86		94	82	**83**	97	38	NR	NR				
Participants with TRAE								NR	NR				
Participants with serious TRAE								NR	NR				
Participants with severe TRAE								NR	NR				
Patients who discontinued due to TEAEs, n (%)								NR	NR				
Patients with a temporary study drug discontinuation due to TEAEs, n (%)								NR	NR				

Outcome	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept		ALLEGRO-2a safety study	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b,c]		Ritlecitinib 200/50 mg n=48	Placebo (combined) n=47	Ritlecitinib[b] 200/50 mg n=36	Placebo[b] n=35
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=131	200/30 mg n=129	50 mg n=130	30 mg n=132	10 mg n=62	*De novo* cohort 200/50 mg n=447	Roll-over cohort 50 mg N=603				
Patients who discontinued due to TRAEs, n (%)								NR	NR				
Patients with a temporary study drug discontinuation due to TRAEs, n (%)								NR	NR				
Total number of deaths	0		0	0	**0**	0	0	NR	NR				
TEAEs up to Week 48, n (%)													
Number of TEAEs								NR	NR	NR	NR	NR	NR
Participants with TEAE								NR	NR	NR	NR	NR	NR
Participants with SAE								NR	NR	NR	NR	NR	NR
Participants with severe AE								NR	NR	NR	NR	NR	NR
Number of TRAEs								NR	NR	NR	NR	NR	NR
Participants with TRAE								NR	NR	NR	NR	NR	NR
Participants with serious TRAE								NR	NR	NR	NR	NR	NR
Participants with severe TRAE								NR	NR	NR	NR	NR	NR
Patients who discontinued due to TEAEs, n (%)								NR	NR	NR	NR	NR	NR

Outcome	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept		ALLEGRO-2a safety study	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b,c]		Ritlecitinib 200/50 mg n=48	Placebo (combined) n=47	Ritlecitinib[b] 200/50 mg n=36	Placebo[b] n=35
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=131	200/30 mg n=129	50 mg n=130	30 mg n=132	10 mg n=62	*De novo* cohort 200/50 mg n=447	Roll-over cohort 50 mg N=603				
Patients with a temporary study drug discontinuation due to TEAEs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR	NR	NR
Patients who discontinued due to TRAEs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR	NR	NR
Patients with a temporary study drug discontinuation due to TRAEs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR	NR	NR
Total number of deaths	0	0	0	0	**0**	0	0	NR	NR	NR	NR	NR	NR
TEAEs over study duration n (%)													
Number of TEAEs	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Participants with TEAE	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Participants with SAE	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Participants with severe AE	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Number of TRAEs	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Participants with TRAE	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Participants with serious TRAE	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A

Outcome	ALLEGRO 2b/3							ALLEGRO-LT		ALLEGRO 2a proof of concept		ALLEGRO-2a safety study	
	Placebo[a] n=131		Ritlecitinib					Ritlecitinib[b,c]		Ritlecitinib 200/50 mg n=48	Placebo (combined) n=47	Ritlecitinib[b] 200/50 mg n=36	Placebo[b] n=35
	Placebo 200/50 mg n=65	Placebo 50 mg n=66	200/50 mg n=131	200/30 mg n=129	50 mg n=130	30 mg n=132	10 mg n=62	*De novo* cohort 200/50 mg n=447	Roll-over cohort 50 mg N=603				
Participants with severe TRAE	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Patients who discontinued due to TEAEs, n (%)	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Patients with a temporary study drug discontinuation due to TEAEs, n (%)	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Patients who discontinued due to TRAEs, n (%)	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Patients with a temporary study drug discontinuation due to TRAEs, n (%)	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A
Total number of deaths	N/A	N/A	N/A	N/A	**N/A**	N/A	N/A	[redacted]	[redacted]	N/A	N/A	N/A	N/A

Note: Data in bold refers to the anticipated licensed dose of ritlecitinib (50 mg once daily).
AE - adverse event; IGA - Investigator Global Assessment (of change); N/A - not applicable; NR - not reported; PGI-C - patient's global impression of change; SAE - serious adverse event; SAS - safety analysis set; TEAE - treatment-emergent adverse event; TRAE - treatment-related adverse event.
[a] *Combined placebo arms for 24-week double-blind period*
[b] *Interim analysis*
[c] *Study duration of ALLEGRO-LT is 24 months*
[d] *Data from the ALLEGRO-2a safety study is from the end of the double-blind period, which was either 6 months or 9 months (depending when the participant entered the extension phase)*

4.2.4.6 Subgroups

Forest plots for response based on SALT ≤ 20 at week 24, according to pre-specified subgroups, were presented in CS Appendix E (see Figure 8). The company considers that treatment effects were consistent across subgroups with the exception of AA severity, for which "*the proportion of patients with AT/AU who achieved SALT ≤10, SALT ≤20 and PGI-C was* [redacted] *than in those with non-AT/AU across all ritlecitinib treatment groups.*"

Differences in the proportion of responses based on SALT ≤ 20 at week 24 between the ritlecitinib 50mg and combined placebo arms were [redacted] for AT/AU and [redacted] for non-AT/AU (CS, Appendix E, Figure 1[32]) [redacted] (though the study was not powered to test for subgroup differences). At 48 weeks the proportion of responders was [redacted] and [redacted] for AT/AU and non-AT/AU respectively (clarification response, question A17, Figure 1[2]), but comparisons with placebo are not available at this time point.

The EAG notes that in addition to AA severity, subgroup analyses also indicate statistically significant differences in treatment effect for [redacted] (though the study was not powered to test for subgroup differences). In the company's clarification response[2] (question A12), the company maintain that treatment effects are consistent across subgroups. The company also stated that an internal longitudinal concentration response (LCR) analysis to explore the relationship between patient characteristics and SALT response concluded that AA severity status (AT/AU at baseline) was the only important covariate, providing evidence for treatment effect modification according to severity status.

Figure 8: [redacted]

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

Not applicable.

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

Not applicable.

4.5 Additional work on clinical effectiveness undertaken by the EAG

No additional work on clinical effectiveness was undertaken by the EAG.

4.6 Conclusions of the clinical effectiveness section

4.6.1 Completeness of the CS with regard to relevant clinical studies and relevant data within those studies

The clinical evidence relating to ritlecitinib for treating severe AA in patients aged 12 years and over is based on four studies: the ALLEGRO 2b/3 study, a double-blind Phase 2b/3 RCT, which examined the efficacy of five doses of ritlecitinib (200 mg once daily for four weeks followed by 50 mg once daily, 200 mg once daily for four weeks followed by 30 mg once daily, 50 mg once daily, 30 mg once daily and 10 mg once daily) for treating severe AA; the ALLEGRO 2a proof of concept study, a double-blind Phase 2a RCT, which examined the efficacy of ritlecitinib (at a dose of 200 mg once daily for four weeks followed by 50 mg once daily) and brepocitinib (not examined in this appraisal) in patients with severe AA; the ALLEGRO-LT study, a Phase 3 single-arm open-label extension study of the ALLEGRO 2b/3 and ALLEGRO 2a proof of concept studies, which also recruited *de novo* patients with moderate to severe AA; and the ALLEGRO-2a safety study, a double-blind Phase 2a RCT, which examined the safety and efficacy of ritlecitinib (200 mg once daily for four weeks followed by 50 mg once daily), with a focus on audiologic TEAEs. The EAG is confident that no additional studies (published or unpublished) of ritlecitinib for treating severe AA in patients aged 12 years and over are likely to have been missed.

4.6.2 Interpretation of treatment effects reported in the CS in relation to relevant population, interventions, comparator and outcomes

The EAG is confident that the relevant population, intervention and comparator have been included in the CS. All outcomes listed in the final NICE scope were reported on, although percentage of area affected by hair loss was not reported at each timepoint, and was instead reported at baseline and then either proportions of people meeting thresholds or in terms of change from baseline.

The trial and FDA primary outcome of the ALLEGRO 2b/3 study was the proportion of participants with a response based on an absolute SALT ≤20 at Week 24. In terms of the anticipated licensed dose

of 50 mg ritlecitinib once daily, [redacted] of patients had a SALT score of ≤20, compared with [redacted] in the placebo arms combined, which represented a statistically significantly difference of [redacted] (p<0.001). By Week 48, the proportion of patients with a SALT score of ≤20 in the ritlecitinib 50 mg arm had reached [redacted].

The EMA primary outcome of the ALLEGRO 2b/3 study was the proportion participants achieving an absolute SALT score ≤10 (response) at Week 24. In terms of the anticipated licensed dose of 50 mg ritlecitinib once daily, [redacted] of patients had a SALT score of ≤10, compared with [redacted] in the placebo arms combined, which represented a statistically significantly difference of [redacted] (p<0.001). By Week 48, the proportion of patients in the ritlecitinib 50 mg arm with a SALT score of ≤10 had reached [redacted].

The primary outcomes of the ALLEGRO-LT study were the incidence of TEAEs, SAEs and AEs leading to discontinuation, clinically significant abnormalities in vital signs, and clinically significant abnormalities in clinical laboratory values, most of which were not reported on in the CS. There were [redacted] and [redacted] TEAEs in the *de novo* and roll-over cohorts, respectively: [redacted] and [redacted] patients in the *de novo* and roll-over cohorts, respectively, had at least one TEAE; [redacted] and [redacted] patients in the *de novo* and roll-over cohorts, respectively, had at least one SAE; and [redacted] and [redacted] patients in the *de novo* and roll-over cohorts, respectively, had at least one TEAE leading to discontinuation.

The primary outcome of the ALLEGRO 2a proof of concept study was the change from baseline in SALT score at Week 24. The LSM change (improvement) from baseline in SALT score at Week 24 was significantly greater [redacted] in the ritlecitinib 200/50 mg arm [redacted] than in the placebo arm [redacted].

The primary outcome of the ALLEGRO-2a safety study was the change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80 dB at Month 9. This particular outcome was not reported on in the CS nor in the interim CSR (the EAG assumes this is because the study is ongoing and the Month 9 results are not yet available), however [redacted] audiological TEAEs were reported at the time of the data cut-off point.

4.6.3 Uncertainties surrounding the reliability of the clinical effectiveness

The key uncertainty is whether the proposed licensed dose of ritlecitinib (50 mg once daily) is effective over the long-term for patients with severe AA, including after treatment discontinuation. The evidence available at the present time only provides comparative data over a 6-month period, and little longer-term data is currently available from the ALLEGRO-LT study at the time of the data cut-off. Much of the data for the effectiveness of ritlecitinib across the four clinical studies conducted to date relates to other dose regimens, most commonly a 200 mg once daily loading dose over four weeks, followed by

50 mg once daily. This increases the uncertainty around the effectiveness of the proposed licensed dose, although data from the ALLEGRO 2b/3 study suggests that those who took the 50 mg dose achieve a similar response in terms of key clinical outcomes (notably a SALT score of ≤20, a SALT score of ≤10 and LSM change from baseline in SALT score) as those who took the 200/50 mg dose. Longer-term data from the ALLEGRO-LT study relates to patients with moderate to severe AA (in particular for *de novo* patients, who were required to have a SALT score of ≥25 at study entry, rather than ≥50 as in the other ALLEGRO studies), and thus it is less clear how effective ritlecitinib may be over the longer term in those with severe AA. In addition, no data on the effect of stopping treatment on AA severity has been presented or reported on, and thus it is unclear whether any gains made are likely to persist once patients discontinue.

5 COST EFFECTIVENESS

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

5.1.1 Objective of cost effectiveness review

The company performed systematic literature searches for: (i) published cost-effectiveness studies of patients who have alopecia areata (CS Appendix G[33]) ii) health-related quality-of-life studies (CS Appendix H[34]) and (iii) cost and resource use studies (CS Appendix I[35]). All three types of searches were undertaken in October 2021. The EAG requested that the company provide an updated search for the three searches; however, this updated search was not provided.

5.1.2 Identification of relevant studies

The cost-effectiveness study search (CS, Appendix G[33]) and cost and health care resource use systematic literature search (CS, Appendix I[35]) were combined in a two-in-one search. The following sources were searched in a single OVID host platform up to October 2021: MEDLINE, Embase, EconLit, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, Cochrane Methodology Register, Health Technology Assessment, NHS Economic Evaluation Database, ACP Journal Club, Cochrane Clinical Answers. In the clarification response to EAG questions B30 (Appendix H: Health-related quality-of-life studies[34]) and B39 (Appendix I: Cost and healthcare resource identification, measurement, and valuation[35]), the sponsor stated that given the lack of data in the original SLR in October 2021, it is unlikely that meaningful studies have been published since the last searches. Instead, targeted searches were conducted to monitor the release of new literature.

Given that the searches were combined and conducted simultaneously in a single database and the MeSH headings/Emtree thesauri were individually identified and fully comprehensive. Having reviewed the company searches, the company combined terms for the AA population with highly sensitive economic evaluation and healthcare utilization search filters and strings (including terms for disease burden, productivity, indirect comparisons and HTA).

The company also searched several HTA agencies: the Canadian Agency for Drugs and Technologies in Health, the Committee to Evaluate Drugs- Ontario Ministry of Health and Long-Term Care Canada, NICE, the Pharmaceutical Benefits Advisory Committee, and the Scottish Medicines Consortium. Hand-searching of review articles, reference tracking of full-text articles, and internet searching were undertaken by the company. The search strategies for these searches were not provided in the submission for review by the EAG.

In the HRQoL studies search, the same database sources were searched in October 2021 (CS, Appendix H,[34] Table 2) as in the cost-effectiveness study search (CS, Appendix G[33]) and the cost and health care resource use systematic literature search (CS, Appendix I[35]). The company combined comprehensive terms for AA with a high sensitivity QoL outcome search filter. There were no consequential errors in the search and the EAG considers that the search is comprehensive.

In Appendix H2 Vignette study (page 29), the company carried out two targeted literature searches to explore the patient and caregiver burden for developing the patient vignettes. Further searches were conducted to identify utility data among AA patients. In search 1 (burden of AA in adults and adolescent patients), the company searched MEDLINE, Embase, and PsycINFO. Terms for AA were combined with QoL and various named instruments. In search 2, (burden and psychosocial impact of caring), the company searched only MEDLINE and Embase. The terms for AA were combined with carer and QoL, and individual scales were named. No consequential errors were identified in these searches.

5.1.3 The inclusion and exclusion criteria used in the study selection

The inclusion and exclusion criteria used by the company are presented in CS Appendix G Table 1 for the cost-effectiveness studies, Appendix H Table 1 for HRQoL studies, and Appendix I Table 1 for cost and healthcare resource studies. The EAG considers the inclusion criteria to be appropriate to capture recent and relevant evidence.

5.1.4 Findings of the cost effectiveness review

The results of the SLR were provided in CS Appendix G1.7 Table 3 for identified economic evaluation studies. The four publications identified were related to disease burden analysis and only reported disability-adjusted life years (DALYs). Therefore, none of the studies were related to the decision problem set out in the final NICE scope.

CS Appendix H1.7 Table 3 summarises the results from 32 studies identified for HRQoL and health utility values. Only one study reported EQ-5D-5L values.[36] The robustness of the company's search for HRQoL and utility studies is critiqued in Section 5.3.4.10 Utility values for patient health states (Key issue 7) as it is part of the evidence used to support the company's choice of utility values for the model.

CS Appendix I1.7 Table 4 describes the 16 included studies for cost and utilisation data where only two were UK studies reporting costs associated with the use of contact immunotherapy.[37, 38]

5.1.5 Conclusions of the cost effectiveness review

None of the published cost-effectiveness analyses identified addressed the specific decision problem outlined in the NICE scope. The company therefore submitted a *de novo* economic analysis.

5.2 Summary of the company's submitted economic evaluation

As part of its submission to NICE, the company submitted a fully executable health economic model programmed in Microsoft Excel®. The company submitted an updated model in response to the clarification request. The EAG report describes the company's updated post-clarification model here but for transparency we have indicated where this differs from the version described in the company's original submission.

5.2.1 Population

The population included in the company's health economic analysis reflects both adult (aged ≥ 18 years) and adolescent (aged 12 to 18 years) patients with AA and a baseline SALT score of 50 or higher. The adolescent population is assumed to be [redacted]% of the total population based on data from the ALLEGRO 2b/3 study. In response to clarification question B40, although the company explained how this proportion was calculated, the EAG noted a discrepancy between the number of adults reported in the model ([redacted]) versus those reported in the CSR Table 14.1.2.1 ([redacted]). There was also a discrepancy in the number of males reported in the CS Table 14.1.2.1 ([redacted]) and the number of males reported in the model ([redacted]); indicating that the adults missing from the figures in the model were adult males. The EAG has corrected this error in their exploratory analyses as described in Section 5.4.2.1 Correction of errors related to percentage of adolescents and PSA implementation for the uncertainty in caregiver disutility and AEs associated with BSC.

In the company's model, the option was available to model both populations separately. In these subgroups analyses adults had a starting age of [redacted] years and were [redacted] female, and adolescents had a starting age of [redacted] years and [redacted] were female. However, in the company's base case analysis a combined cohort was modelled with an average age of [redacted] years, and [redacted] being female. These characteristics were based on the population of the ALLEGRO 2b/3 study.

The model included two sets of transition matrices as described in Section 5.2.5.8Adverse events and 05.2.5.2 Interim stopping rule for ritlecitinib at 24 weeks one for adults only and the other for the whole population of ALLEGRO 2b/3 study, with the latter used for both adolescents and the mixed cohort.

5.2.2 Interventions and comparators

The intervention is ritlecitinib administered orally at a dose of 50 mg once daily.

The comparator modelled by the company was BSC which the company defines as non-pharmacological management including wigs and other prosthetic supports alongside routine appointments with healthcare professionals. described as none of the pharmacological options. Further details on the resource use assumed in the BSC arm can be found in Section 5.2.5.8 Adverse

events. Outcomes from the placebo arm of the ALLEGRO 2b/3 study at 24 weeks are assumed to be applicable to patients receiving BSC in the model.

The EAG notes that the comparator in the final NICE scope is defined as "established clinical management" depending on the severity of hair loss.[11] This ranges from cosmetic options and watchful waiting to pharmaceutical interventions such as topical corticosteroids, the only licensed treatment option. Non-responders may be referred to a dermatologist where more intensive therapies can be given such as contact immunotherapy or off-label systemic treatments such as oral corticosteroids and immunosuppressive drugs. However, as previously discussed (see Section 3.3) the EAG accepts that there is significant variability in the treatments offered within the NHS currently making it difficult to define "established clinical management." In this context, BSC is considered an acceptable comparator because it is likely to be the only treatment available consistently to patients with severe AA.

5.2.3 Perspective, time horizon and discounting

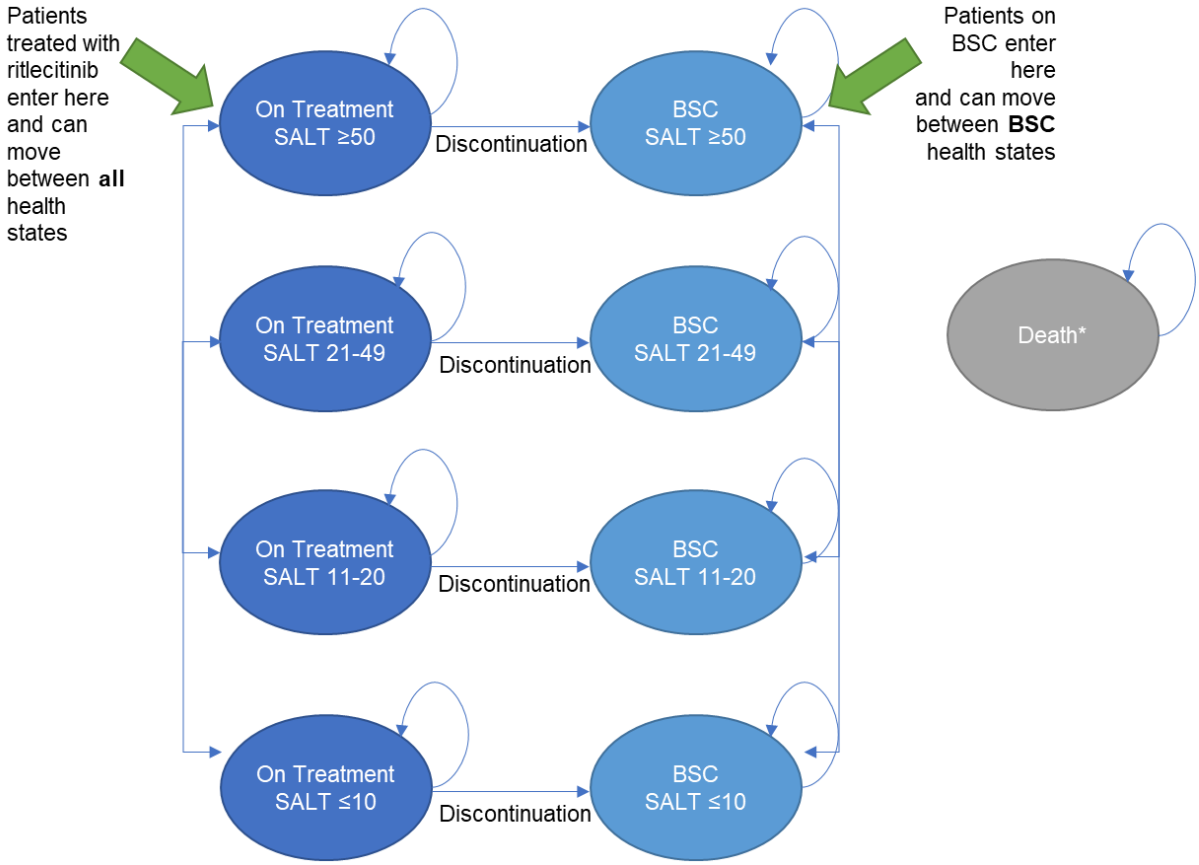
The base case model adopts an NHS and Personal Social Services (PSS) perspective although a societal perspective is explored as a scenario. The base case model uses a lifetime horizon although an alternative time horizon of 5 years was included in a scenario analysis. Both costs and QALYs were discounted at 3.5% per annum as recommended by NICE.[39]

5.2.4 Model structure

The submitted model adopts a semi-Markov state approach which determines health state occupancy based on data from the ALLEGRO 2b/3 and ALLEGRO-LT studies and consists of nine health states. There are four health states for patients on ritlecitinib treatment covering different SALT scores (SALT ≥50; SALT 21-49; SALT 11-20; SALT ≤10). There are four equivalently defined health states for patients on BSC. There is also a death state to capture all-cause mortality. The company's model structure is shown in Figure 9. The model adopts a cycle length of 12 weeks and includes half cycle correction.

The company used the SALT score to define the four AA severity-based health states based on feedback from clinicians. Clinical advice to the EAG suggested that these groupings were broadly appropriate as a change in SALT score was likely to be an outcome important to patients. However, they also they also commented that patients with a similar SALT score could have different experiences and there was not always a direct correlation between SALT score and the impact of AA of quality of life as other factors may influence the degree to which the disease impacts HRQoL. These factors include active versus stable disease, stage of life (e.g., adolescence versus middle or later life), gender, eyelash and eyebrow involvement and the degree of acceptance and adaptation

Figure 9: The company's model structure



As the population in question is patients with severe AA, all patients enter the model in either the on-treatment SALT ≥50 state or the BSC SALT ≥50 state, depending on the treatment arm. Transitions were allowed between any of the four AA severity-based health states. These are mainly informed by data from the ALLEGRO 2b/3 and ALLEGRO-LT studies with assumptions applied where data are lacking as described in Sections 5.2.5.1 Short-term health state transitions (first 48 weeks on ritlecitinib and 24 weeks for BSC) to 0

Patients on ritlecitinib who discontinue treatment transition from the on-treatment states to the BSC states. Patients could discontinue treatment due to lack of response (Section 5.2.5.2 Interim stopping rule for ritlecitinib at 24 weeks and 5.2.5.3 Final stopping rule for ritlecitinib at 48 weeks) or other reasons (Section 5.2.5.5 Discontinuation of ritlecitinib following 48 weeks). Patients who discontinue are assumed to gradually lose any treatment benefit they have accrued to this time point as described in Section 5.2.5.6 Transitions after discontinuation of ritlecitinib. The company's base case analysis implements response-based stopping rules at both 24 and 48 weeks. These assume that any patient whose SALT score worsens between baseline and 24 weeks will stop treatment, and any patient with a SALT score >20 will stop treatment at 48 weeks. Alternative stopping combinations of stopping rules are explored in scenario analyses. The implementation of stopping rules is further described in Section

5.2.5.2 Interim stopping rule for ritlecitinib at 24 weeks and 5.2.5.3 Final stopping rule for ritlecitinib at 48 weeks. The clinical appropriateness of the stopping rules is discussed in Section 5.3.4.2 Clinical appropriateness of stopping rules.

Spontaneous remission (i.e., transitioning to SALT ≤10) is assumed to occur to a fixed proportion of patients ([redacted]) starting treatment with BSC. This applies both to the BSC treatment arm at week 24, and to patients transitioning to BSC who are discontinuing from ritlecitinib in the first cycle after they start BSC (which does not occur before week 24). The transitions for patients experiencing spontaneous remission are different from those for patient reaching a SALT score ≤10 through active treatment and this is further described Section 5.2.5.7 Spontaneous remission on BSC.

Patients in any of the eight alive health states can transition to the absorbing death health state, with a transition probability that was deemed equal across all health states. General mortality rates as reported in the ONS National Life Tables for the years 2018 – 2020 were applied. No excess mortality was assigned to AA.[40]

The company's model employs the following structural assumptions in the base case analysis:

- All patients enter the model in the most severe health state, SALT ≥50.
- The placebo arm from the ALLEGRO 2b/3 study is equivalent to BSC in the model, therefore efficacy data from the placebo arm were used to model patients' progression on BSC.
- AA severity for patients on BSC is based on data from ALLEGRO 2b/3 for the first 24 weeks; thereafter it is assumed that then they gradually get worse till they reach SALT ≥50.
- AA severity for patients on ritlecitinib during the first two years is informed by data from the ALLEGRO 2b/3 and ALLEGRO-LT studies but after 2 years the company's base case analysis assumes no change in AA severity until treatment discontinuation.
- Patients are assumed to discontinue treatment with ritlecitinib if their SALT score has worsened at 24 weeks or if their SALT score is >20 at week 48 or anytime thereafter.
- Patients on ritlecitinib who discontinue and switch to BSC gradually worsen until they reach SALT ≥50.
- [redacted] of patients on BSC, whether from baseline or after discontinuing treatment on ritlecitinib, were assumed to have SALT scores ≤10; this was classed as spontaneous remission and it was assumed patients experiencing remission are in equilibrium with patients relapsing.
- Carer disutility is applied to all patients regardless of their age.
- Utility for patients with AA is assumed not to be impacted by average declines in utility observed in the general population.
- No deaths occur in the first 48 weeks of the model in line with data from the ALLEGRO 2b/3

- Adverse events included in the model are assumed to be managed through admission rather than primary care (GP appointments).
- Resource use includes wigs, psychological support consultation, visits to dermatology nurses, GPs, and dermatologists.

5.2.5 Evidence used to inform the company's model parameters

Table 17 summarises the evidence sources used to inform the parameters of the company's model. The derivation of the model parameter values using these sources is described in further detail in the following sections.

Table 17: Summary of evidence sources used to inform the model parameters

Parameter type	Parameter	Source(s)
Patient characteristics	Age at base line	ALLEGRO 2b/3[18]
	Percent male	
	Percent adolescents	
Transition probabilities	First 24 weeks for patients on BSC and ritlecitinib	Health state occupancy informed by patient-level data from ALLEGRO 2b/3[18]
	Weeks 24-48 for patients on ritlecitinib	
	Weeks 48-96 for patients on ritlecitinib	Transition probability matrices derived from ALLEGRO LT[19] (patients on ritlecitinib regardless of dose)
	All-cause mortality	ONS National Life Tables for the years 2018 – 2020[40]
AE frequency	Incidence of TEAEs - ritlecitinib	ALLEGRO 2b/3[18]
	Incidence of TEAEs - BSC	A rate ratio applied to the incidence of TEAEs on ritlecitinib estimated from ALLEGRO 2b/3[18]
Health-related quality of life	Utility for severity-based health states	The company's vignette study
	Caregiver disutility	
Resource use	Wigs, psychological support consultation, visits to dermatology nurses, GPs, and dermatologists	Expert opinion
Unit costs	Drug acquisition - ritlecitinib	The company's submission[1]
	Monitoring costs associated with ritlecitinib	National Schedule of NHS Costs - Year 2020-21[46]
	Wigs and visits to healthcare team members	Expert opinion, National Schedule of NHS Costs - Year 2020-21[46] & PSSRU 2021[45]
	Management of AEs	National Schedule of NHS Costs - Year 2020-21[46]

AE - adverse event; BSC - best supportive care; ONS - Office of National Statistics; TEAE - treatment-emergent adverse event; PSSRU - Personal Social Services Research Unit

5.2.5.1 Short-term health state transitions (first 48 weeks on ritlecitinib and 24 weeks for BSC)

Patient-level SALT score data from the ALLEGRO 2b/3 study were used to inform the health state occupancy between the four AA severity-based health states. The SALT score was assessed at weeks 4, 8, 12, 18, 24, 28, 34, 40, and 48. Data were available relating to the first 48 weeks for ritlecitinib, and for the first 24 weeks for placebo (i.e., BSC) after which patients previously receiving placebo were switched over to one of the ritlecitinib dosing regimens. Therefore, ritlecitinib data at 12, 24, 34 (as a proxy for 36 weeks), and 48 weeks and placebo data at 12 and 24 weeks were utilised to derive the distribution of patients between AA severity states for the first four and two model cycles respectively. Table 18 shows the distribution of patients over the SALT-based health states in the first 48 weeks for ritlecitinib and 24 weeks for BSC. However, in the company's base case analysis, an interim stopping rule was applied and therefore the data in Table 1 were adjusted before being applied in the model (see Section 05.2.5.2 Interim stopping rule for ritlecitinib at 24 weeks

Table 18: Distribution of patients in the first four cycles for those on ritlecitinib and first two cycles for those on BSC/placebo based on data from ALLEGRO 2b/3 (without stopping rule)

	SALT score			
	50-100	21-49	11-20	≤ 10
Patients on ritlecitinib (without interim stopping rule applied)				
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]
Week 36	[redacted]	[redacted]	[redacted]	[redacted]
Week 48	[redacted]	[redacted]	[redacted]	[redacted]
Patients on BSC				
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]

5.2.5.2 Interim stopping rule for ritlecitinib at 24 weeks

The company's base case model adopts an 'interim stopping rule' whereby ritlecitinib could be stopped for 6.45% of the cohort after two model cycles (24 weeks). This represents [redacted] patients whose SALT scores worsened by week 24 while on ritlecitinib out of the 124 patients whose data were available for analysis at 24 weeks.

Table 19 shows the distribution of patients over the SALT-based health states in the first 48 weeks for ritlecitinib when the 'interim stopping rule' was applied. The model also includes options to present results without the interim stopping rule in which the data from Table 18 was applied without adjustment. The EAG notes that the proportions achieving a SALT score ≤20 or a SALT score ≤10 are higher in Table 19 than in Table 18 and this is further critiqued in Section 5.3.4.3.

Table 19: Distribution of patients on ritlecitinib in the first four cycles assuming the 'interim stopping rule' is applied at 24 weeks

	SALT score				No treatment
	50-100	21-49	11-20	≤ 10	
Patients on ritlecitinib (with interim stopping rule applied at 24 weeks)					
Week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 36	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 48	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

5.2.5.3 Final stopping rule for ritlecitinib at 48 weeks

The company's base case analysis also applies a 'final stopping rule' whereby patients on ritlecitinib who do not achieve a target SALT score of ≤ 20 by 48 weeks are assumed to stop treatment and transition to the BSC health states. The model also has the functionality to apply this rule at 36 weeks instead of 48 weeks and to apply a definition of SALT ≤ 10 instead of ≤ 20 at either 36 or 48 weeks. These are explored in scenario analyses.

5.2.5.4 Long-term health state transitions (48+ weeks for ritlecitinib and 24+ weeks for BSC)

After week 48, any patients on ritlecitinib who fail to achieve the target SALT score of ≤ 20 (i.e., at the two ritlecitinib health states with SALT scores of 21-49 and ≥ 50) are assumed to discontinue treatment, transition to BSC health states, and gradually worsen as described in Section 5.2.5.6.

Patients who stay on treatment are handled according to the transition matrices presented in Table 20 which were derived from the ALLEGRO-LT trial. Each transition matrix contained 3-month probabilities which were assumed equivalent to 12-week probabilities.

In response to clarification question B1,[2] the company explained that the matrices accounted for time on treatment dependent on when ritlecitinib was started. This meant that for patients starting on placebo in the ALLEGRO 2b/3 trial and switching to ritlecitinib after 24 weeks, the matrices started including these patients after 6 months in the ALLEGRO-LT trial. Similarly, *de novo* patients who started ritlecitinib in the ALLEGRO-LT trial were included in the long-term matrices only after completing 48 weeks in ALLEGRO-LT. In contrast, patients receiving ritlecitinib from baseline in ALLEGRO 2b/3 contributed data from the start of ALLEGRO-LT.

After 2 years, patients on ritlecitinib were assumed to remain within the same health state unless they discontinue treatment or die. This was based on the company's assertion that "*there is no waning effect*

seen in the ALLEGRO-LT study" (CS, p127). The validity of this assumption is further discussed in Section 5.3.4.4 Critique of assumption of lack of waning (Key issue 3)5.3.4.2 Clinical appropriateness of stopping rules.

Table 20: Transition matrices derived from ALLEGRO-LT follow-up and applied from week 48 for patients on ritlecitinib

To / From	SALT score			
	50-100	**21-49**	**11-20**	**≤ 10**
Transitions from month 12 (assumed equivalent to 48 weeks) to month 15				
SALT 50-100	Patients discontinue treatment and transition to BSC			
SALT 21-49				
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT ≤ 10	[redacted]	[redacted]	[redacted]	[redacted]
Transitions from month 15 to month 18				
SALT 50-100	Patients discontinue treatment and transition to BSC			
SALT 21-49				
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT ≤ 10	[redacted]	[redacted]	[redacted]	[redacted]
Transitions from month 18 to month 21				
SALT 50-100	Patients discontinue treatment and transition to BSC			
SALT 21-49				
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT ≤ 10	[redacted]	[redacted]	[redacted]	[redacted]
Transitions from month 21 to month 24				
SALT 50-100	Patients discontinue treatment and transition to BSC			
SALT 21-49				
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT ≤ 10	[redacted]	[redacted]	[redacted]	[redacted]
Transitions from month 24 onwards				
SALT 50-100	Patients discontinue treatment and transition to BSC			
SALT 21-49				
SALT 11-20	0.0%	0.0%	100.0%	0.0%
SALT ≤ 10	0.0%	0.0%	0.0%	100.0%

For patients starting the model on BSC, all patients who achieved some response and a SALT score < 50 by week 24 were assumed to return to the severe AA state with a SALT ≥ 50 and remain in this state

there until death, with the exception of those experiencing spontaneous remission as described in Section 5.2.5.7 Spontaneous remission on BSC.

5.2.5.5 Discontinuation of ritlecitinib following 48 weeks

Kaplan-Meier (K-M) estimates of time on treatment for ritlecitinib were available for an additional 1.64 years from the ALLEGRO-LT trial following the initial 48 weeks from the ALLEGRO 2b/3 trial. This population included all ALLEGRO 2b/3 patients who were treated with any dosing regimen of ritlecitinib (except the 10 mg dose) or began on placebo and transitioned to ritlecitinib, followed by a 50 mg dose in the ALLEGRO-LT trial. It also included *de novo* patients who started ritlecitinib treatment in the ALLEGRO-LT trial with a 200 mg dose followed by a 50 mg dose. However, patients treated in China were excluded "*because the patient-level data was not available due to confidentiality requirements*" (CS, p135).

The company's model includes four different sets of K-M data for time on treatment. These four different sets are discussed further in Section 5.3.4.7 Time to discontinuation (Key Issue 6) and this description focuses on the set used in the company's base case. The company's base case analysis uses a SALT >20 definition of non-response and excludes patients from the K-M dataset if they have a SALT>20 at 48 weeks of anytime thereafter. The EAG notes that patients are excluded at all time points if their SALT score is >20 rather than being censored at the time their SALT score first becomes >20. The rationale given for taking this approach is that the discontinuation estimates were only applied to responders whose SALT scores were below 20 as all patients with higher scores are assumed to discontinue at any given cycle.

Six parametric survival curves (exponential, Weibull, Gompertz, log-logistic, lognormal, generalized gamma) were fitted to the K-M dataset, and the exponential distribution was selected in the company's base case based on convergence, and statistical fitting (AIC and BIC scores). Figure 10 presents the K-M data used plus the selected exponential curve. A critique of the methods used by the company to generate the K-M data and a discussion on the choice of survival curve is provided in Section 5.3.4.7 Time to discontinuation (Key Issue 6).

Figure 10: K-M data plus the exponential fit for time (after 48 weeks) to treatment discontinuation (inset shows detail for first 2 years)

5.2.5.6 Transitions after discontinuation of ritlecitinib

On discontinuing ritlecitinib, patients were transitioned to the BSC health state of equivalent SALT score for one cycle and then were assumed to gradually worsen by one health state each cycle. This means that patients who discontinue while having a SALT score ≤ 10 would take three cycles, equivalent of 36 weeks, to return to their initial health state of severe AA (i.e., staying in each of the two intermediate health states with SALT scores 11-20 and 20-49 for 1 cycle per each). Patients who stop treatment due to reaching a SALT score >20 will spend one cycle (12 weeks) on BSC in SALT 20-49 before returning to their original severe AA health state (SALT>50).

5.2.5.7 Spontaneous remission on BSC

In the ALLEGRO 2b/3 trial, [redacted] patients ([redacted]) out of the 131 treated with placebo reached SALT scores ≤10 by week 24. Therefore, the company's base case assumes that out of the cohort on BSC, [redacted] are at remission at any given cycle. The same proportion was also applied to patients who discontinue on ritlecitinib. The proportion of patients on BSC in spontaneous remission was assumed constant over time as it was assumed that over time the number of patients who lose remission is equivalent to that of patients who achieve it.

5.2.5.8 Adverse events

The company has applied AEs from the ALLEGRO 2b/3 trial in their base case analysis. This included TEAEs occurring in greater than 5% of patients in the ritlecitinib 50 mg arm at 48 weeks as presented in Table 35 of the CS. The probabilities from week 48 for ritlecitinib were adjusted to reflect the probability over a model cycle (12 weeks).

As placebo was only given for 24 weeks in the ALLEGRO 2b/3 trial, the company initially calculated a risk ratio (RR) between the risk of all adverse events happening to patients on placebo and that happening to patients on ritlecitinib by week 24, and this was equivalent to [redacted]. In response to clarification question B14, the company updated the RR to reflect only the TEAEs selected to appear in the model. These specific AEs occurred in [redacted] of 131 patients at 24 weeks in the placebo arm and [redacted] of 130 patients at 24 weeks in the ritlecitinib 50 mg arm (RR=[redacted]). The revised probabilities for TEAEs from the model provided post-clarification, when incorporating the revised RR are shown in Table 21.

Table 21: Probability of adverse events per cycle [adapted from CS, Table 37 to reflect post-clarification model]

Adverse event	Ritlecitinib 50 mg	BSC
Acne	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Folliculitis	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]
Urticaria	[redacted]	[redacted]

Abbreviations: BSC, best supportive care

5.2.5.9 Mortality

All-cause mortality estimates were extracted from the UK national life tables for years 2018-2020.[40] No adjustments were done to these estimates although no patient was assumed to die at the first 48 weeks in line with the ALLEGRO 2b/3 trial. The model includes functionality to allow for deaths to occur within the first 48 weeks as a scenario option.

5.2.5.10 Utility values

The company's economic model does not use any of the HRQoL collected directly from patients in the ALLEGRO phase 2b/3 trial (EQ-5D-5L, EQ-VAS, SF-36 or AAPPO) to estimate health utility values. The EQ-5D can be used to estimate utility values using UK general population preferences and the SF-36 can be used to generate utility values by mapping to the SF-6D. UK general population valuation sets are available for both of these tools, but the EQ-5D is the measure preferred by NICE. However, neither of these approaches have been used to inform the company model. The CS states that both the EQ-5D-5L and the SF-36, "*lack content validity and potential responsiveness to changes in HRQoL*," for patients with AA. The company identified high scores for the SF-36 and a high proportion of patients reporting no problems on the EQ-5D at baseline (see Section 4.2.4.4). The company interpreted this as demonstrating a ceiling effect making it difficult to demonstrate improvement in HRQoL for patients whose clinical outcomes improved during the study. They noted that the high reported HRQoL values could be due either to the exclusion of patients with 'major psychiatric conditions' from the trial or to a high level of patient adaptation, given that the mean time since diagnosis was 10.1 years in enrolled patients (CS, p162). The CS states that the EQ-5D lacks content validity as it was missing domains on social functioning, relationships, emotional aspects, physical appearance and financial elements (CS, p169). The CS also describes the SF-36 as lacking content validity because a range of factors that severely impact social functioning in patients with AA are not included in the SF-36.

The company stated that AAPPO had good content validity, citing the qualitative study by Winnette *et al.* (2021), which describes the development of the AAPPO.[41] However, AAPPO is not a preference-based measure and therefore cannot be used to estimate health utility directly. The company did not use the data from the trials to map from the AAPPO to either the EQ-5D-5L or the SF-36 due to their concerns regarding the appropriateness of these generic measures (CS, p164).

The company also conducted a systematic literature review to identify studies reporting DLQI, generic measures of HRQoL and utility values in AA. This identified three studies reporting preference-based measures of utility. One study (Burge 2021) reported EQ-5D-5L scores for three AA severity levels (mild, moderate, severe).[36] The EQ-5D-5L estimates were considered unsuitable by the company due to the insensitivity of the EQ-5D as previously discussed in relation to the ALLEGRO 2b/3 trial results. The remaining two studies reported utility scores derived from the AQoL-8D. The first study was an RCT comparing ciclosporin to placebo in patients with moderate-to-severe AA[42]. In addition to plotting utility values by trial arm over time, this paper also reported utility values for subgroups with AU/AT and patchy AA.[43] It also reported utility values by age and gender. The second study reporting AQoL-8D was a roll-over study of patients from the first trial in which respondents remained on their previously allocated treatment (ciclosporin/placebo with blinding maintained) and non-responders were offered open-label tofacitinib.[44] This paper only reported the mean change from baseline in utility for tofacitinib. The company does not use any of these estimates from the literature in the economic model. The estimates from the two trials reporting AQoL-8D were described as not useful as they are not presented in a way which maps to the model health states.

Given the company's concern regarding the direct measures of utility included in the trial outcomes, and the lack of data in the literature which they considered acceptable for use in the model, the company conducted a vignette study to inform the economic analysis. This vignette study consisted of three phases. The first phase was to develop draft vignettes informed by qualitative semi-structured interviews with patients (N=3 adults; N=3 adolescents) and carers (N=5), a detailed literature review and a retrospective analysis of the HADS and AAPPO outcomes from the ALLEGRO 2b/3 trial. In the second phase, feedback was obtained on the draft vignettes from patients (N=5 adults), caregivers (N=5) and healthcare professionals. Vignettes were developed to describe the experiences of patients for four different AA severity levels denoted by SALT scores (0 to 10, 11-20, 21-49 and 50-100) which correspond to the modelled health states. (A vignette was also developed for a SALT score of 50-100 with significant eyelash/eyebrow loss but this did not correspond to any of the modelled health states). In addition, a vignette was also developed for a carer of an adolescent with a SALT score >50. In the third phase, a time-trade-off (TTO) exercise was conducted to generate utility values for the final vignettes using a UK general population sample. In addition, the utility for this carer vignette was

compared to an age-matched general population utility value to estimate a utility decrement for carers. The resultant utility values are summaries in Table 22. A critique of the vignette study is provided in Section 5.3.4.10 Utility values for patient health states (Key issue 7).

Table 22: Summary of health state utility values and the caregiver utility value for the cost-effectiveness analysis obtained by TTO valuation of vignettes [adapted from CS, Table 46]

Health state	TTO utility value: mean (standard deviation)	Standard error	Range	95% confidence interval
N=120 members of the general population				
SALT 0-10	[redacted]	[redacted]	[redacted]	[redacted]
SALT 11-20	[redacted]	[redacted]	[redacted]	[redacted]
SALT 21-49	[redacted]	[redacted]	[redacted]	[redacted]
SALT 50-100	[redacted]	[redacted]	[redacted]	[redacted]
N=57* members of the general population				
Carer utility for patients with a SALT score >50	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutility** for patients with a SALT score >50	[redacted]	[redacted]	[redacted]	[redacted]

**The caregiver vignette was finalised and introduced into valuation interviews after fieldwork with patient health states was initiated resulting in a reduced sample size.*

***Caregiver disutility for patients with a SALT score >50 calculated by subtracting the utility for the UK population norm for people aged 35-44 (0.91)[45] from the carer utility for patients with a SALT score >50.*

Abbreviations: SALT, Severity of Alopecia Tool; TTO, time-trade-off

5.2.5.11 Resource use and costs

The costs and resource use included in the base case model comprised: drug acquisition costs; health state related costs and resource use; and those related to AE management. In addition, details of the productivity and out-of-pocket costs included in the societal perspective scenario can be found in the CS (Section B.3.5.4), but these are not described here as they are considered by the EAG to fall outside of the NICE reference case (see Section 5.3.2).

Drug acquisition costs

The cost of ritlecitinib 50 mg is £[redacted] per a pack containing 30 capsules, this includes the PAS discount. This is equivalent to £[redacted] per 12 weeks considering a compliance of [redacted]% as observed in the ALLEGRO 2b/3 trial. Being a JAK inhibitor, patients on ritlecitinib need frequent monitoring of for the [redacted]. In their response to clarification question B31, the company assumed that

[redacted] would be conducted by patients and by clinicians during the routine appointments included within resource use for each health state. Monitoring resource use and unit costs were presented in the CS Table 49 and Table 50 respectively. This equated to £7.33 per 12 weeks.

As BSC was assumed to be non-pharmacologic treatment, it was assumed to include only wigs and appointments with healthcare professionals, as discussed in Section 3.3.

Health state related costs and resource use

Resources used for different health states included wigs, a fitting and collection service for wigs, psychological support consultation, visits to a dermatology nurse in an outpatient setting, dermatology-related visits to a nurse in a primary setting, dermatology-related visits to a GP, and outpatient visits to a dermatologist. In response to clarification question B37,[2] the company described how the resource use frequency was calculated for both ritlecitinib and BSC where the estimates were presented in Table 15 and Table 16, respectively (these replace CS Tables 51 and 52). The resource use was estimated using expert elicitation via a Delphi panel of ≥5 respondents as detailed in the CS, Appendix I.[35]

The EAG notes that the use of wigs and wig services was similar between ritlecitinib and BSC; however, the resource use relating to appointments for patients on ritlecitinib was higher due to the additional monitoring for JAK inhibitors. Generally, patients on ritlecitinib needed more outpatient visits to a dermatology nurse ([redacted] extra per annum regardless of SALT score), fewer GP visits, and more outpatient visits to a dermatologist although the exact number of additional or fewer appointments varied across SALT scores. The number of psychological support appointments was higher for BSC, but only for patients with a SALT score of 11 to 20.

The unit costs for the aforementioned resources were mainly sourced from PSSRU unit cost report 2021 and the National Schedule of NHS Costs 2020-21.[46, 47] In response to clarification question B34, Table 14 presents a corrected version for the unit costs. Table 23 presents the 12-week costs per each health state as used in the company's base case.

Table 23: Health state costs used in the company's base case every 12 weeks

	Ritlecitinib	BSC
SALT ≥50	[redacted]	[redacted]
SALT 21-49	[redacted]	[redacted]
SALT 11-20	[redacted]	[redacted]
SALT ≤10	[redacted]	[redacted]

Costs related to AE management

TEAEs were assumed to be managed through admission rather than primary care, thus the unit costs were sourced from the National Schedule of NHS Costs 2020-21, as clarified in the company's response to clarification question B11.[2] In the same response, the company presented and amended version of the unit costs used in Table 9. These costs were multiplied by the frequency of AE occurrence per treatment arm as explained in Section 5.2.5.7 Spontaneous remission on BSC and resulted in estimates of £74 and £57 for managing AEs associated with ritlecitinib and BSC every 12 weeks.

5.2.6 Model validation and face validity check

The company describes their validation approach as including a discussion with UK clinical dermatologists regarding the anticipated positioning of ritlecitinib, relevant comparators, key clinical inputs and model assumptions. The CS reports that model verification was undertaken via internal and external health economists and model functionality was checked together with formulae and visual basic application (VBA) implementation. No assessment of cross validity was possible due to a lack of published cost-effectiveness analyses addressing the same decision problem to compare against.

5.2.7 Cost effectiveness results

All results presented in this section include the company's PAS for ritlecitinib and include the company's amendments to the model following the clarification process.

Central estimates of cost-effectiveness

The company's base case cost-effectiveness results are presented in Table 24, which shows the probabilistic estimates of the company's base case estimated using the average costs and QALYs across 1000 probabilistic sensitivity analysis (PSA) samples when the model was rerun by the EAG. Total costs, QALYs and ICERs were judged to converge after running the PSA over 1,000 iterations.

The probabilistic version of the model suggests that ritlecitinib is expected to generate an additional [redacted] QALYs at an additional cost of £[redacted] per patient compared to BSC resulting in an ICER of £13,394 per QALY gained. The deterministic version of the model produces a slightly lower ICER (£13,179 per QALY gained). All QALY gains relate to differences in HRQoL as the model assumes general population all-cause mortality in both arms.

Table 24: The company's base case results

Technology	Life years gained*	QALYs accrued	Total costs incurred	Incremental			ICER
				Life years	QALYs	Costs	
Probabilistic model (1000 runs by the EAG)							
BSC	NR	[redacted]	[redacted]	-	-	-	
Ritlecitinib	NR	[redacted]	[redacted]	0.00	[redacted]	[redacted]	£13,394
Deterministic model							
BSC	48.69	[redacted]	[redacted]	-	-	-	
Ritlecitinib	48.69	[redacted]	[redacted]	0.00	[redacted]	[redacted]	£13,179

** Values reported here are higher than those reported by the company as the latter were reported as discounted life years*
NR - not reported

The company presents disaggregated outcomes for the deterministic model in terms of costs accrued by different elements and QALYs accrued in different health states. These results are presented in Table 25. The differences in costs are primarily associated with the acquisition cost of ritlecitinib and costs associated with more resource use for patients on BSC whilst the additional QALY gain is mainly a consequence of additional time spent in SALT ≤10 on ritlecitinib compared to BSC and less time spent in SALT ≥50.

Table 25: Base case disaggregated outcomes for company's base case (deterministic model)

Description	Ritlecitinib	BSC	Incremental
Disaggregated costs (discounted)			
Acquisition costs	[redacted]	[redacted]	[redacted]
Monitoring costs	[redacted]	[redacted]	[redacted]
AE costs	[redacted]	[redacted]	[redacted]
Resource use costs related to health states	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]
Disaggregated QALYs (discounted)			
SALT ≥50	[redacted]	[redacted]	[redacted]
SALT 21-49	[redacted]	[redacted]	[redacted]
SALT 11-20	[redacted]	[redacted]	[redacted]
SALT ≤10	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Uncertainty around the central estimates of cost-effectiveness from the probabilistic sensitivity analysis

Figure 11 presents the cost-effectiveness plane for the company's base case PSA, and Figure 12 shows the corresponding cost-effectiveness acceptability curve (CEAC) (both based on the EAG's re-run of 1000 PSA samples). The EAG's re-run of the company's PSA suggests that the probability that

ritlecitinib generates more net monetary benefit than BSC at a WTP threshold of £20,000 per QALY gained is approximately 0.99. The company also presented the results of the PSA using cost-effectiveness planes and CEACs for ritlecitinib compared with BSC (see clarification response,[2] Figures 12 and 13). These were consistent with the EAG's re-run, with a mean ICER of £13,178 (see Table 20 of the clarification response) and 99% of samples having an ICER under £20,000 per QALY (extracted by the EAG from the model).

Figure 11: [redacted]

Figure 12: [redacted]

5.2.8 Company's deterministic sensitivity analyses

The company's deterministic sensitivity analyses are presented using a tornado plot (see Figure 14 of the company's clarification response). The analyses are performed by using the lower and upper bounds of 95% confidence intervals assuming that the standard error was set as 20% of the mean when it was not available or could not be calculated from the standard deviation and sample size.

The company's results show that the parameters which had the biggest impact on the ICER were: the utility value used for patients with SALT ≥50; caregiver disutility; health state costs associated with SALT ≥50 for patients on BSC; and the utility value used for patients with SALT ≤10. None of the parameter ranges explored increased the ICER above £20,000 per QALY gained or led to ritlecitinib dominating BSC.

5.2.9 Company's scenario analyses

The company carried out several scenario analyses that were updated post-clarification in addition to other scenario analyses requested by the EAG. These are presented in Table 22 of the company's clarification response. The scenarios with the largest impacts were the use of a 5-year time horizon (ICER of £17,323) and the inclusion of an assumption that patients who achieved a SALT score of 11-49 on BSC at 24 weeks would not revert back to a higher score (ICER to £17,223). It was also noted that changing assumptions regarding the extrapolation of long-term transition matrices after 24 months, caregiver disutility, and spontaneous remission increased the ICER above £14,000 per QALY gained. The scenario including costs not incurred by the NHS or PSS (i.e., societal perspective) reduced the ICER to less than £10,000 per QALY gained, but this was considered to fall outside of the NICE reference case by the EAG.

The following scenarios had less impact on the ICER compared with the above mentioned scenarios: allowing deaths in the first 48 weeks, selecting age subgroups, removing the interim stopping rule at 24 weeks, using SALT ≤10 as the response threshold, applying the final stopping rule at 36 weeks, using a sensitivity analysis from the vignette TTO study, using resource use for SALT=100 for all patients with SALT>50, and using other parametric distributions for time to ritlecitinib discontinuation. The company also presented some additional scenarios requested by the EAG which explored excluding patients who had dosing regimens other than the licensed dose from contributing to the long-term transition matrices, excluding wig fitting costs and applying primary care costs for TEAS. All of these resulted in ICERs under £14,000 per QALY gained.

5.3 Critique of company's submitted economic evaluation by the ERG

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

The EAG examined the company's implementation of the model within Microsoft Excel® and compared to parameters in the model with the sources in the CS and the company's response to clarification. The EAG identified some errors in the model which were later corrected by the company following clarification; these errors are not further described here. The EAG was not able within the time available to conduct any duplicate programming of the model; however, the EAG believes that the company's post-clarification version of the model to be generally well programmed. No significant programming errors were identified post-clarification, although some minor errors in the sampling of

two parameters for the PSA were identified in the numbers used to calculate the baseline characteristics (see Section 5.3.4.16 PSA errors and Section 5.2.1 respectively).

5.3.2 Adherence of the company's model to the NICE reference case

The extent to which the company's model adheres to the NICE reference case[48] is summarised in Table 26. The company's economic analysis is partially in line with the Reference Case; the main deviations that are specific to the economic analysis relate to: (i) the use of a vignette study to inform the health utility values when EQ-5D outcomes were measured directly in patients in the pivotal RCT (see Section 5.3.4.10 Utility values for patient health states (Key issue 7) for critique), and (ii) the inclusion of carer disutilities which the EAG does not consider to be sufficiently justified in this case (see Section 5.3.4.11 Utility values for carer HRQoL decrements (Key issue 8) for critique).

The EAG also notes its previous comments in Section 3 regarding the company's use of average starting characteristics to generate an ICER for the whole population specified in the decision problem (aged 12 years and over) rather than using a weighted average across age-based subgroups (12-17 years and ≥18 years). This is further discussed in Section 5.3.4.1 Age and severity subgroups (Key issues 1 and 2). It also notes its previous comments in Section 3 regarding the fact that some of the data informing the cost-effectiveness analysis is based on patients who received doses of ritlecitinib which differ from the licensed dose.

Table 26: Adherence of the company's economic analysis to the NICE reference case

Element	Reference case	EAG comments
Population	The scope developed by NICE	The company's economic analysis has a population of people aged 12 years and over with severe AA as specified in the updated NICE final scope (24th January 2023 version). The company's base case is for the whole population aged 12 years and over. The scope does not specify that subgroups based on age will be considered. However, the company's model has the functionality to present cost effectiveness estimates separately for adults (aged over 18 years) and adolescents (aged 12 to 18 years) and results are also presented separately for these groups. The EAG believes that it is appropriate for the model to capture differences in expected model outcomes based on whether the patient is an adolescent or an adult. However, the EAG prefers to take a weighted average of outcomes across the age-based subgroups, rather than using a model based on average patient characteristics, as has been done in the company's base case.
Intervention	As listed in the scope developed by NICE	The intervention is ritlecitinib administered orally at a dose of 50 mg once daily. This is consistent with the anticipated marketing authorisation. The EAG notes that some of the data informing the model were estimated in patients who had received other dosing regimens. This is further discussed in Section 5.3.4.5.
Comparator(s)	As listed in the scope developed by NICE	The comparator is BSC which the company considers to be non-pharmacological management of severe AA. The NICE final scope specified the comparator as "established clinical management without ritlecitinib". Although the scope describes a range of treatments used for severe AA, the company argues that none of these are suitable comparators (see Section 3.3). The EAG accepts that there is significant variation in current NHS clinical practice and many of the treatments currently used are either off-label/unlicensed for severe AA or are only available at a limited number of sites (e.g., contact immunotherapy). Based on this, the EAG accepts that BSC is the only comparator consistently available within current NHS clinical practice.

Element	Reference case	EAG comments
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health gains accrued by patients are valued in terms of QALYs gained. Health impacts on caregivers were included for adolescent patients up to the age of 18. The EAG is not convinced that the inclusion of caregiver QALYs is sufficiently justified in this case (see section 5.3.4.11).
Perspective on costs	NHS and PSS	The company's base case analysis adopts an NHS and PSS perspective. This is therefore consistent with the NICE reference case. The company also provides a scenario including productivity losses and out-of-pocket expenses for patients which falls outside of the NICE reference case.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The CS is consistent with the NICE reference case.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	A lifetime horizon is adopted which is considered by the EAG to be consistent with the NICE reference case.
Synthesis of evidence on health effects	Based on systematic review	No synthesis has been conducted. The main efficacy evidence comes from the ALLEGRO 2b/3 study. Whilst some data are available comparing ritlecitinib to placebo from the ALLEGRO 2a study, these data have not been synthesised with outcomes from the ALLEGRO 2b/3 study due to the inclusion of a loading dose in the ALLEGRO 2a study. The EAG considers this to be reasonable given that the dose of ritlecitinib in ALLEGRO 2a was not consistent with the anticipated marketing authorisation.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Health gains are valued in terms of QALYs. Utility values obtained from the EQ-5D-5L in the ALLEGRO 2b/3 study have not been incorporated in the company's economic analysis as the company argues that the EQ-5D lacks sensitivity in severe AA.

Element	Reference case	EAG comments
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	Instead, utility values for the model have been taken from a vignette study in which members of the general population were asked to complete a TTO exercise for a series of vignettes that reflect the model health states, and one that reflects the carer of an adolescent patient with severe AA. The EAG does not consider that the approach used to estimate health utilities is consistent with the NICE reference case. The EAG accepts that the company has provided some evidence supporting a lack of content validity for the EQ-5D in this population, and some evidence that there may be a ceiling effect within the particular population enrolled in the ALLEGRO 2b/3 trial. The EAG also notes that the duration of the trial may not be sufficient for changes in clinical measures to translate into quality of life gains given that much of the burden of severe AA is psychosocial. The EAG also accepts that given the high baseline EQ-5D values observed in the trial, using the EQ-5D data directly from the trial may underestimate the potential for utility gain in the long-term in a population with a broader range of disutility than demonstrated in the selective group enrolled in the trial. However, the EAG considers that other sources of data from the literature would support the construct validity of the EQ-5D in differentiating between patients with different severities of AA. Therefore, the EAG argues that the company could have used EQ-5D data from the literature to inform the model rather than the vignette study which is not consistent with the NICE reference case.
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to estimated QALY gains. The EAG considers this to be consistent with the NICE reference case. The company has not made a case for QALY weighting based on the severity modifier.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The company's base case cost-effectiveness analysis generally used appropriate estimates of resource use and unit costs that were consistent with the NICE reference case. The company also included a scenario analysis which included societal costs in the form of productivity losses and out-of-pocket costs. The EAG notes that the inclusion of societal costs is not consistent with the NICE reference case.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at a rate of 3.5% per annum. This is consistent with the NICE reference case.

5.3.3 The main issues identified by the critical appraisal

The issued identified in the EAG's critical appraisal are discussed in detail in sections 5.3.4 but the key issues identified in the critical appraisal which are associated with the greatest decision uncertainty are summarised in Box 1. Issue that have a smaller impact on the ICER are fully described in Section 5.3.4 but are not included in Box 1.

Box 1: Summary of the main issues identified within the company's health economic model

Issue 1 - Lack of subgroup analysis for AT/AU (Section 5.3.4.1 Age and severity subgroups (Key issues 1 and 2))
Issue 2 - ICERs should be calculated using weighted average of outcomes across age subgroups (Section 5.3.4.1 Age and severity subgroups (Key issues 1 and 2))
Issue 3 - Long-term extrapolation assumes no treatment waning (Section 5.3.4.4 Critique of assumption of lack of waning (Key issue 3))
Issue 4 - Long-term extrapolation should exclude patients receiving 30mg prior to ALLEGRO LT (Section 5.3.4.5 Generalisability of the long-term data to the 50 mg dose (Key Issue 4))
Issue 5 - There should be no spontaneous remission for patients switching to BSC from ritlecitinib (Section 5.3.4.6 Spontaneous remission (Key issue 5))
Issue 6 - Company's estimate of discontinuation leads to unrealistically high mean duration on treatment (Section 5.3.4.7 Time to discontinuation (Key Issue 6))
Issue 7 - Use of vignettes to estimate utility values for health states (Section 5.3.4.10 Utility values for patient health states (Key issue 7))
Issue 8 - Carer HRQoL decrements for carers of adolescents applied during adult years (Section 5.3.4.11 Utility values for carer HRQoL decrements (Key issue 8))
Issue 9 – Utility values are not age-adjusted (Section 5.3.4.15 Utilities not age-adjusted)

5.3.4 EAG critique of the modelling performed by the company

5.3.4.1 Age and severity subgroups (Key issues 1 and 2)

The EAG believes that a subgroup analysis should have been conducted for patients with AT/AU based on the evidence of lower clinical efficacy in this group (see Section 5.3.4.1) during the 48-week follow-up of ALLEGRO 2b/3. The EAG has estimated outcomes for this subgroup in their exploratory analyses by applying the short-term efficacy data for the AT/AU subgroup and assuming the same long-term efficacy as for the severe AA population. This subgroup analysis is further described in Section 5.4.2.14 Subgroup analysis for AT/AU.

The EAG also prefers to estimate the average costs and QALYs for adolescents (aged 12 to 17) and adults (aged ≥ 18 years) separately and then estimate a combined ICER for the whole population

considered in the decision problem. This allows the appropriate time horizon and all-cause mortality estimates to be applied to each group in addition to allowing age specific estimates of carer disutility and clinical effectiveness.

5.3.4.2 Clinical appropriateness of stopping rules

The clinical advice sought by the EAG indicated that clinicians would expect to see some treatment response by 6 months, but they would define this as a lack of progression (worsening SALT) and some evidence of regrowth, rather than a SALT score ≤ 20. Based on this, the interim stopping rule appears reasonable. Furthermore, the EAG's clinicians said that significant regrowth may take up to a year and if there was some evidence of regrowth at 6 months, then they would continue treatment up to a year and stop if there was not an adequate response at 1 year. The EAG's clinical experts had different opinions of what was considered an adequate response including a SALT score <20, a SALT score <10 or 50% regrowth which could in principle cover a range of SALT scores from 25 to 50 depending on the starting score. The clinical experts also commented that it may depend on the starting SALT score, how easily the remaining areas can be covered by the patient and whether there was a significant improvement in quality of life (e.g., a DLQI improvement of 50%). Based on this, the EAG considered that the clinical acceptability of the SALT<20 at 48 weeks stopping rule was uncertain and considers the impact of both more and less restrictive stopping rules being implemented in clinical practice in its exploratory analyses (see Section 5.4.2.2).

5.3.4.3 Clinical evidence used when implementing the interim stopping rule

The EAG is concerned that the proportion of patients achieving a response is greater when implementing the interim stopping rule compared with the proportion when not implementing the interim stopping rule. This appears to be because of the company's handling of missing data when estimating response rates for the group of patients who did not worsen in the first 24 weeks of treatment. The company states, in response to clarification question B4, that [redacted] of the 130 patients randomised to ritlecitinib 50 mg had data missing at week 24 due to COVID-19, and were assumed missing at random, leaving [redacted] available for the analysis of whether patients were steady/improved or worsened. [redacted] of these [redacted] had worsened leaving [redacted] who had steady or improved SALT scores at 24 weeks. However, the company then says, in response to question B5, that only [redacted] patients were carried forward for the analysis that informs the response rates when implementing the interim stopping rule. Based on additional information provided by the company at the factual accuracy check, the EAG understands that there were [redacted] patients with missing SALT scores at 24 weeks who were missing due to reasons other than COVID-19 and these were assumed by the company to be missing at random. (130 [redacted]). However, this is inconsistent with the company's stated approach in the effectiveness analysis (see CS, Tables 14 and 15), that patients with data missing for non-COVD-19 reasons are

assumed to be non-responders. This is also inconsistent with the assumption that these patients are steady/improved at 24 weeks for the purposes of the interim stopping rule. The proportion of responders observed in the [redacted] patients with data available at 24 weeks has been applied to the overall proportion who are steady/improved ([redacted] of the starting population). Implementing this assumption results in a higher proportion achieving SALT≤20 and SALT≤10 in the analysis with the interim stopping rule compared to the analysis without the interim stopping rule (see Table 18 and Table 19). This is despite the fact that the absolute numbers reporting this outcome is the same in both data sets (see Table 27). The EAG has explored the potential impact of assuming that the patients who were assumed steady/improved but who then did not contribute to the response rates have a SALT score ≥50 in section 5.4.2.13.

Table 27 Summary of numbers contributing to the response rates under the two alternative stopping rules

Timepoint	Data informing analysis using final stopping rule (i.e. without interim rule), N=130				Data informing analysis including interim stopping rule			
	N missing due to		N with SALT response		N missing due to		N with SALT response	
	COVID-19[a]	Non COVID-19 reasons	Informing response	Responded	COVID-19[a]	Non COVID-19 reasons	Informing response	Responded
24 weeks	[redacted]	[redacted]	[redacted]	[redacted]	N=[redacted] with [redacted][c] steady/improved; 8 worsened N=[redacted] carried forward for response			
34 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	[redacted]	[redacted]
48 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NR	[redacted]	[redacted]

[a] *Patients with SALT scores missing due to COVID-19 are assumed missing at random and excluded*

[b] *Response to clarification question B4 says that there were [redacted] missing due to COVID-19 and [redacted] contributing to response data, and response to clarification question B5 gave these figures as [redacted] but the latter was clarified as a typo at the factual accuracy check.*

[c] *patients without SALT data at week 24 missing due to reasons other than COVD-19 'may not have worsened and, therefore, are assumed steady'*

5.3.4.4 Critique of assumption of lack of waning (Key issue 3)

The CS cites the follow-up data from the *de novo* cohort of the ALLEGRO-LT study as demonstrating a plateau in SALT scores (CS, p133) and claim that this justifies the assumption that patients remain in the same health state after 96 weeks (2 years). However, the EAG notes that as CS Figure 23 only includes *de novo* patients from ALLEGRO-LT and the maximum follow-up reported in this Figure is for 24 months (data were only available for [redacted] patients at 28 months according to clarification response, Table 7). Furthermore, the median follow-up duration for SALT scores appears to be between [redacted] months as the number with data falls to less than 50% of the starting sample at [redacted] months.

In response to clarification (question A25),[2] the company also provided follow-up data from ALLEGRO-LT for patients who rolled over from the ALLEGRO phase 2b/3 and ALLEGRO phase 2a studies as a single combined 'roll-over' group. These patients had longer follow-up, with the majority ([redacted]) coming from ALLEGRO 2b/3; the exact treatment received prior to starting ALLEGRO-LT in the [redacted] rolled over from ALLEGRO 2a was variable (see clarification response Table 6 and Section 4.2.1.1). The CSR for ALLEGRO-LT (p.36) states that the "*majority of roll-over participants had up to 48 weeks of exposure to ritlecitinib prior to entering Study B7981032*" (i.e., ALLEGRO-LT). However, the EAG notes that only a minority ([redacted]) had been allocated to a 50 mg dose of ritlecitinib at the start of the ALLEGRO 2b/3 study and therefore would have had the opportunity to receive a 50 mg dose for the full 48 weeks (clarification response, Table 5). In the roll-over group as a whole ([redacted]), median follow-up appears to be between [redacted] months, with only [redacted] of the starting cohort providing data at [redacted] months and [redacted] at [redacted] months. Therefore, the efficacy data beyond a total treatment time of 2 years is limited to the [redacted] roll-over patients with data at 15 months and the [redacted] *de novo* patients with data at month 28. The EAG would argue that further follow-up data from ALLEGRO-LT would be beneficial to assess the appropriateness of assuming no waning of treatment effect in the long-term. The EAG considers that the assumption of no treatment waning is currently poorly supported and therefore explores alternative scenarios for long-term efficacy in their exploratory analyses using the average transition matrices estimated across the whole second year of treatment in ALLEGRO-LT. It should be noted that the EAG's preferred approach to estimating the average transition matrix across the second year of treatment is slightly different to the company's 'average transition matrix' scenario analysis (see Section 5.4.2.3). A scenario is also conducted by the EAG to explore the company's 'last observation carried forward' approach in which the final transition matrix from 21 to 24 months is repeatedly applied beyond 24 moths. However, this was not preferred in the EAG base case because it is informed by less data and is therefore more uncertain than using the average transition matrix.

5.3.4.5 Generalisability of the long-term data to the 50 mg dose (Key Issue 4)

The EAG is concerned that the data used to inform the transition matrices applied from 48 weeks included the *de novo* patients who received a loading dose during ALLEGRO-LT and patients who received doses other than 50 mg during ALLEGRO 2b/3. In response to clarification question B2, the company provided two scenarios which addressed these concerns. The first excluded the *de novo* patients from the long-term transition matrices. The second excluded all patients who had not transitioned from the 50 mg arm of the ALLEGRO 2b/3 study. Following clarification, the EAG is less concerned that the *de novo* patients would have had less time on treatment, as the company clarified (in response to question B1) that these patients only contribute to the long-term transition matrices after 12 months of treatment with ritlecitinib in ALLEGRO-LT. Therefore, they are equivalent, in terms of treatment received, to patients who have rolled over from the 200 mg/50 mg arm of the ALLEGRO

2b/3 study. However, the EAG is still concerned about the fact that the majority of the patients in ALLEGRO-LT did not receive the 50 mg dose from the start of their treatment (including *de novo* patients). In some cases, a higher loading dose may have contributed to a greater treatment effect, and in other cases receiving a lower 30 mg does may have contributed to a lower treatment effect (NB: the 10mg dose arm of ALLEGRO 2b/3 were excluded). Based on SALT≤20 responses reported for the different dosing regimens in ALLEGRO 2b/3 (CS, Figure 13), the EAG considers that the impact of the loading dose in the first month is likely to be minimal in patients who have had at total of 48 months on treatment. Therefore, data from those who had a loading dose before moving on to a 50mg dose can be combined with those having a 50mg dose from base line of the purposes of estimating the long-term efficacy. However, the EAG is not convinced that patients rolling over from the 200mg/30mg or 30mg arms of the ALLEGRO 2b/3 study will have outcomes similar to those rolling over form either 200mg/50mg or 50mg. This is because they will be experiencing a dose increase from 30mg to 50mg at the start of the ALLEGRO-LT study which may result in transition matrices that capture improvement in response to a dose increase instead of the expected outcomes for patients on a consistent 50mg dose. Therefore, the EAG has conducted exploratory analysis in which patients receiving the 30mg dose before joining the ALLEGRO-LT study are excluded from the data used to estimate the long-term treatment efficacy (see Section 5.4.2.4).

5.3.4.6 Spontaneous remission (Key issue 5)

The EAG would argue that any cases of spontaneous remission for patients in the ritlecitinib arm of the ALLEGRO 2b/3 trial would have been incorporated within the group of patients deemed to have responded to ritlecitinib and therefore would have already been captured within the modelled outcomes for 48 weeks and the status of these patients as continued responders in the ALLEGRO-LT study. Therefore, allowing for a further incidence of spontaneous remission when treatment non-responders switch to BSC is not appropriate and the EAG preferred to exclude spontaneous remission on discontinuation from ritlecitinib in their base case (see Section 5.4.2.5).

The EAG's clinical experts advised that spontaneous remission was extremely unlikely in their experience in patients with severe AA and became less likely beyond the first 6 months to a year. Therefore, the company's scenario that explored higher rates of spontaneous remission was not considered realistic.

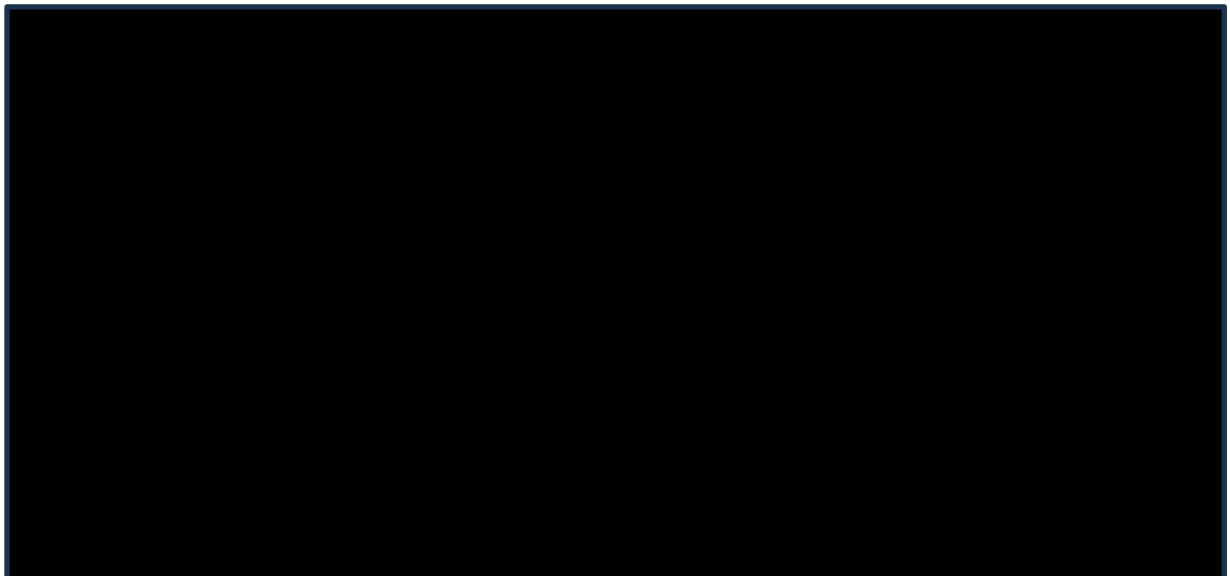
5.3.4.7 Time to discontinuation (Key Issue 6)

The company's model includes four different sets of K-M data for time on treatment. In all four sets, patients were excluded from the K-M dataset if they were non-responders at 48 weeks. The definition of non-response was SALT>20 for two of the K-M datasets and SALT>10 for the other two datasets. For each definition of response, there is one analysis excluding patients who fail to meet the response

criteria after 48 week and another where these patients are not excluded. The company's base case used the K-M dataset where patients were excluded if they had a SALT score >20 at 48 weeks or at any time thereafter. The four different K-M curves can be seen in Figure 13.

The EAG believes that it is incorrect for the company to exclude patients who have a SALT score >20 after 48 weeks on treatment from the analysis of time to discontinuation. The EAG accepts that in the model, patients reaching a SALT score of >20 would discontinue treatment, and for that reason would not remain at risk of discontinuation for another reason not related to SALT score. However, the EAG would prefer to censor those patients at the time they reached a SALT score >20 rather than excluding them from the K-M plots altogether. In addition, the EAG is concerned that the company's analysis assumes a constant risk of discontinuation over time (i.e., an exponential survival curve for time to discontinuation), whereas the K-M data included in the model suggest a higher rate of discontinuation around 1.5 years for 3 of the four possible K-M datasets. Although this may be due to smaller numbers informing the sample at that time, it is impossible to know if this is the cause as no information has been given by the company on the numbers at risk over time. In addition, the K-M curve showing time to discontinuation for patients with a SALT score <20 at 48 weeks, with no exclusion based on later SALT scores, shows a marked increase in the rate of discontinuation around 1 year which is not accurately captured by the company's preferred survival curve. This K-M curve also appears to have an additional 1 year of data compared with the curves in which the dataset was restricted to those with a SALT score <10 at 48 weeks or where patients with a SALT score >20 or >10 at later time points were also excluded. The EAG is concerned that this could either be because this K-M did not have data from the first 48 weeks of follow-up excluded, in which case the higher discontinuation rate from later on should have been used, or it is because the duration of follow-up was significantly reduced by the exclusion of patients with a SALT score >20 after 48 weeks. If the latter is the explanation then this is inconsistent with the company's claim that SALT scores are stable in the long-term (i.e., after 96 weeks, which would be around 1 year on Figure 13).

Figure 13: K-M datasets provided in the model for time to discontinuation after 48 weeks and proportion of responders at 48 weeks remaining on treatment in the company's base case



The net effect of the company's approach is that there is a very slow decline in the proportion of responders at 48 weeks who remain on treatment after 48 weeks (see % on treatment in base case model in Figure 13). This slow decline is based on few discontinuations being observed in the group who never have a SALT score exceeding 20 during the ALLEGRO-LT follow-up period. However, there is clearly a much higher rate of discontinuation occurring in the ALLEGRO-LT study when discontinuations for any reason after 48 weeks are included in the K-M data. The EAG would argue that the very low rate of discontinuation observed in the model leads to a predicted mean time on treatment of 5.9 years, which is unrealistically high. Long-term follow-up studies of other JAK inhibitors for rheumatoid arthritis with around 9.5 years of patient follow-up are reporting median durations on treatment of 3.1 years and 4.6 years for tofacitinib and baricitinib, respectively (mean durations of 3.2 and 3.9 years, respectively).[49, 50] The EAG believes that a more realistic estimate of time to discontinuation would be provided by censoring rather than excluding those having a SALT score >20 from the K-M data before fitting the survival curves. In the absence of such an analysis, the EAG would prefer to apply a higher discontinuation rate. It has explored an arbitrary doubling of the hazard for discontinuation in Section 5.4.2.6. This results in a mean time on treatment of 3.6 years for the ritlecitinib arm which is more consistent with the mean time on JAK inhibitors seen in rheumatoid arthritis where long-term follow-up studies are available. Although the duration of time on treatment may be subject to different factors in severe AA and rheumatoid arthritis, in the absence of long-term studies to assess the likely duration of treatment, the EAG believes that their approach provides a more reasonable estimate.

The EAG also notes that the company has provided minimal details regarding the methods used to fit the parametric survival curves to the time on treatment data. However, given that the alternative curves presented have similar AIC/BIC values, and similar predictions during the period of observed data, the EAG believes that the choice of curve should be based on the one that provides the most realistic long-term extrapolation of time on treatment. The exponential curve, which was chosen by the company, provides a lowest mean duration of time on treatment, but the EAG believes that this results in a predicted mean duration of treatment within the model that is unrealistically high. Therefore, the EAG does not prefer any of the alternative parametric survival functions.

The EAG also notes that it was unable to match the K-M estimates in Figure 29 of the CS to any of the four K-M curves provided in the model, despite being able to match the survival curves shown by overlaying the graphs generated in the model with CS, Figure 29 (see Appendix 1, Figure 14 which shows the base case K-M curve from the model overlaid on CS, Figure 29). This appears to be because the time axis is one datapoint out of step with the survival data. In light of the EAG's concerns regarding

the company's decision to exclude rather than censor patients with a SALT score >20 and the apparent discrepancy between the K-M data in the model and the data plotted in CS, Figure 29, which is being used to assess model fit, the EAG is more concerned about the validity of the K-M dataset to which the curves have been fitted than the selection of the appropriate curve based on fit to that dataset.

5.3.4.8 Adverse event rates

The company's cost-effectiveness analysis only included adverse events occurring in greater than 5% of the study population receiving the licensed dose of ritlecitinib (50 mg). However, the EAG would argue that rare but serious adverse events, should be included regardless of their frequency if they have the potential to have high costs. The EAG is particularly concerned about the risk of serious infection which was a safety event of interest in the ALLEGRO 2b/3 study. [redacted] Given that serious infections are a recognised risk for JAK inhibitors, [redacted] the EAG would have preferred to see some exploration of the impact these events could have on the cost-effectiveness of ritlecitinib. The EAG notes that CS, Appendix F summarises safety evidence across a broader group of studies (ALLEGRO 2a proof of concept, ALLEGRO 2b/3, ALLEGRO 2a safety study, ALLEGRO-LT and study in vitiligo patients – B7981019) described as the AEP population, with the purposes of examining "*events of low frequency, longer latency and for sub-group analysis*" (Appendix F, Table 4). The EAG believes that this AEP population should be used to explore serious adverse events in the model as it draws on a wider group exposed to ritlecitinib across multiple clinical trials.

5.3.4.9 Mortality in the first 48 weeks

Whilst it is unlikely to have a substantial impact on the ICER, the EAG would prefer to include all-cause mortality during the first 48 weeks of the model as per the scenario analysis provided by the company. It is standard practice to include all-cause mortality in cost-effectiveness analyses even for conditions without an increased risk of mortality. The lack of deaths observed within the ALLEGRO 2b/3 trial simply reflects the inability of a trial to detect a low all-cause mortality rate in a sample of this size. This does not mean that we would not expect any deaths from any cause in the population likely to be treated within the NHS. The EAG has therefore included mortality during the first 48 weeks in their base case analysis (see Section 5.4.2.7).

5.3.4.10 Utility values for patient health states (Key issue 7)

The CS states that generic measures of HRQoL are not specific or sensitive enough and therefore the company proposes that the utility values from vignettes valued using TTO are a more appropriate source of utility estimates for the cost-effectiveness analysis. The CS states that this is in accordance with the hierarchy of preferred HRQoL evidence described in the NICE methods guide.[48] The EAG agrees that vignettes are one of the options included in the hierarchy presented in the NICE methods guide, but notes that vignettes valued using preference elicitation methods, such as TTO, are only an option where evidence shows that the EQ-5D is not appropriate. The guide also states that to make the case that the EQ-5D is inappropriate, evidence should be provided on the content validity, construct validity and responsiveness of the EQ-5D in the population of interest. In situations where the EQ-5D has been demonstrated as inappropriate, the hierarchy in the NICE methods guide would recommend using a different generic preference-based measure or a condition-specific preference-based measure over the use of vignettes.

The company's claim that the EQ-5D lacks content validity is supported by a systematic literature review, interviews with patient advocacy group representatives and interviews with clinicians. Informed by these, the CS presents a diagram of aspects of HRQoL not covered by generic quality of life measures (CS, Figure 44). The EAG believes that some of the elements would map well to items on the EQ-5D. For example, worry, sadness, anxiety and hopefulness should be covered by the anxiety and depression domain of the EQ-5D, and academic performance /productivity should be covered by the usual activities domain which includes work and study as examples. In addition, there are other aspects included in Figure 44, such as the financial impact of hair replacements, that the EAG would consider as falling outside of HRQoL. Whilst the purpose of the EQ-5D is to cover the broad domains of health that are common across many conditions, the EAG recognises that patients may not see how these domains are relevant in the context of their specific condition and this is reflected in the company's qualitative research with patient advocacy group representatives (n=9).

The CS describes the EQ-5D as being insensitive in AA due to high baseline scores. However, two of the reasons given for this insensitivity relate to the population enrolled in the ALLEGRO 2b/3 study in which people with psychiatric comorbidities were excluded and there was a mean 10.1 years since diagnosis of AA in the trial cohort. Therefore, the observed insensitivity of the EQ-5D may not apply in other studies where the EQ-5D has been used in populations with a shorter duration of time since diagnosis or a higher prevalence of psychiatric comorbidities. For example, the cross-sectional study reported by Edson-Heredia *et al.* 2022 had a mean time since diagnosis of 2.4 years (SD 4.05), an average HADS-A score of 5.36 and an average HADS-D score of 6.21 despite including a mixture of mild, moderate and severe AA.[51] This study reported a significant difference in EQ-5D by severity

ranging from 0.90 in mild AA to 0.79 in severe AA. The HADS-D and HADS-A scores in the severe AA subgroup were 6.63 and 7.62 respectively. These findings suggest that the high baseline EQ-5D values in ALLEGRO 2b/3 study, which the company states are similar to that of population norms, may be due to selection bias in the ALLEGRO 2b/3 specific study.

The EAG considers that the study by Edson-Heredia (2022)[51] indicates that the EQ-5D demonstrates construct validity within a cohort of AA patients with mixed severity because statistically significant differences were observed in EQ-5D between mild, moderate and severe groups. This demonstrates 'known group' content validity for the EQ-5D-5L. The anxiety/depression domain of the EQ-5D-5L is reported to differ statistically significantly across the severity levels and the differences are in the direction expected based on significant differences in HADS-A and HADS-D scores across the severity levels. However, the EAG notes that Edson-Heredia *et al.* did not explicitly define mild, moderate and severe AA using SALT scores and therefore these groups may not correlate exactly with changes in SALT score and may be more related to a physicians global impression of severity which could encompass a broader evaluation of the patient than hair loss alone.[51] The mean hair loss in the severity groups was 8.2%, 26.2% and 72% in patients with mild, moderate, and severe AA, suggesting that physician assessed severity was strongly correlated with hair loss. The authors also report that the degree of hair loss from the scalp was described as the main indicator of severity for physicians, followed by the patient's level of distress. The EAG believes that this study would support the use of the EQ-5D in detecting quality of life differences related to different degrees of hair loss in AA. Although Edson-Heredia *et al.* reports outcomes from Japanese patients enrolled in the Adelphi AA Disease Specific Programme (DSP)™, similar outcomes were reported in an abstract by Burge *et al.*, which reported outcomes from the same database, but for US patients.[36] Utilities based on EQ-5D-5L in this US cohort were 0.95, 0.93 and 0.87 for mild, moderate and severe AA respectively and the difference was statistically significant (p=0.007). An abstract by Bewley *et al.* which also reports outcomes from the Adelphi database, but for a European cohort, provides EQ-5D values of 0.90±0.10, 0.85±0.14 and 0.78±0.17 (UK value set), for mild (N=91), moderate (N=267) and severe (N=174) patients respectively.[52] Whilst fewer details are available on the methods used in the US and European studies, due to only abstracts being available, the results suggest consistent findings across Japanese, US and European cohorts.

A second psychometric test for assessing the appropriateness of a tool such as the EQ-5D in a specific population is measures of responsiveness. That is the ability to detect differences in patients over time when other measures of disease are known to have changed. To assess this, one needs to either follow patients over time as their disease changes severity through a longitudinal observational study or to detect changes in EQ-5D within a trial setting. Therefore, it may be difficult to assess responsiveness in a disease with few effective treatments unless a prospective study exists following patients over time.

The company's SLR for HRQoL data did not identify any longitudinal data sets reporting EQ-5D over time, or any clinical trials reporting EQ-5D changes in response to treatment for treatments other than ritlecitinib. Therefore, assessment of responsiveness is limited to the outcomes reported in the ritlecitinib trials included in the CS.

Whilst the company has presented changes from baseline in EQ-5D for responders and non-responders these were initially provided only by study arm and were therefore limited by sample size (CS, Table 39). In response to clarification (question B16, Table 11), the company also provided these values when aggregated over all ritlecitinib treatment arms and all treatment arms (i.e., ritlecitinib arms and placebo arms). In this analysis, although there was a trend for higher EQ-5D utility scores for responders, the difference between the means was small ([redacted] at both 24 and 48 weeks when using either all ritlecitinib arms or all study arms) and the 95% CI for responders and non-responders were overlapping at each time point. A better assessment of responsiveness would have been to estimate the mean change from baseline in EQ-5D for responders and non-responders across all study arms, but the company declined to provide such an analysis in response to clarification (question B16). The company did provide an estimate of Cohen's d effect size comparing EQ-5D utilities (US valuation) for SALT<10 with SALT=100. This used data from multiple timepoints in the ALLEGRO-2b/3 trial to provide a large number of observations. They reported utilities of [redacted] and [redacted], respectively, and a medium effect size (d=0.24) (CS, p143)[53].

Responsiveness can be difficult to assess in a clinical trial setting as an inability to demonstrate a change in EQ-5D associated with a treatment when a change is measured in other trial outcomes could demonstrate either a lack of responsiveness in the EQ-5D or that those clinical outcomes do not have a significant correlation with HRQoL. There may also be a time lag whereby HRQoL improves following response to treatment but not within the timeframe of the study. In this case, it may be that the SALT score improves due to regrowth over a significant proportion of the scalp, but hair regrown within the timeframe of the trial is still short at the end of the trial period and the HRQoL improvements of not needing to use wigs may not be realised immediately. This may be especially true if the main impact of severe AA is psychosocial. This could account for the lack of statistically significant differences between trial arms in the ALLEGRO 2b/3 study. The EAG notes that longer term EQ-5D data were collected in the ALLEGRO-LT study, but these have not been provided by the company making it difficult to assess if the short duration of the ALLEGRO 2b/3 study is a significant factor in the lack of difference between trial arms in EQ-5D.

Assessment of responsiveness can also be affected by ceiling effects in that improvements in EQ-5D may be difficult to assess in a population with very high starting values and this finding does not mean that the HRQoL tool would not be responsive in another population. The EAG also notes that the EQ-

5D has been assessed as having good validity and responsiveness in people with other skin diseases, based on a 2015 review of 16 papers covering a range of skin conditions, with most of the studies (12 of 16) being in patients with psoriasis.[54]

Overall, the EAG considers that whilst the company has made the case that patients with severe AA may not be able to easily relate the impact of AA to the questions contained in the EQ-5D tool when specifically asked about its relevance, the availability of data showing variation in EQ-5D utility scores by AA severity suggests that it could be used to estimate utility scores for responders and non-responders in the cost-effectiveness analysis. Therefore, the company has not sufficiently made the case that EQ-5D is not appropriate in severe AA.

The company's SLR of HRQoL stated that it included 24 studies reporting DLQI (Appendix H, post-clarification version). However, the CS states that none of the studies reported utility values obtained by mapping from the DLQI. The EAG notes that such mapping algorithms do exist and that the company does not appear to have explored using any of the published studies reporting DLQI to estimate utility values for the economic model.[55] The CS raises concerns regarding the relevance of the DLQI in AA, as the DLQI is primarily focused on the impact of skin conditions on quality of life, and therefore not specifically on the impact of hair loss. However, the CS reported that 6/8 clinicians considered the DLQI to be a good method for assessing HRQoL in patients with AA, although 4/8 said it did not fully capture the impact of AA because some domains were irrelevant. When the EAG asked their clinical experts about the relevance of DLQI, they had a similar response, with two out of three experts saying it was useful and used regularly in clinical practice and the other expert saying that they did not use it because it was too skin focused. In response to clarification (question B27), the company stated that existing DLQI to EQ-5D mapping algorithms were not considered appropriate or fit-for-purpose and that any mapping study would be limited by the insensitivity of the EQ-5D in AA. Overall, the EAG considers that DLQI is potentially useful and that the company could have explored the potential for using DLQI to estimate utility values for the health states. However, the lack of DLQI measurement in the trials themselves means that DLQI values for the various modelled health states would need to be obtained from the literature. Furthermore, taking this approach would be reliant on using mapping algorithms developed in other skin conditions, such as psoriasis.[55]

The EAG is not confident that the company's systematic review of HRQoL studies has identified all of the relevant evidence. Firstly the review has not been updated since October 2021 and the EAG have identified at least one publication since then (Edson-Heredia *et al.* in March 2022) which has reported relevant EQ-5D which was missed by the company.[51] Secondly, the EAG identified four studies[56-59] from a published systematic review[60], which the company had incorrectly excluded and which were included by the company in their updated Appendix H following clarification. Whilst none of these four

studies provided alternative utility estimates, their incorrect exclusion from the original review, combined with the omission of the Edson-Heredia paper which did provide an alternative source of utility values, lowers the EAG's confidence about whether the review process was robust and identified all relevant information.

The CS describes the development of the vignettes in Appendix H. Overall, the EAG considers that many of the best practice recommendations described by Matza *et al.* have been followed.[61] The company has used multiple sources of evidence to inform the vignettes including published literature, qualitative evidence, such as interviews with patients and clinicians, and quantitative data on AAPPO and HADS from the clinical trials and a separate preference study. Vignettes were refined and validated using input from clinical experts and patients and the process of vignette development has been fully reported. The methods used to value the vignettes appears to be appropriate, in that a TTO valuation approach has been used in a general population sample in accordance with NICE's preferred approach when EQ-5D is not considered appropriate.

However, the EAG has some concerns regarding the methods used in the vignette study. Firstly, the vignettes do not report the absence of symptoms in domains unaffected by AA such as self-care and mobility. In contrast these are included in the carer vignettes. The absence of positive messages stating a lack of problems in the domains unaffected by AA may cause a focusing effect leading to an overestimation of the importance of condition-specific symptoms in the members of the general population being asked to complete the TTO exercise (DSU hierarchy p30).[62]

Secondly, the patients involved in the qualitative interviews were all required to have had specific treatments previously or be interested in receiving systemic treatment. This requirement may have led to a selective group who are more motivated to find an effective treatment because they experience a greater HRQoL impact from their severe AA than others who would not seek treatment. The EAG would therefore argue that the vignettes may not be applicable to patients who have similar SALT scores but are currently managing their alopecia without receiving or seeking to receive systemic treatments. Furthermore, the debrief interviews were conducted in an entirely female sample, which could have biased the refinement of the health states to reflect experiences more representative of women.[63]

Finally, the EAG believes that the vignettes lack face validity when compared with the quantitative data the company has used to inform them. For example, the vignette for SALT>50 (A3) states, [redacted] whereas the data from the ALLEGRO 2b/3 trial (Table 8, Appendix H) suggest that the most common response to, "over the past week, how often did you feel embarrassed about your hair loss?" in patients with a

SALT score of 50-100 was never ([redacted]), followed by rarely ([redacted]). In fact, only [redacted] responded sometimes, often, and always respectively. Similarly, [redacted] of patients with a SALT score of 50 to 100 responded 'not at all' when asked, "over the past week, how much did you limit your interactions with others because of your hair loss?" This contrasts with a description of, "[redacted] for the relevant vignette (A3). The draft vignette for SALT>50 used the wording, "[redacted] which seems less definite than the wording in the final vignette. The impact of patient and clinician feedback on the vignettes, as summarised in Appendix H Table 12, shows an increase in frequency for many of the domains with the CS stating that "there was consensus that the severity of the impacts was understated across all health states." Although it is best practice to have draft vignettes validated by patients and clinicians, this does in this case appear to have resulted in quite a discrepancy between the quantitative data used to inform the vignettes and the final wording. On this basis, the EAG would question whether the utility values estimated accurately reflect the average experience of patients with the relevant SALT score.

The company states that the utility estimates for AA from the vignette study are within a similar range to those reported for other dermatological conditions such as psoriasis, hidradenitis suppurativa, atopic dermatitis and venous leg ulcers. However, the company only provides utility values for one of these conditions, hidradenitis suppurativa for which it cites a utility range of 0.35 to 0.80 for most to least severe hidradenitis suppurativa. This is based on a cross-sectional study of patients with hidradenitis suppurativa where severity was categorised by Hurley stage (1 to 3). The EAG is not convinced that it is meaningly to compare utility scores in AA with other dermatological conditions, as each condition will define severity differently and each will have different domains of HRQoL affected. For example, severe hidradenitis suppurativa is associated with significant pain which is not present in severe AA.

Overall, the EAG would have preferred to have seen a reference case analysis using the EQ-5D data from the ALLEGRO 2b/3 trial, preferably also supplemented with longer-term EQ-5D from ALLEGRO-LT, to estimate utility values for the economic analysis. The EAG is not convinced that the company has demonstrated that the EQ-5D is not appropriate in AA as it considers that the ceiling effects observed in the trial may be the result of selection bias in the study population. Furthermore, the availability of EQ-5D data from the literature showing that utility scores derived from the EQ-5D vary by AA severity would support the appropriateness of the EQ-5D for the purposes of informing the modelling. Therefore, in the absence of a reference case analysis, the EAG prefers to use the data from Bewley *et al.* and this has been incorporated in the EAG's base case analysis (see Section 5.4.2.8). In addition, the EAG considers that any analysis based on the vignette study should be treated with caution

given their concerns regarding the lack of correspondence between the quantitative data used to inform the vignettes and the wording in the final vignettes.

5.3.4.11 Utility values for carer HRQoL decrements (Key issue 8)

The EAG is concerned that the vignette for caregivers specifically asks the TTO participant to imagine that they are a caregiver to a family member between the ages of 12 and 17 years with severe AA (SALT 50 to 100), but in the company's base case analysis the utility decrement for caregivers is applied regardless of whether the person with AA is an adult or an adolescent. Furthermore, a carer vignette has only been developed for patients with severe AA and it is therefore not possible to estimate how much quality of life for caregivers might improve in response to achieving an improvement in SALT score. The carer utility decrement has simply been estimated as the difference between the vignette for a carer for an adolescent with severe AA and the general population utility for a person of a similar age. It is possible that much of the impact of AA on caregivers might remain even in the event of a treatment response, particularly if there were parental concerns regarding the safety of the treatment and the likelihood that symptoms could return if treatment were stopped.

The application of caregiver disutilities explicitly within cost-effectiveness analyses in TAs has typically been quite rare, occurring in only 12 out of 414 TAs conducted up to January 2019.[64] It has been more common in HSTs, occurring in 4 out of 8 HST conducted up to the same time point but this likely reflects the severity of conditions reflected within the HST programme. The majority of the TAs including carer HRQoL decrements were in multiple sclerosis. Other diseases represented were Alzheimer's, myelofibrosis, juvenile idiopathic arthritis and moderate to severe atopic dermatitis. However, the vast majority of TAs conducted in the period covered by this review did not include carer HRQoL decrements. It is unclear to the EAG why carer HRQoL decrements should be more relevant in AA than in the many other diseases considered by NICE. The NICE methods guide states that evidence should be provided to show that the condition is associated with a substantial effect on carer's HRQoL, but the EAG notes that the company has not provided any utility values directly measured in caregivers and has not attempted to estimate how caregiver HRQoL might vary in response to treatment. Therefore, their estimate of caregiver disutility is reliant on the assumption that the impact of severe AA on caregivers is accurately described in the vignette and that all the impact will be resolved if patients move out of the severe AA state, whether or not they achieve SALT<20.

In response to clarification (Question B21), the company provided a scenario analysis in which the carer utility decrement is limited to the proportion aged under 18 at baseline and to the model cycles in which they remain under 18. Given the way in which the vignette study was framed, the EAG considers this to be much more appropriate than the base case analysis in which it was applied to all patients regardless of age. The EAG prefers to apply this approach in their base case but also explores excluding the carer

HRQoL decrement altogether in a scenario analysis on the basis that the vignette study did not explore how the carer HRQoL decrement would change in response to treatment (see Section 5.4.2.9).

5.3.4.12 Resource use for health states

The EAG is concerned that a higher level of psychological support is assumed for patients having BSC who have a SALT score of 11 to 20. The EAG believes that this lacks clinical face validity as the need for psychological support is likely to be related to the severity of AA rather than the treatment being received. The EAG prefers to assume the same psychological support for patients with the same SALT score receiving either BSC or ritlecitinib in its base case analysis (see Section 5.4.2.10)

The EAG's clinical experts advised that access to wig provision for patients with severe AA was inconsistent across the NHS and access to free wigs was limited to specific groups such as young people in full-time education and those receiving specific benefits.[65] Therefore, the EAG believes that the cost to the NHS of wig provision assumed in the model may not reflect actual current practice and explored scenarios where wig provision was half that assumed in the company's base case for adults (see Section 5.4.2.11).

Furthermore, the EAG's clinical experts noted that access to psychological provision for dermatology patients with severe AA was limited currently in the NHS. Therefore, the EAG has explored scenarios in which psychological support is half that assumed in the company's base case.

5.3.4.13 Resource use for adverse events

The EAG notes that the company has assumed that adverse events will result in admission. However, none of the adverse events included in the model were SAE and instead were TEAE. The EAG believes it is more likely that these TEAEs will be managed in primary care as per the company's scenario analysis (see Section 5.4.2.12).

5.3.4.14 Utilities for adverse events

The EAG was unable to corroborate the disutility of 0.07 applied for acne, folliculitis and urticaria which the company claims was based on the disutility for severe redness reported by Stein *et al.* (2018).[66] The EAG did identify a disutility of 0.06 for severe redness/skin peeling as a severe adverse event (i.e. Grade 3 or 4) in response to chemotherapy. The EAG does not consider that this is likely to be applicable to the adverse events it is applied against in the company's model which were all described in the CS as either mild or moderate severity. The EAG notes that Stein *et al.* report a disutility of 0.218 for serious infection, which is an adverse event that the company has not included in their analysis because it did not occur in the 50mg arm of ALLEGRO 2b/3, even though serious infections did occur in ritlecitinib treated patients in the broader AEP population. The EAG was unable to corroborate the

disutility of 0.01 applied for nasopharyngitis for which the company claims to have used a disutility of 'sense organ disorder' reported by Sullivan *et al.*[67] However, the EAG believes that it is unlikely that mild to moderate nasopharyngitis would result in significant utility loss. The disutility values for diarrhoea and headache seem appropriate and could be corroborated from Sullivan *et al.* Overall, the EAG questions the need to include disutilities in the model for TEAEs which are mild to moderate in severity and believes the impact may have been overestimated which would be favourable to BSC, but the likely impact will be small. For this reasons, the EAG has not explored this issue further in their exploratory analyses.

5.3.4.15 Utilities not age-adjusted (Key issue 9)

The EAG notes that utility values are constant across time for patients in a particular health-state and are therefore not adjusted to reflect declining utilities in the general population over time. The EAG considers that the AA specific utility weights should have been applied as multipliers to the expected utility values in the general population. The EAG has been unable to implement a scenario analysis which incorporates the age-adjustment accurately and fully to their satisfaction within the company's existing model structure. However, the EAG believes that the lack of age-adjustment has led to an overestimation of the incremental QALYs gained and therefore an underestimation of the ICER.

5.3.4.16 PSA errors

The EAG notes that caregiver disutility ('Utility weights' sheet cell D79) may assume a positive value with running PSA as there was no capping to it. It was also noted that following clarification process question B14, the company included the uncertainty in the calculated rate ratio for AEs happening with BSC in the PSA; however, the probability of each AE occurrence per cycle ('Adverse Events' sheet cells E54:E61) takes the deterministic value (G50) instead of the probabilistic value (L50). These errors have been corrected by the EAG in their exploratory analyses (see Section 5.4.2.1 Correction of errors related to percentage of adolescents and PSA implementation for the uncertainty in caregiver disutility and AEs associated with BSC)

5.4 Exploratory analyses undertaken by the ERG

5.4.1 Overview of the EAG's exploratory analyses

The exploratory analyses performed by the EAG are described in Section 5.4.2. These included correcting implementation errors in the model (Section 5.4.2.1 Correction of errors related to percentage of adolescents and PSA implementation for the uncertainty in caregiver disutility and AEs associated with BSC) and exploring alternative plausible data and assumption (Sections 5.4.2.2 to 5.4.2.13). Results for individual changes are reported in Section The results of the EAG's exploratory analysis are shown in Table 29. The EAG's corrected company base case implements the corrections described in Section 5.4.2.1. This is followed by implementing individual changes using the EAG's corrected

company base case as the starting point, which are described as EAG exploratory analyses 1 to 9. All of these individual changes are implemented in the model for the whole population (≥12 years of age) which uses average baseline characteristics, and these are reported using the deterministic model. These are then combined in an EAG base case for the whole population, for which a deterministic result is presented. EAG deterministic base case results are then presented for the adolescent and adult subgroups (aged 12 to 17 years and aged ≥18 years respectively) and the EAG base case analysis using the weighted average across these subgroups is presented using a both a deterministic and probabilistic approach. A deterministic subgroup analysis is presented for the AT/AU subgroup using the EAG's base case preferences and a weighted average approach for the age subgroups. In this scenario the efficacy data up to 48 weeks are the same for both age groups but other aspects of the modelling differ by age as described under Section 5.4.2.14. Finally, deterministic scenario analyses 1 to 8 are presented using the EAG base case as the starting point and using a weighted average approach to estimate the ICER across the whole population covered in the decision problem. It is worth noting that probabilistic ICERs are significantly higher than their deterministic counterparts because the continuity correction applied to the long-term transition matrices, to handle transitions with no observations (e.g. month 12 to month 15 transition from SALT ≥50 to SALT ≤10), is only applied in the probabilistic version of the model and this reduces the relative treatment effect of ritlecitinib.

5.4.3.1 Impact of individual changes. The EAG has also combined their preferred data and assumptions to give an EAG preferred base case reported in Section 5.4.3.2. The EAG has also presented an exploratory analysis the AT/AU subgroup using their preferred base case assumptions combined with AT/AU specific short-term efficacy data (methods in Section 5.4.2.14 Subgroup analysis for AT/AU and results in Section 5.4.3.3)

5.4.2 EAG's exploratory analyses - methods

5.4.2.1 Correction of errors related to percentage of adolescents and PSA implementation for the uncertainty in caregiver disutility and AEs associated with BSC

The EAG corrected the identified errors described in Section 5.2.1 and Section 5.3.4.16 PSA errors as follows:

(a) Patients missing from numbers used to calculate base line characteristics

The EAG identified that there were [redacted] adult males missing from the numbers used in the company's model to estimate the proportion of the patients who are adolescent and the proportion of patients who are male (see Section 5.2.1). The EAG corrected this resulting in the percentage who are adolescent reducing from [redacted]% to [redacted]% and the percentage of adults who are female reducing from [redacted]% to [redacted]%.

(b) Capping the maximum of the caregiver disutility by zero

The EAG introduced a cap to the sampled values for the caregiver disutility values in the PSA so that they never exceed zero (see Section 5.3.4.16).

(c) Inclusion of the uncertainty in rate ratio for AEs associated with BSC in the PSA

The EAG applied the probabilistic value for rate ratio ('Adverse Events' sheet cell L50) in calculating the per cycle probability for AEs associated with BSC when running the PSA (see Section 5.3.4.16).

5.4.2.2 Stopping rules for ritlecitinib

The EAG base case adopts the company's stopping rule for patients at week 48 if SALT ≤20 was not achieved. However, two scenarios are explored; the first is a more restrictive scenario where patients not achieving SALT ≤10 at week 36 and thereafter would discontinue, and a less restrictive scenario where an interim stopping rule is not applied at 24 weeks and a threshold of SALT ≤20 is applied at 48 weeks and thereafter. The EAG did not include the option to relax the stopping rule after 48 weeks in the less restrictive stopping rule scenario, as when this option is selected the transition matrices which excluded patients whose SALT score increased over 20 are still applied and the EAG had concerns regarding the face validity of discontinuation rates when this option is selected.

5.4.2.3 Transition probabilities for patients on ritlecitinib beyond 2 years

In its base case, the EAG pooled the transitions observed across the four 3-month transition matrices used to inform year 2 transitions on ritlecitinib and used these pooled counts to estimate a single 3-month transition matrix, which was applied from 2 years onwards. This is slightly different from the 'average transition matrix' scenario provided by the company as the company estimated four transition matrices and then estimated the numerical average across these. In a scenario analysis, the EAG used the final transition matrix from 21 to 24 months to inform the transition probabilities after 2 years on treatment.

5.4.2.4 Transition probabilities for patients on ritlecitinib between year 1 and year 2

The EAG used the transitions estimated only from patients on the 50 mg dose (regardless of whether there was a loading dose) to inform the matrices applied from 48 weeks in its base case.

5.4.2.5 Spontaneous remission in the ritlecitinib arm

We have assumed no spontaneous reemission in the ritlecitinib arm when patients transition from ritlecitinib to BSC.

5.4.2.6 Time to discontinuation

In the EAG's preferred base case we have continued to use an exponential time to discontinuation curve but have assumed double the hazard of discontinuation applied in the company's base case. This results in a mean time on treatment of 3.6 years across the whole ritlecitinib arm (responders and non-responders).

5.4.2.7 All-cause mortality in first 48 weeks

We have applied all-cause mortality in the first 48 weeks as per the company's scenario analysis.

5.4.2.8 Utility values for patients by severity

We have used the utility values reported by Bewley *et al.*,[52] applying the value for mild AA (mean 0.90, SD 0.10) to both the SALT≤10 and SALT11-20 health states. We have applied the value for moderate (mean 0.85, SD 0.14) to SALT 21-49 and the value for severe (mean 0.78, SD 0.17) to SALT ≥50.

5.4.2.9 Utility value for carers

In the EAG's base case we have applied the carer disutility for carers of adolescents only to those patients in their adolescent years (ages 12-17) as per the company scenario provided at clarification. We have also conducted an exploratory analysis in which carer disutility is excluded altogether.

5.4.2.10 Same resource use for psychological support across treatments

In the EAG's base case we have set the number of appointments for psychological for patients receiving ritlecitinib to the same values applied to patients receiving BSC.

5.4.2.11Reduced access to psychological support and wig services

We have also conducted exploratory analyses in which we halve the number of appointments for psychological support and the number of wigs received in both arms.

5.4.2.12 Resource use for adverse events

In the EAG base case we have assumed that the TEAEs included in the model result in a GP appointment rather than a hospital admission as per the company's scenario analysis provided at clarification.

5.4.2.13 Assumed outcomes in the interim stopping rule analysis for patients with missing SALT scores

Given the EAG's uncertainty regarding the handling of patients with missing data in the implementation of the interim stopping rule, the EAG has explored an alternative assumption in which the [redacted] patients who were assumed to be steady/improved at week 24, but who did not contribute to response rates thereafter were assumed to have a SALT score >50. Therefore, an additional [redacted] patients with SALT >50 were included in the health state occupation calculations for patients aged 12 and over.

When using the data specific to adults for the subgroup analysis, there were [redacted] adults who were steady/improved at 24 weeks, with only [redacted] contributing to the response data at 34 weeks and 48 weeks, but we don't know how many of the [redacted] adults were missing due to COVID-19 reasons. In the absence of this information, we've assumed only [redacted] of the [redacted] patients were missing due to non-COVID reasons and we've assumed an additional [redacted] non-responders in this exploratory analysis.

As the EAG are not entirely clear why these patients have been excluded, we are unsure how many of these patients were adults and it has been necessary to assume a SALT score for those with missing data, these changes are only considered in exploratory analyses.

5.4.2.14 Subgroup analysis for AT/AU

For the AT/AU subgroup, the EAG would have preferred to extract the transition matrices for the first two years from the patient level data provided in the CSR if time permitted. Instead, the EAG extracted the figures reported in the CSR for ALLEGRO 2b/3 regarding response rate at week 24 for SALT ≤10 (Figure 14.2.2.2.5.1) and the response rates up to week 48 for SALT ≤20 (Figure 14.2.1.3.4.4). This meant that breakdown of patients between health states SALT 11-20 and SALT ≤10 for timepoints 12, 36, and 48 weeks was not available. Neither was the distribution of non-responders between SALT scores of 21-49 and ≥50. Therefore, the EAG had to make assumptions as shown in

Table 28. In addition, the EAG did not have AT/AU subgroup data that was specific to adults and therefore the data from all ages was assumed applicable to both adults and adolescents.

The long-term transition matrices were assumed the same as with the whole cohort as detailed in Section 5.2.5.2 Interim stopping rule for ritlecitinib at 24 weeks; however, the EAG notes that this may overestimate the efficacy of ritlecitinib in this subgroup in patients with AT/AU are more likely to experience treatment failure. In the subgroup analysis for AT/AU, the EAG has applied the same resource use and costs as in the EAG's preferred base case.

Table 28: Distribution of patients in the AT/AU subgroup for ritlecitinib up to week 48 and BSC up to week 24 based on ALLEGRO 2b/3 data and additional assumptions

	SALT score				No treatment
	50-100	21-49	11-20	≤ 10	
Patients on ritlecitinib (with interim stopping rule applied at 24 weeks)					
Week 12	[redacted]	[redacted]	[redacted]	[redacted]	0.0%
Week 24	[redacted]	[redacted]	[redacted]	[redacted]	0.0%
Week 36	[redacted]	[redacted]	[redacted]	[redacted]	0.0%[3]
Week 48	[redacted]	[redacted]	[redacted]	[redacted]	0.0%[3]
Patients on BSC					
Week 12	[redacted]	[redacted]	[redacted]	[redacted]	
Week 24	[redacted]	[redacted]	[redacted]	[redacted]	

[1] 3.4% achieved SALT ≤20, these were assumed equally distributed between 11-20 and ≤10. The breakdown for the remaining patients (96.6%) was based on the proportions for the whole population.
[2] The breakdown for the remaining patients (92.8%) was based on the proportions for the whole population
[3] No worsening was assumed as these patients are starting from SALT 100
[4] 20.0% achieved SALT ≤20, these were assumed equally distributed between 11-20 and ≤10. The breakdown for the remaining patients (80.0%) was based on the proportions for the whole population
[5] 30.9% achieved SALT ≤20, these were assumed equally distributed between 11-20 and ≤10. The breakdown for the remaining patients (69%) was based on the proportions for the whole population
[6] No patients achieved SALT <20 or SALT <10 in placebo arm of ALLEGRO 2b/c for AT/AU subgroup. The breakdown between the SALT≥50 and SALT 21 to 49 health states was based on the proportions for the whole population

5.4.3 Results of the EAG's exploratory analyses

The results of the EAG's exploratory analysis are shown in Table 29. The EAG's corrected company base case implements the corrections described in Section 5.4.2.1. This is followed by implementing individual changes using the EAG's corrected company base case as the starting point, which are described as EAG exploratory analyses 1 to 9. All of these individual changes are implemented in the model for the whole population (≥12 years of age) which uses average baseline characteristics, and these are reported using the deterministic model. These are then combined in an EAG base case for the whole population, for which a deterministic result is presented. EAG deterministic base case results are then presented for the adolescent and adult subgroups (aged 12 to 17 years and aged ≥18 years respectively) and the EAG base case analysis using the weighted average across these subgroups is presented using a both a deterministic and probabilistic approach. A deterministic subgroup analysis is presented for the AT/AU subgroup using the EAG's base case preferences and a weighted average approach for the age subgroups. In this scenario the efficacy data up to 48 weeks are the same for both age groups but other aspects of the modelling differ by age as described under Section 5.4.2.14. Finally, deterministic scenario analyses 1 to 8 are presented using the EAG base case as the starting point and using a weighted average approach to estimate the ICER across the whole population covered in the decision problem. It is worth noting that probabilistic ICERs are significantly higher than their deterministic counterparts because the continuity correction applied to the long-term transition matrices,

to handle transitions with no observations (e.g. month 12 to month 15 transition from SALT ≥50 to SALT ≤10), is only applied in the probabilistic version of the model and this reduces the relative treatment effect of ritlecitinib.

5.4.3.1 Impact of individual changes

After correcting errors in the company's deterministic model, the ICER for ritlecitinib versus BSC is estimated to be £13,179 per QALY gained. The largest change in the ICER was seen when the EAG used the utility values reported by Bewley *et al.* This increases the ICER to £33,945 per QALY gained. Using pooled counts of patients throughout the second year on treatment to inform transitions beyond year 2 increases the ICER to over £15,600 per QALY gained.

Three changes increased the ICER above £14,000 per QALY; assuming no spontaneous remission in the ritlecitinib arm, doubling the discontinuation hazard, and applying a carer disutility only during adolescent years. Using only patients who were on the 50 mg dose to inform the long-term matrices and assuming the same psychological support for ritlecitinib as for BSC had minimal impact on the ICER. Assuming TEAEs are managed in primary care and allowing mortality in the first 48 weeks of the model marginally decrease the ICER to approximately £13,000 per QALY and £13,100 per QALY respectively.

When including all the changes preferred by the EAG, the deterministic ICERs using average starting characteristics increased to £60,735 per QALY. The base case deterministic ICERs for adults and adolescents were £69,246 per QALY and £55,349 per QALY respectively, resulting in a weighted average of £66,674 per QALY (probabilistic ICER = £82,152 per QALY). The deterministic ICER using the weighted average approach differs significantly from the deterministic ICER when using the average starting characteristics for the whole population (£66,674 per QALY versus £60,735 per QALY). The weighted average approach more accurately captures the expected outcome in both age subgroups which is the reason the EAG prefers to use this approach in its base case. However, the EAG notes that the impact of using this approach is smaller when applying the company's preferred assumptions where it increases the ICER from £13,179 per QALY to £13,235 per QALY. The AT/AU subgroup had a lower ICER compared to the EAG base case for the whole population, because although the response rates were lower in the AT/AU subgroup, this resulted in fewer patients staying on treatment due to the stopping rule, and the reduction in incremental cost exceeded the reduction in incremental QALYs. The EAG notes that this exploratory subgroup analysis for patients with AT/AU conducted by the EAG required many assumptions and should therefore be interpreted with caution. The EAG would prefer to see the model was properly populated with data specific to the AT/AU subgroup by the company.

The following four scenario analyses showed a further increase of approximately £11,000 to £23,000 per QALY in the deterministic ICER: applying a more restrictive stopping rules with a response threshold of SALT ≤10 from 36 weeks; using only patients who had started and continued trials on the licensed 50 mg dose to inform the long-term TMs; applying month 21-24 transition matrix to inform transitions beyond year 2 for patients on ritlecitinib. All other scenarios increased the ICER less significantly.

Table 29: Results of the EAG's exploratory analyses [a]

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
Company base case (Deterministic)					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,179
EAG's corrected company base case: correcting implementation errors in the company's economic model					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,179
EAG exploratory analysis 1: Using pooled counts from the second year to estimate the 3-month transition matrix applied from 2 years onwards					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£15,676
EAG exploratory analysis 2: Using only patients who were on the 50 mg dose to inform the long-term matrices					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,294
EAG exploratory analysis 3: Assuming no spontaneous remission in the ritlecitinib arm					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£14,578
EAG exploratory analysis 4: Assuming double the hazard of discontinuation applied in the company's base case					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£14,217
EAG exploratory analysis 5: Allowing mortality in the first 48 weeks of the model					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,139
EAG exploratory analysis 6: Using the utility values reported by Bewley *et al.*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£33,945
EAG exploratory analysis 7: Carer disutility applies only during adolescent years					

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£14,192
EAG exploratory analysis 8: Assuming the same psychological support for ritlecitinib as for BSC					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,170
EAG exploratory analysis 9: Assuming TEAEs are managed in primary care					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£12,976
EAG base case applying analyses 1-9 (Deterministic)					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£60,735
EAG base case subgroup 1 (adults only)					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£69,246
EAG base case subgroup 2 (adolescents only)					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£55,349
EAG base case average weighted average across age subgroups (Deterministic)					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£66,674
EAG base case applying analyses 1-9 (Probabilistic)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£82,152
EAG subgroup (AT/AU only)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£60,293
EAG scenario 1 (Restrictive stopping rules applied; interim stopping rule in addition to response threshold of SALT ≤10 applying from 36 weeks onwards)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£89,888
EAG scenario 2 (Relaxed stopping rules applied; no interim stopping rule)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£69,220
EAG scenario 3 (the last long-term TM from month 21-24 is applied to transitions beyond 2 years for patients on ritlecitinib)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£77,806

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
EAG scenario 4 (Using only patients who had the licensed 50 mg dose throughout the two trials with no placebo or loading dose to inform the long-term TMs)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£84,238
EAG scenario 5 (Assuming no disutility for caregivers)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£68,960
EAG scenario 6 (Assuming half the resource use for wigs)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£67,267
EAG scenario 7 (Assuming half the resource use for psychological consultation)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£68,451
EAG scenario 8 (Assuming missing patients were to transition to SALT ≥50)*					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£68,425

[a] *Deterministic and for the whole population covered by the decision problem (aged ≥ 12 years) unless otherwise stated*
[*]Using the average weighing between adults and adolescents

5.4.3.2 The EAG's estimate of the ICER (whole population covered in scope)

The exploratory analyses conducted by the EAG, which are provided in Table 29, indicate that there are plausible changes to parameter values which would considerably increase the company's estimate of the ICER but where the most appropriate value remains uncertain. Such parameters include: the utility values for AA severity states; the rate of discontinuation for reasons other than loss of response; and the long-term transition matrices. Uncertainty regarding these parameters could be reduced in future by additional follow-up from the ALLEGRO-LT study and further analysis of the long-term EQ-5D that this study will provide.

The EAG also had concerns regarding the handling of missing data in the analysis used to estimate response rates when implementing the interim stopping rule. This could be addressed by the company providing further information on the number of patients with data available at the different time points and further analyses exploring alternative assumptions for those with missing data.

The EAG also had concerns regarding the method for estimating the carer disutility which could be addressed by either providing estimates of utility measured directly in carers of patients with different

SALT scores, or by directly measuring the change in carer utility when patients respond to treatment. However, both of these options would require further primary data collection.

Some areas of decision uncertainty in the model are unlikely to be addressed by further evidence collection. For example, the ICER is quite sensitive to the choice of stopping rule with both the more and less restrictive scenarios explored by the EAG resulting in an increase in the ICER; interpretation of this is complicated by the more favourable response rates when implementing the interim stopping rule. If the stopping rule proposed by the company is not broadly acceptable to clinicians, then the cost-effectiveness of ritlecitinib when implemented in clinical practice may differ from that estimated in the EAG's preferred base case analysis. In addition, if the age distribution of patients receiving ritlecitinib in clinical practice differs significantly from that observed in the ALLEGRO 2b/3 trial then that may mean that the average cost and QALY outcomes estimated in the model for the eligible population as a whole may not be realised in clinical practice.

5.4.3.3 The EAG's estimate of the ICER for the AT/AU subgroup

The EAG's estimate of the ICER in the AT/AU subgroup is lower than in the broader population of severe AA included in the scope (£60,293 per QALY versus £66,674 per QALY). The EAG anticipated a higher ICER given the lower response rates observed in the AT/AU subgroup in the ALLEGRO 2b/3 study. Although the EAG notes that when the efficacy data for the AT/AU subgroup were applied to the company's base case scenario the ICER increased from £13,179 per QALY to £15,207 per QALY. The EAG notes that their analysis for the AT/AU subgroup made several assumptions, and a more accurate estimate of the ICER could be generated by fully populating the model with data specific to this subgroup.

6 OTHER FACTORS

The company has not submitted any evidence to support the implementation of a severity modifier in this appraisal. Although the company has not done the necessary calculations to estimate the absolute and proportional QALY losses required to evaluate whether a severity modifier should be applied in this case, the EAG does not believe that it is likely that the requirements would be met in this appraisal. A managed access scheme has not been proposed.

7 OVERALL CONCLUSIONS

7.1 Clinical effectiveness conclusions

The clinical evidence relating to ritlecitinib for treating severe AA in patients aged 12 years and over is based on the ALLEGRO 2b/3 RCT, the ALLEGRO-LT single-arm study, the ALLEGRO 2a proof of concept RCT, and the ALLEGRO-2a safety RCT, with the majority of the evidence coming from the ALLEGRO 2b/3 study. The EAG's clinical advisors confirmed that the eligibility criteria of these studies are representative of the severe AA patients seen in clinical practice, with the exception that patients might be seen at an earlier stage in the clinic than among the samples included in the ALLEGRO studies. In the ALLEGRO 2b/3 study, ritlecitinib demonstrated significantly greater efficacy than placebo in terms of clinically important outcomes. A significantly greater proportion of patients in the ritlecitinib 50 mg arm than the combined placebo arms had a SALT score of ≤20 (and , respectively) and a SALT score of ≤10 (and respectively), at Week 24. By Week 48, the proportions of patients in the ritlecitinib 50 mg arm with a SALT score of ≤20 and ≤10 were and (although comparison with a placebo control was not possible at that timepoint). In the ALLEGRO-LT study, and of *de novo* and rollover participants, respectively, had a SALT score of ≤20 at Week 24, and by Week 48 the proportions were and , respectively. At Week 24, and of *de novo* and rollover participants, respectively, had a SALT score of ≤10 at Week 24, and by Week 48 the proportions were and , respectively. In terms of LSM change from baseline in SALT score, in the ALLEGRO 2b/3 study patients in the ritlecitinib 50 mg arm attained significantly greater reductions ( ) than those in the combined placebo arms ( ) at Week 24, with even greater reductions at Week 48 among those in the ritlecitinib 50 mg arm ( ). In the ALLEGRO 2a proof of concept study, significantly greater reductions were also seen at Week 24 in the ritlecitinib 200/50 mg arm ( ) than in the placebo arm ( ). Similarly, in the ALLEGRO-2a safety study, there were greater reductions in SALT score (LSM change from baseline) in the ritlecitinib 200/50 mg arm ( ) than in the placebo arm ( ) at Month 6.

The key issue relating to evidence for the clinical effectiveness of ritlecitinib is uncertainty over whether the proposed licensed dose of ritlecitinib (50 mg once daily) is effective over the long-term for patients with severe AA, including after treatment discontinuation. Further evidence from the ALLEGRO-LT study may elucidate this, although the EAG notes that, as the inclusion criteria allowed for those with milder AA (SALT ≥25), it would be difficult to assess the long-term effects of ritlecitinib on those with severe AA.

7.2 Cost-effectiveness conclusions

The economic analysis submitted by the company was broadly in line with the decision problem specified in the scope, with the exception of two issues related to the handling of subgroups. Firstly, the company did not provide a subgroup analysis for the AT/AU population. The EAG considers this omission to be important as response rates in the ALLEGRO 2b/3 study were lower for this subgroup indicating the potential for higher ICERs in this group. Whilst the EAG has conducted an exploratory analysis for this subgroup, which estimated an ICER that is lower than the ICER in the overall population (£60,293 per QALY versus £66,674 per QALY), this analysis relies on several assumptions. The EAG would prefer to see a company analysis populated with appropriate data for the AT/AU subgroup. Secondly, although the company provides subgroup analyses for adolescents (aged 12 to 17 years) and adults (aged 18 years and over), their base case ICER is based on average characteristics across the whole population (aged 12 years and over). The EAG prefers to use a weighted average approach to estimate expected outcomes in the whole population from the estimated outcomes in the age-specific subgroups. The importance of this issue is greater in the EAG's preferred base case scenario in which carer disutility is restricted to adolescent patients.

The economic analysis submitted by the company largely complied with the NICE reference case with the most important exception being that the EQ-5D data from the pivotal ALLEGRO 2b/3 trial were not used to inform the model. The EAG was not persuaded that the company had sufficiently demonstrated that the EQ-5D was not appropriate in AA in general, although it accepts that there appears to be some issues with high baseline utility values in the ALLEGRO 2b/3 trial, which may be specific to the trial population. The EAG also noted the availability of EQ-5D data from the literature which suggest that the EQ-5D does have construct validity in AA. The EAG also considered that the utility values generated from the vignette study should be treated with caution given their concerns regarding the lack of correspondence between the quantitative data used to inform the vignettes and the wording in the final vignettes. The EAG also had reservations regarding whether the inclusion of disutility for carers was sufficiently justified in this case because the company has not measured utility values directly in carers and has not demonstrated that any disutility for carers is reduced when the patient they care for responds to treatment. The company's base case scenario took an NHS and PSS perspective, but the company also conducted a scenario analysis which included societal costs in the form of productivity losses and out-of-pocket costs. The EAG considers this scenario to fall outside of the NICE reference case.

The EAG's clinical advisors considered that a model structure based on SALT score was broadly appropriate. However, they also noted that whilst a change in SALT score is likely to be important to patients, there may not be a direct correlation between SALT score and the impact of AA on HRQoL and therefore a model based on SALT score may miss important heterogeneity in the experience of

patients. The EAG considered that the inclusion of a stopping rule based on lack of response was likely to be clinically appropriate, but there may be variation in how clinicians assess response in clinical practice and therefore the impact of alternative stopping rules on cost-effectiveness should be considered.

The company's base case deterministic analysis estimates that ritlecitinib is expected to generate an additional [redacted] QALYs, at an additional cost of [redacted] per patient, with a corresponding ICER of £13,179 per QALY versus BSC across the whole population covered by the scope. The company's base case probabilistic analysis provided similar results to their deterministic analysis with an ICER of £13,394 per QALY.

The EAG had concerns regarding several key model assumptions including: the application of disutility values for carers of adolescent patients to carers of patients of all ages; the inclusion of spontaneous remission as an outcome for patients switching from ritlecitinib to BSC; the assumption of no treatment waning for patients remaining on ritlecitinib in the long-term and the lack of age-adjustment for utilities. The EAG had concerns regarding several parameters informing the economic analysis including: the utility values from the vignette study; the data used to estimate long-term efficacy; the method used to estimate treatment discontinuation rates; and the handling of missing values for the analysis informing the interim stopping rule. The EAG also had concerns regarding several other issues which were found to have less impact on the estimates of cost-effectiveness. These related to adverse events, resource use for different severities of AA, and all cause-mortality in the first year.

The EAG conducted analyses to explore the impact of the uncertainties described above on the estimates of cost-effectiveness. The EAG's exploratory analyses indicate that the ICER could be potentially much higher than that estimated in the company's base case with the most important areas of decision uncertainty being as follows: the choice of utilities values; the data used to estimate long-term efficacy; the inclusion of spontaneous remission for ritlecitinib; the rate of treatment discontinuation; and the inclusion of carer disutility. Under the EAG's preferred assumptions, ritlecitinib is expected to generate an additional [redacted] QALYs, at a cost of £[redacted] per patient, with a corresponding ICER of £66,674 per QALY versus BSC across the whole population covered by the scope. The corresponding probabilistic estimates were an additional [redacted] QALYs, at a cost of £[redacted] per patient, with a corresponding ICER of £82,152 per QALY. However, the EAG considers there to be considerable uncertainty in this estimate as the deterministic ICER increased to £89,888 per QALY when implementing stricter stopping criteria.

8 REFERENCES

1. Pfizer. Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over. Company evidence submission summary to NICE. *Document B* 2023.
2. Pfizer. Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over [ID4007] - Response to EAG Clarification Questions. In: Excellence NIfHaC, ed. Single Technology Appraisal; 2023.
3. National Institute for H, Care E. Clinical Knowledge Summaries: alopecia areata. 2018.
4. Madani S, Shapiro J. Alopecia areata update. *Journal of the American Academy of Dermatology* 2000;42:549-66.
5. 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. *Br J Dermatol* 2022;186:257-65.
6. Macbeth AE, Holmes S, Harries M, Chiu WS, Tziotzios C, de Lusignan S*, et al.* The associated burden of mental health conditions in alopecia areata: a population-based study in UK primary care. *Br J Dermatol* 2022;187:73-81.
7. Trueb RM, Dias MFRG. Alopecia Areata: a Comprehensive Review of Pathogenesis and Management. *Clin Rev Allergy Immunol* 2018;54:68-87.
8. Harries MJ, Sun J, Paus R, King LE, Jr. Management of alopecia areata. *BMJ* 2010;341:c3671.
9. Messenger AG, McKillop J, Farrant P, McDonagh AJ, Sladden M. British Association of Dermatologists' guidelines for the management of alopecia areata 2012. *Br J Dermatol* 2012;166:916-26.
10. Xu L, Liu KX, Senna MM. A practical approach to the diagnosis and management of hair loss in children and adolescents. *Front Med (Lausanne)* 2017;4:112.
11. National Institute for Health and Care Excellence. Final scope: Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over. London, UK.: National Institute for Health and Care Excellence; 2022.
12. British Association of Dermatologists. Ritlecitinib for moderate-to-severe alopecia areata (aged 12 and over) [ID4007]. Professional organisation submission. Submission to NICE. In; 2023.
13. Pfizer. Identification, selection and synthesis of clinical evidence. *Appendix D of Company evidence submission summary to NICE* 2023.
14. Banno M, Tsujimoto Y, Kataoka Y. Using the Cochrane Central Register of Controlled Trials to identify clinical trial registration is insufficient: a cross-sectional study. *BMC Medical Research Methodology* 2020;20.
15. Centre for Reviews and Dissemination. Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination; 2008.
16. 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. *British Medical Journal* 2011;343:d5928.
17. Critical Appraisal Skills Programme. CASP Cohort Study Checklist.; 2018.
18. Pfizer. A Phase 2b/3 Randomized, Double-Blind, Placebo-Controlled, Dose-Ranging Study to Investigate the Efficacy and Safety of PF-06651600 in Adult and Adolescent Alopecia Areata (AA) Subjects With 50% or Greater Scalp Hair Loss; 2022.
19. Pfizer. A Phase 3 Open-Label, Multi-Center, Long-Term Study Investigating the Safety and Efficacy of PF-06651600 in Adult and Adolescent Participants With Alopecia Areata; 2022.
20. King B, Guttman-Yassky E, Peeva E, Banerjee A, Sinclair R, Pavel AB*, 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;85:379-87.
21. Pfizer. A Phase 2a, Randomized, Double-Blind, Placebo-Controlled Study Investigating the Safety of Ritlecitinib (PF-06651600) in Adult Participants With Alopecia Areata; 2022.
22. Pfizer. A Phase 2b/3 randomised, double blind, placebo-controlled,dose ranging study to investigate the efficacy and safety of PF-06651600 in adult and adolescent alopecia areata (AA)

subjects with 50% or greater scalp hair loss. Clinical trial registration no. NCT03732807: clinicaltrials.gov; 2022.
23. Pfizer. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT). Clinical trial registration no. NCT04006457: clinicaltrials.gov; 2022.
24. Pfizer. Study To Evaluate The Efficacy And Safety Profile Of PF-06651600 And PF-06700841 In Subjects With Alopecia Areata. Clinical trial registration no. NCT02974868: clinicaltrials.gov; 2020.
25. Pfizer. Placebo-controlled safety study of ritlecitinib (PF-06651600) in adults with alopecia areata (Allegro2a). Clinical trial registration no. NCT04517864: clinicaltrials.gov; 2022.
26. Pfizer data on f. Clinical study report: study B7931005 (ALLEGRO); 2019.
27. European Medicines A. General Considerations for Clinical Trials. London: European Medicines Agency; 1998.
28. Pfizer. Adverse reactions. *Appendix F of Company evidence submission summary to NICE* 2023.
29. Sinclair R, Lesiak A, Mehlis B. Long-term safety and efficacy of ritlecitinib in adults and adolescents with alopecia areata: interim results from the ALLEGRO-LT phase 3, open-label study. Presented at: European Academy of Dermatology and Venereology (EADV) Congress; September 7-10, 2022; Milan, Italy. Oral presentation. In: EADV; 2022.
30. Hamre HJ, Glockmann A, Kienle GS, Kiene H. Combined bias suppression in single-arm therapy studies. *Journal of Evaluation in Clinical Practice* 2008;14:923-9.
31. Loke YK, Golder SP, Vandenbroucke JP. Comprehensive evaluations of the adverse effects of drugs: importance of appropriate study selection and data sources. *Therapeutic Advances in Drug Safety* 2011;2:59-68.
32. Pfizer. Subgroup analysis. *Appendix E of Company evidence submission summary to NICE* 2023.
33. Pfizer. Published cost-effectiveness studies. *Appendix G of Company evidence submission summary to NICE* 2023.
34. Pfizer. Health-related quality of life studies. *Appendix H of Company evidence submission summary to NICE* 2023.
35. Pfizer. Cost and healthcare resource identification, measurement and valuation. *Appendix I of Company evidence submission summary to NICE* 2023.
36. Burge R, Anderson P, Austin J, et al. The patient-reported burden of alopecia areata by current severity: a real-world study in the United States [Poster 26158]. American Academy of Dermatology (AAD); 2021/03/19/, abstract no. 341.
37. Abstracts for the British Association of Dermatologists 98th Annual Meeting, Edinburgh, U.K., 3-5 July 2018. *The British Journal of Dermatology* 2018;179 Suppl 1:3-211.
38. Bhat A, Sripathy K, Wahie S, Carr M. Efficacy and cost-efficiency of diphencyprone for alopecia areata. 2011, abstract no. 98, p. 43-4.
39. National Institute for H, Care E. The reference case | Guide to the methods of technology appraisal 2013 | Guidance. 2013.
40. Ons. National Life Tables. Based on data for the years 2018-2020. 2020.
41. Winnette R, Martin S, Harris N, Deal LS. Development of the Alopecia Areata Patient Priority Outcomes Instrument: A Qualitative Study. *Dermatology and Therapy* 2021;11:599-613.
42. Lai VWY, Chen G, Gin D, Sinclair R. 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.
43. Lai VWY, Chen G, Sinclair R. Impact of cyclosporin treatment on health-related quality of life of patients with alopecia areata. *Journal of Dermatological Treatment* 2021;32:250-7.
44. Lai VWY, Bokhari L, Sinclair R. Sublingual tofacitinib for alopecia areata: a roll-over pilot clinical trial and analysis of pharmacokinetics. *International Journal of Dermatology* 2021;60:1135-9.
45. Kind P, Hardman G, Macran S. UK population norms for EQ-5D; 1999.
46. PSSRU. Unit costs of Health & Social Care; 2021.
47. England N. National Cost Collection for the NHS - National schedule of NHS costs 2020/21.

48. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. London, UK.: National Institute for Health and Care Excellence,; 2022.
49. 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.
50. Cohen SB, Tanaka Y, Mariette X, Curtis JR, Lee EB, Nash P, *et al.* Long-term safety of tofacitinib up to 9.5 years: a comprehensive integrated analysis of the rheumatoid arthritis clinical development programme. *RMD Open* 2020;6.
51. 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.
52. Bewley A, Galvan SV, Johansson E, Durand F, Petto H. Measuring the Burden of Alopecia Areata with the European Quality of Life-5 Dimensions (EQ-5D): Results from a real-world survey in 5 European countries. *Value in Health*;25:S428-9.
53. Pfizer data on f. Does the EQ-5D measure the full impact of alopecia areata on patients' quality of life? In; 2022.
54. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *Eur J Health Econ* 2015;16:927-39.
55. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health Qual Life Outcomes* 2018;16:31.
56. 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.
57. Endo Y, Miyachi Y, Arakawa A. Development of a disease-specific instrument to measure quality of life in patients with alopecia areata. *Eur J Dermatol* 2012;22:531-6.
58. Fabbrocini G, Panariello L, De Vita V, Vincenzi C, Lauro C, Nappo D, *et al.* Quality of life in alopecia areata: a disease-specific questionnaire. *J Eur Acad Dermatol Venereol* 2013;27:e276-81.
59. Ozturkcan S, Ermertcan AT, Eser E, Sahin MT. Cross validation of the Turkish version of dermatology life quality index. *Int J Dermatol* 2006;45:1300-7.
60. 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.
61. Matza LS, Stewart KD, Lloyd AJ, Rowen D, Brazier JE. Vignette-Based Utilities: Usefulness, Limitations, and Methodological Recommendations. *Value Health* 2021;24:812-21.
62. Rowen D, Brazier J, Wong R, Wailoo A. Measuring and valuing health-related quality of life when sufficient EQ-5D data is not available: NICE DSU Report. Sheffield: The University of Sheffield, UK.; 2020.
63. Pfizer data on f. Vignette study for utility estimation in Alopecia Areata. In; 2022.
64. Pennington B, Wong R. Modelling carer health-related quality of life in NICE technology appraisals and highly speciliased technologies. NICE DSU Report. Sheffield, UK: The University of Sheffield,; 2019.
65. NHS Digital. Wigs and fabric supports on the NHS. 2020. https://www.nhs.uk/nhs-services/help-with-health-costs/wigs-and-fabric-supports-on-the-nhs/ (Accessed 28/02/2023).
66. Stein EM, Yang M, Guerin A, Gao W, Galebach P, Xiang CQ, *et al.* Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States. *Health and Quality of Life Outcomes* 2018;16:1-12.
67. Sullivan PW, Valuck R, Saseen J, MacFall HM. A comparison of the direct costs and cost effectiveness of serotonin reuptake inhibitors and associated adverse drug reactions. *CNS drugs* 2004;18:911-32.

9 APPENDICES

Appendix 1

Figure 14: [redacted]

Single Technology Appraisal

Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over [ID4007]

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 23 March 2023** 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 '[redacted]' in turquoise, all information submitted as '[redacted]' in yellow, and all information submitted as '[redacted]' in pink.

Issue 1 EQ-5D is not an appropriate measure of quality of life in patients with AA

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 1.5, Page 13: "The EQ-5D has demonstrated construct validity and responsiveness in other skin diseases and estimates from the literature (Bewley et al.) suggest that it has construct validity in AA." **Section 5.3.2, Page 115:** "However, the EAG considers that other sources of data from the literature would support the construct validity of the EQ-5D in differentiating between patients with different severities of AA." **Section 5.3.4, Page 124:** "The EAG considers that the study by Edson-Heredia (2022) indicates that the EQ-5D demonstrates	**Section 1.5, Page 13:** "Despite the EQ-5D having demonstrated construct validity and responsiveness in other skin diseases, estimates from the literature (Bewley et al.) have not shown that it has construct validity in AA." **Section 5.3.2, Page 115:** "However, the EAG considers that other sources of data from the literature **may have been used for patients with different severities of AA but there is limited evidence to support the construct validity of the EQ-5D for patients with AA** ~~would support the construct validity of the EQ-5D in differentiating between patients with different severities of AA~~." **Section 5.3.2, Page 124:** "The EAG considers that the study by Edson-Heredia (2022) indicates that EQ-5D demonstrates **some**	The company CS outlines the challenges of reflecting the lived burden of AA in utility estimates. Our position continues to be that EQ-5D captures some domains broadly, but some domains are omitted. Moreover, the domains that are captured are not able to capture the level of detail to reflect the burden that is important to patients; therefore, the benefit in improving AA is being missed regardless of the source of EQ-5D data. The EAG acknowledges the difficulties in connecting the generic domains associated with EQ-5D with the AA patient (page ref). EQ-5D has been validated by clinicians and patients to lack construct validity when used in AA, misses major domains	This is not a matter of factual accuracy.

construct validity within a cohort of AA patients with mixed severity because statistically significant differences were observed in EQ-5D between mild, moderate and severe groups." **Section 5.3.4, Page 127:** "The EAG also notes that the EQ-5D has been assessed as having good validity and responsiveness in people with other skin diseases, based on a 2015 review of 16 papers covering a range of skin conditions, with most of the studies (12 of 16) being in patients with psoriasis." **Section 5.3.4, Page 127:** "the availability of data showing variation in EQ-5D utility scores by AA severity suggests that it could be used to estimate utility scores for responders and non-responders in the cost-	**differentiation by severity of HRQoL in patients with AA when measured using EQ-5D. However, there are missing elements which means HRQoL may not be fully captured** ~~construct validity within a cohort of AA patients with mixed severity because statistically significant differences were observed in EQ-5D between mild, moderate and severe groups~~." **Section 5.3.4, Page 127:** "The EAG also notes that the EQ-5D has been assessed as having good validity and responsiveness in people with other skin diseases, based on a 2015 review of 16 papers covering a range of skin conditions, with most of the studies (12 of 16) being in patients with psoriasis, **though acknowledged on p.124 the limited relevance of skin conditions to AA**." **Section 5.3.4, Page 127:** "the availability of data showing variation in EQ-5D utility scores by AA severity suggests that it ~~could be used~~ **it has some sensitivity** to estimate utility scores for responders and non-responders in the cost-effectiveness	such as social functioning, and is therefore not sensitive enough to detect potential responsiveness HRQoL.[1] We do not believe this is reflected in the statements from the EAG and would reinforce that this is not a criticism of EQ5D but that it misses important elements and therefore undervalues treatments. Moreover, justification for the construct validity of EQ-5D given that it is suitable to determine quality of life in patients with skin conditions is inappropriate. This is further established by the EAG's support of this in Section 5.3.4.10, Page 124, and again on Page 130, stating that the "*EAG is not convinced that it is [meaningful] to compare utility scores in AA with other dermatological conditions as each condition will define severity differently and each will have different domains of HRQoL affected*". Whilst	

effectiveness analysis. Therefore, the Company has not sufficiently made the case that EQ-5D is not appropriate in severe AA." **Section 5.3.4, Page 130:** "Furthermore, the availability of EQ-5D data from the literature showing that utility scores derived from the EQ-5D vary by AA severity would support the appropriateness of the EQ-5D for the purposes of informing the modelling." **Section 7.2, Page 144:** "The EAG also noted the availability of EQ-5D data from the literature which suggest that the EQ-5D does have construct validity in AA."	analysis. **However, given the relevant domains to patients with AA that are missing as advised by patients and clinical experts, the** ~~Therefore, the Company has not sufficiently made the case that~~ EQ-5D is not appropriate in severe AA." **Section 5.3.4, Page 130:** "**Whilst there is** ~~Furthermore, the availability of~~ EQ-5D data **available** from the literature showing that utility scores derived from the EQ-5D vary by AA severity**, there is a lack of construct validity in the tool to encapsulate all elements of HRQoL relevant to patients with AA** ~~would support the appropriateness of the EQ-5D for the purposes of informing the modelling~~." **Section 7.2, Page 144:** "The EAG also noted the availability of EQ-5D data from the literature which suggest that the EQ-5D ~~does have construct validity in AA~~ **may be able to detect some differences in HRQoL by severity, though the**	consideration of the utility values reported in other dermatology conditions can be useful to verify the plausible range that utility values for patients with AA may fall in, the Company agrees with the EAG's statement on current evidence and lack of direct comparability due to misalignment in the domains of HRQoL assessed in other dermatological conditions. The Company argues that there should be some caution when comparing and or overlapping AA with other skin/dermatological conditions for this reason. AA is an autoimmune dermatological disease that presents with hair loss; therefore, there is no evidence that EQ-5D sufficiently covers the burden of hair loss and associated symptoms despite the appropriateness of EQ-5D in other skin conditions.	

	burden of AA may not be captured in its entirety."		
Section 5.3.2, Page 112: "the use of a vignette study to inform the health utility values when EQ-5D outcomes were measured directly in patients in the pivotal RCT" **Section 5.3.2, Page 115:** "Therefore, the EAG argues that the Company could have used EQ-5D data from the literature to inform the model rather than the vignette study which is not consistent with the NICE reference case." **Page 5.3.4, Page 135:** "Therefore, in the absence of a reference case analysis, the EAG prefers to use the data from Bewley et al. and this has been incorporated in the EAG's base case analysis"	**Section 5.3.2, Page 112:** The Company requests that the EAG remove this sentence. **Section 5.3.2, Page 115:** "Therefore, the EAG argues that the Company could have used EQ-5D data from the literature to inform the model rather than the vignette study ~~which is not consistent with the NICE reference case~~." **Page 5.3.4, Page 135:** "**Despite this,** ~~Therefore, in the absence of a reference case analysis,~~ the EAG prefers to use the data from Bewley et al. and this has been incorporated in the EAG's base case analysis" **Section 7.2, Page 150:** "The economic analysis submitted by the Company largely complied with the NICE reference case ~~with the most important exception being that the EQ-5D data from the pivotal ALLEGRO 2b/3 trial were not used to inform the~~	It is within the NICE reference case to use alternative utility data if the Company has demonstrated that data collected within the trial is not suitable. The EAG report (page 13) acknowledges that "there may be some underestimation of QALY gains when using the utility values obtained directly from ALLEGRO 2b/3". Given this, the use of EQ-5D data from the study would undervalue the benefit in reducing the severity of AA. Therefore, this further demonstrates that the EQ-5D data from ALLEGRO 2b/3 is not suitable. Moreover, the use of a vignette is within the bounds of the NICE reference case given the preferred hierarchy of evidence. As such, the use of a vignette is in accordance with the NICE reference case	This is not a matter of factual inaccuracy. The NICE methods guide (2022) states in Table 41, which summarises the reference case, that the source of data for measurement of health-related quality of life should be those reported directly by patients or carers, or both. This does not stop the committee considering non-reference case scenarios. The EAG considers that using vignettes is outside of the reference case because the utility values are not measured in patients themselves. The EAG is simply reminding the

Section 7.2, Page 144: "The economic analysis submitted by the Company largely complied with the NICE reference case with the most important exception being that the EQ-5D data from the pivotal ALLEGRO 2b/3 trial were not used to inform the model." **Section 7.2, Page 144:** "The EAG considers this scenario to fall outside of the NICE reference case."	~~model~~**, but the EAG would have preferred to apply EQ-5D values for utility from the literature**." **Section 7.2, Page 144:** The Company requests that the EAG remove this sentence.	and it is inappropriate to suggest otherwise.	committee that the company's approach is a departure from the reference case so they can consider if it is an acceptable and justified departure.

Issue 2 The relevance of carer disutility to caregivers of patients with AA has been misrepresented

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 5.3.2, Page 112: "The EAG is not convinced that the inclusion of caregiver QALYs is sufficiently justified in this case (see section 5.3.4.11)." **Section 5.3.4, Page 130:** "The application of caregiver disutilities explicitly within cost-effectiveness analyses in TAs has typically been quite rare, occurring in only 12 out of 414 TAs conducted up to January 2019. It has been more common in HSTs, occurring in 4 out of 8 HST conducted up to the same time point but this likely reflects the severity of conditions reflected within the HST programme. The majority of the TAs including carer HRQoL decrements were in multiple sclerosis. Other diseases represented were Alzheimer's, myelofibrosis, juvenile idiopathic arthritis and moderate to severe atopic dermatitis. However, the vast majority of TAs conducted in the period covered by this review did not include carer HRQoL	The Company requests that the EAG remove these statements.	The impact on carers and caregivers, adults and adolescents is also clearly outlined in the company submission (CS B.1.3.2 and summarised in Figure 5). Furthermore, the relevance of the burden to HRQoL of carers to patients with AA was validated by clinical experts, all of whom agreed that it was relevant to adults and adolescent patients with AA. Given this, the application of caregiver disutility is in line with the NICE reference case, which states that all direct health effects of technologies should be considered. Moreover, given that the application of caregiver disutilities is within the NICE reference case, the choice of	This is not a matter of factual inaccuracy. Section 4.3.17 of the NICE methods guide (2022) states that "When presenting health effects for carers, evidence should be provided to show that the condition is associated with a substantial effect on carer's health-related quality of life and how the technology affects carers." The EAG is stating its opinion that the company has not provided sufficient evidence to meet this requirement. It is also providing the committee with some context regarding other appraisals where carer

decrements. It is unclear to the EAG why carer HRQoL decrements should be more relevant in AA than in the many other diseases considered by NICE.		other submitting companies to exclude this from their base case should bear no influence on whether caregiver disutilities are relevant for the appraisal of a new treatment. Therefore, this statement is unjustified and should not be presented in the final EAG report.	disutility has been included. It is up to the committee to decide if this context is informative.

Issue 3 There is no evidence to support that spontaneous remission is not possible in patients who have previously been treated with ritlecitinib

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 1.1, Page 5: "The Company has assumed that spontaneous remission can occur after treatment discontinuation of ritlecitinib, but the EAG prefers to assume that any ritlecitinib patients having spontaneous remission are	**Section 1.1, Page 5:** Remove this statement and include spontaneous remission for patients treated with ritlecitinib in the EAG model. **Section 1.5, Page 11:** Issue to not be presented. **Section 5.3.4, Page 120:** Issue to not be presented.	As stated by the EAG, the response of some patients to ritlecitinib may have coincided with instances of spontaneous remission. For patients treated with ritlecitinib, spontaneous remission is applied immediately on discontinuation which accounts for these cases	This is not a matter of factual accuracy. It is an alternative interpretation of the trial evidence that the EAG considers to be more plausible than the company's interpretation. Randomisation within ALLEGRO 2b/3 should have resulted in an

already included in the ritlecitinib responders". **Section 1.5, Page 11:** "The EAG believes that any cases of spontaneous remission in the ritlecitinib arm at 24 weeks will have been classified as treatment responders. Therefore, these cases are already accounted for in the model and no additional cases of spontaneous remission need to be accounted for in the model when patients discontinue ritlecitinib treatment." **Section 5.3.4, Page 120:** "The EAG would argue that any cases of spontaneous remission for patients in the ritlecitinib arm of the ALLEGRO 2b/3 trial would have been incorporated within the group of patients deemed to have responded to ritlecitinib and therefore would have already been		where response coincides with spontaneous remission. In line with patients treated with BSC, the proportion of patients with spontaneous remission is assumed to be constant reflect that some patients will lose and others will gain spontaneous remission over time. Moreover, there is no evidence to support that treatment with ritlecitinib will impact the likelihood of a patient achieving spontaneous remission over time. If a patient does respond to treatment and then discontinues treatment, they may well go on to experience spontaneous remission in the future. This is accounted for in the assumption that a constant proportion of patients have remission to account for patients gaining and losing spontaneous remission over time.	equal incidence of spontaneous remission in both arms. In the ritlecitinib arm these patients will have been classed as treatment responders and their outcomes will already be captured in the proportion remaining on treatment and responding long-term in the model. The EAG does not believe that it is correct to apply the spontaneous remission rate to patients stopping treatment as most stop treatment due to non-response. It would be more accurate to ensure that a fixed proportion of ritlecitinib responders remain on treatment life-long because they are in spontaneous remission but have been misclassified as

captured within the modelled outcomes for 48 weeks and the status of these patients as continued responders in the ALLEGRO-LT study. Therefore, allowing for a further incidence of spontaneous remission when treatment non-responders switch to BSC is not appropriate and the EAG preferred to exclude spontaneous remission on discontinuation from ritlecitinib in their base case (see Section 5.4.2.5). The EAG's clinical experts advised that spontaneous remission was extremely unlikely in their experience in patients with severe AA and became less likely beyond the first 6 months to a year. Therefore, the Company's scenario that explored higher rates of spontaneous remission was not considered realistic."		Furthermore, if a patient discontinues treatment with ritlecitinib due to a lack of response, this does not prevent them from going on to experience spontaneous remission in the future. Therefore, exclusion of the assumption that patients previously treated with ritlecitinib could not experience spontaneous remission is biased against ritlecitinib. Further to this, the assumption to apply spontaneous remission to discontinuers of ritlecitinib was justified with clinicians, who agreed it was relevant to include for patients treated with ritlecitinib.[1] This is because some patients will discontinue treatment with ritlecitinib following response to treatment, which may be indefinitely sustained in an unknown number of patients.	responders. However, the company's model does not accommodate this as an option as there is no separate health state for patients in spontaneous remission who remain on treatment. The company could address this by applying a treatment

Issue 4 It is incorrect to state that the long-term evidence from ALLEGRO-LT is insufficient to support the long-term efficacy of ritlecitinib in patients with severe AA

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 1.4, Page 6: **"**Further evidence from the ALLEGRO-LT study may elucidate this, although the EAG notes that, as the inclusion criteria allowed for those with milder AA (proportion of scalp hair loss ≥25%), it would be difficult to assess the long-term effects of ritlecitinib on those with severe AA (proportion of scalp hair loss ≥50%)." **Section 7.1, Page 143** "Further evidence from the ALLEGRO-LT study may elucidate this, although the EAG notes that, as the inclusion criteria allowed for those with milder AA (SALT ≥25), it would be difficult to assess the long-term effects of ritlecitinib on those with severe AA."	The Company requests that the EAG remove these statements.	Whilst the ALLEGRO-LT study included patients with a baseline SALT score ≥25, only data from patients in the modified *de novo* group were included in the economic modelling. This was stated in response to question B1.E at clarification, the modified *de novo* group in the ALLEGRO-LT trial considered which was used to inform transition matrices and treatment discontinuation after 48 weeks in the economic model included "*Participants with a screening or baseline SALT score ≤50"*. This means that the economic model provides assessment of the long-term effects of ritlecitinib on those with severe AA at baseline and, therefore, the statements are factually inaccurate.	This is not a matter of factual accuracy. These statements relate to the clinical effectiveness evidence not the data used in the model. Figures 23 and 24 of the CS pertain to the whole *de novo* cohort not the modified *de novo* cohort used to inform the modelling. The company has not presented any clinical effectiveness data for the subgroup of participants in the ALLEGRO-LT study with a proportion of scalp hair loss of ≥50%. The size of the 'modified de novo group' is not clear from the baseline characteristics reported in either CS, Table 12 or section 10.4 of the CSR and this subgroup does not appear to have been defined a priori.

Section 7.2, Page 145 "the data used to estimate long-term efficacy"	**Section 7.2, Page 145** The Company requests that the EAG remove this statement.	As above, the data used to estimate long-term efficacy is aligned with the anticipated positioning of ritlecitinib. As such, the Company do not believe that the choice of data for ritlecitinib long-term effectiveness is a limitation.	The quote from section 7.2, page 145 does not relate to the inclusion of patients with SALT score 25 to 50 within the data informing the economic model as these were excluded from the long-term transition matrices as the company states. The key issue being referred to here is related to the doses received in ALLEGRO-LT (see EAG report section 5.3.4.5) and the evidence to support the assumption of no treatment waning (see EAG report section 5.3.4.4)

Issue 5 It is incorrect to state that discontinuation of a JAK inhibitor in patients with severe AA would not be aligned with the retention observed in the ALLEGRO-LT study

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 1.1, Page 5: "the EAG prefers to assume a higher treatment discontinuation rate than the rate estimated by the	The Company requests that the EAG remove these statements.	It is factually inaccurate to assume that mean time on treatment with a JAK inhibitor for patients with AA will be equal to mean time on	This is not a factual inaccuracy. The EAG acknowledges that there are no data to inform duration of

company as the company's approach results in a mean time on treatment that is much higher than that observed when JAK inhibitors have been used long-term in other indications" **Section 1.5, Page 12:** "In addition, the EAG believes the discontinuation rate in the company's base case analysis results in a mean time on treatment that is too high in comparison to the mean duration on treatment when JAK inhibitors are used in other indications where longer follow-up is available." **Section 5.3.4.7, Page 122:** "The EAG would argue that the very low rate of discontinuation observed in the model leads to a predicted mean time on treatment of 5.9 years, which is unrealistically high.		treatment with a JAK inhibitor for patients with other conditions. It is not known how time on treatment with JAK Inhibitors varies across different conditions and, therefore, there is no evidence to support this assumption. This was supported by dermatologists with a specialist interest in hair disorders.[1] For example, one dermatologist commented that "*rheumatoid arthritis is very different"* due to factors including the average age of patients being older and more patients having comorbidities on average compared to patients with AA. This suggest treatment with a JAK inhibitor could be better tolerated in patients with AA compared to patients with rheumatoid arthritis. They added that "*it is important to note the difference in comparing apples with pears*	treatment in the long-term for JAK inhibitors in this indication because it is a new treatment. In the absence of these data, the EAG prefers to assume that mean duration on treatment will be similar to other indications where JAK inhibitors have been used long-term.

Long-term follow-up studies of other JAK inhibitors for rheumatoid arthritis with around 9.5 years of patient follow-up are reporting median durations on treatment of 3.1 years and 4.6 years for tofacitinib and baricitinib, respectively (mean durations of 3.2 and 3.9 years, respectively)." **Section 5.3.4.7, Page 123:** "This results in a mean time on treatment of 3.6 years for the ritlecitinib arm which is more consistent with the mean time on JAK inhibitors seen in rheumatoid arthritis where long-term follow-up studies are available."		*– same class of treatment but different kinase and a different disease."* Furthermore, dermatologists with a specialist interest in hair disorders considered extrapolation of data from the ALLEGRO-LT trial to be appropriate to estimate long-term treatment discontinuation (3/3 clinicians, 100%).[1]	

Issue 6 Positioning of ritlecitinib is incorrectly summarised

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG response
Section 2.3, Page 19:	To update text as follows:	The Company has not proposed that topical corticosteroids or contact	The EAG based this statement on the company's positioning in

“Ritlecitinib is proposed as a second- or third-line treatment for AA, after topical corticosteroids and contact immunotherapy (where available), as a systemic treatment option”	“Ritlecitinib is proposed ~~as a second- or third-line treatment for AA, after topical corticosteroids and contact immunotherapy (where available),~~ as a systemic treatment option **for patients with AA**.”	immunotherapy should be defined as lines of therapy prior to treatment with ritlecitinib due to variability in access to treatment options, patient preference and the limitations of existing options described at length throughout Section B.1.3.3 in the CS. The Company recommend that the positioning of ritlecitinib should be aligned to the technology’s full marketing authorisation for this indication to avoid restricting or delaying access to treatment with ritlecitinib for patients with AA.	Figure 10 of the CS. However, it accept that these topical treatments may not always be available or suitable in severe AA as indicated by the left hand arrow in Figure 10 which bypasses these treatments. The EAG has amended its text to say “Ritlecitinib is proposed as a systemic treatment option for severe AA either after topical corticosteroids and contact immunotherapy, or when these topical treatments are not suitable or available.”

Issue 7 Factually inaccurate statements

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Section 1.2, Page 6: "The modelling assumptions that have the greatest effect on the ICER are: … inclusion of age-adjustment for utilities."	Inclusion of age-adjustment for utilities to be removed from the list.	The "inclusion of age-adjustment for utilities" has not been conducted as a scenario meaning the statement is assumed and there is no evidence to support the assumption that it will be impactful.	The EAG has removed the bullet from the list as suggested. However, the EAG maintains that the lack of age-adjustment in the utilities is likely to have led to an underestimation of the ICER as stated in section 5.3.4.15.
Section 1.5, Page 14: "The EAG is also not convinced that the caregiver disutility provided by the vignettes is accurate as the company assumed that caregivers of adolescents with mild to moderate AA would have no disutility."	The company request that this statement is removed.	Exclusion of caregiver disutility for patients with mild and moderate disease is conservative, as the potential incremental benefit for ritlecitinib is not fully presented. Moreover, exclusion of patients with mild and moderate patients from the vignette study does not	This is not a factual inaccuracy. Assuming a carer disutility for severe AA and none for mild/moderate AA means that there is a QALY gain achieved from moving patients

		preclude errors for the patients with severe AA.	from severe to mild/moderate AA. If the company had estimated the disutility for mild/moderate AA and this was found to be a non-zero value, then the carer QALY gain achieved from patients responding to treatment would be smaller. Therefore, the company's approach is not conservative and may in fact have overestimated QALY gains in carers.
~~Section 1.7, Page 15:~~ ~~"due to the continuity corrections required to handle missing observations in the long-term transition matrices, which is only applied in the probabilistic analysis."~~			The EAG is unclear what the company's believes to be factually inaccurate in this statement and has therefore not

			amended the statement.
Section 2.1, Page 17: "Most patients in remission will experience further episodes of alopecia areata and around 30% of patients with patchy hair loss will eventually progress to complete hair loss."	To update text as follows: "Most patients in remission will experience further episodes of alopecia areata and around ~~30% of patients with patchy hair loss will eventually progress to complete hair loss~~ **5% of people with patchy disease will progress to develop alopecia totalis**."	We could not identify 30% in the reference provided.[2] The statement added is sourced from Xu *et al.* (2017).[3]	We had cited the wrong paper – the 30% figure is given in Harries *et al.* (2010).[1] Therefore, to consider both estimates, the text has been amended to: "Most patients in remission will experience further episodes of alopecia areata and around it is estimated that between 5% and 30% of patients with patchy hair loss will eventually progress to alopecia totalis.[1, 2]"
Section 3.1, page 20: "The EAG notes that the Company's base case model reflects the overall population (aged 12	Statement to be updated to:	The baseline characteristics in the base case of the economic model are a weighted average	The company has misunderstood the point being made.

years and over) using average baseline characteristics rather than based on a weighted average of age subgroups."	"The EAG notes that the Company's base case model reflects the overall population (aged 12 years and over) using **weighted** average baseline characteristics ~~rather than based on a weighted average of age subgroups~~."	according to the ALLEGRO 2b/3 trial population. As stated in response to B40 during clarification, the mean age was calculated using the formula displayed below: $\bar{x}_{all} = (\bar{x}_{adolescent\ age} \times \%_{adolescent}) + (\bar{x}_{adult\ age} \times (1 - \%_{adolescent}))$	The EAG has amended the sentence to make this clearer but notes that the reader is already referred to section 5.3.4.1 for a fuller explanation of this issue.
Section 4.1.1, page 25: "The Company searched several electronic bibliographic databases in October 2021 (original review) followed by an update in November 2022 (Appendix D.1 Identification and selection of relevant studies)."	To update the text as follows: "The Company searched several electronic bibliographic databases in October 2021 (original review) followed by an update in ~~November~~ **September** 2022 (Appendix D.1 Identification and selection of relevant studies)."	The SLR was initially performed in October 2021 and was updated in September 2022.	The text has been amended as suggested by the company.
Section 4.1.1, page 25:	To update the text as follows:	For the original review, conference abstracts were	The text has been amended as

"The Company searched several key conference abstract and society presentation websites in the last three years (2020-2021)."	"The Company searched several key conference abstract and society presentation websites ~~in the last three years (2020-2021)~~ **from 2020-2022**."	searched from 2020 to 2021 to retrieve the latest studies. For the SLR update, the same conference proceedings were searched to identify conference abstracts published between October 2021 and September 2022.	suggested by the company.
Section 4.1.2, page 26: "The Company's clarification response (questions A1 and A2) clarified that non-pharmacological clinical management was not considered as a comparator in the SLR, however, placebo was considered as a comparator"	To update the text as follows: "The Company's clarification response (questions A1 and A2) clarified that non-pharmacological clinical management was not considered as a comparator in the SLR, however, placebo was considered as a comparator **if the intervention was a comparator of interest"**	Missing information.	The text has been amended as suggested by the company.
Section 4.2.1, page 29 (Table 4):	To update the text as follows:	Misspecification.	The text has been amended as

"Adolescents and adults aged ≥12 years with clinical diagnosis of AA, ≥50% scalp hair loss (SALT ≥50) including AT and AU, no regrowth ≤6 months, current episode ≤10 years."	"Adolescents and adults aged ≥12 years with clinical diagnosis of AA, ≥50% scalp hair loss (SALT ≥50) including AT and AU, no regrowth ~~≤6 months~~ **within 6 months**, current episode ≤10 years."		suggested by the company.
Section 4.2.1.1, page 31: "Therefore, it seems unlikely that large numbers of relevant patients were excluded based on this criterion, although the EAG notes that this is uncertain as no details were provided as to how patients were identified and recruited into the study."	The Company requests that the EAG remove this sentence.	The EAG provide no justification for the inclusion of this sentence. Without knowing the method of identification, there is no evidence to support this statement.	This is not a matter of factual inaccuracy. The company did not provide this information that the EAG requested at clarification. The information that the EAG required regarding the identification and recruitment of patients was not in the CSRs, as described at clarification response. The EAG needs to know details on how the

			patients recruited into the trials were identified, for instance, were all patients attending a specific clinic screened for recruitment into the study or were certain patients selected by trials clinicians based on clinicians' opinion about their prognosis? Did every patient who met the inclusion criteria stand an equal chance of being recruited? The EAG views this as important because it could potentially impact on the trial results.
Section 4.2.3.1, page 49: "The EAG concludes that the ALLEGRO 2b/3 study is at moderate to high risk of bias; the	To update the text as follows:	The only apparent issue identified by the EAG during their quality assessment was that the method randomisation	The text has been amended as suggested by the company as the

Company did not provide a summary appraisal of risk of bias."	"The EAG concludes that the ALLEGRO 2b/3 study is at **an unclear** ~~moderate to high~~ risk of bias; the Company did not provide a summary appraisal of risk of bias."	and allocation concealment is unclear, due to lack of clarity relating to how patients were identified and recruited into the study. Their assessment of risk of bias being 'moderate to high risk' is disproportionate considering what was identified.	EAG agrees that a rating of 'unclear' would be more appropriate. Text in the paragraph that follows the amended one has also been amended accordingly to: "Based on the information available, the EAG judged risk of bias of the ALLEGRO 2a proof of concept study and the ALLEGRO-2a safety study to be unclear, due to the lack of information available on the conduct of the study."
Section 4.2.4.2, page 69: "At Week 24 was significantly greater in the ritlecitinib arms than in the placebo arms"	The Company requests that the EAG remove this sentence.	The proportion of patients with an EBA response (≥2-grade improvement from baseline/score of 3) at Week	The word 'significantly' has been deleted from this sentence, as

		24 was not statistically significant as it was not powered.	the sentence is otherwise factually accurate.
Section 4.2.4.2, page 69: "Which represented a statistically significantly difference (p<0.01)."	The Company requests that the EAG remove this sentence.	It did not represent a statistically significant difference as it was not powered.	The text has been amended as suggested by the company.
Section 4.2.4.2, page 69: "At Week 24 was significantly greater in the ritlecitinib arms than in the placebo arms"	The Company requests that the EAG remove this sentence.	The proportion of patients with an ELA response (≥2-grade improvement from baseline/score of 3) at Week 24 was not statistically significant as it was not powered.	The word 'significantly' has been deleted from this sentence, as the sentence is otherwise factually accurate.
Section 4.2.4.2, page 69: "Which represented a statistically significantly difference (p<0.001)."	The Company requests that the EAG remove this sentence.	It did not represent a statistically significant difference as it was not powered.	The text has been amended as suggested by the company.
Section 4.2.4.3, page 73: "The proportion of patients with an IGA rating of no change or further loss at Week 24 was greater in the combined placebo arm ([redacted]) than the ritlecitinib 200/50 mg arm ([redacted]), whereas the proportion of patients with an IGA rating of 1-24% regrowth ([redacted]), 25-49%	The Company requests that the EAG rephrase this sentence, as the IGA rating of 1-24% regrowth for placebo [redacted] is higher than the	Misrepresentation of data.	This text has been amended to: "The proportion of patients with an IGA rating of no change or further loss at Week 24

regrowth ([redacted]), 50-74% regrowth ([redacted]), 75-99% regrowth ([redacted]) and 100% regrowth ([redacted]) at Week 24 was greater than for the combined placebo arm ([redacted]).”	[redacted] that is outlined for ritlecitinib.		was greater in the combined placebo arm ([redacted]) than the ritlecitinib 200/50 mg arm ([redacted]), whereas the proportion of patients with an IGA rating of 25-49% regrowth ([redacted]), 50-74% regrowth ([redacted]), 75-99% regrowth ([redacted]) and 100% regrowth ([redacted]) at Week 24 was greater than for the combined placebo arm ([redacted]), and the prortion of patients with an IGA rating of 1-24% regrowth was greater for the combined placebo arm ([redacted]) than for the ritlecitinib

			200/50 mg arm ([redacted]).”
Section 4.2.4.6, page 86: “With no overlap in the 95% CI, indicating a statistically significant difference.”	The Company requests that the EAG remove this sentence.	The study is not powered to detect a difference in a sub group and therefore this statement may be misleading without qualification.	The ERG does not consider this a factual inaccuracy as non-overlapping CI are indicative of statistically significant differences. The report has been amended to clarify that the study was not powered to test for subgroup differences.
Section 4.2.4.6, page 86: “The EAG notes that in addition to AA severity, subgroup analyses also indicate statistically significant differences in treatment effect for gender”	The Company requests that the EAG remove this sentence.	The study is not powered to detect a difference in a sub group and therefore this statement may be misleading without qualification.	As above. Qualification has been added to the report.
Section 5.2.4, page 96 “Spontaneous remission (i.e., transitioning to SALT ≤10) is assumed to occur to a fixed proportion of patients (1.54%) starting treatment with BSC.”	The Company encourages the EAG to clarify that this occurs from Week 24 onwards.	Some relevant data is omitted.	To clarify the timing, the EAG has amended the following statement to say “This applies

			both to the BSC treatment arm at week 24, and to patients transitioning to BSC who are discontinuing from ritlecitinib in the first cycle after they start BSC (which does not occur before week 24)."
Section 5.2.5, page 97 (Table 17) "An estimated rate ratio applied to the incidence of TEAEs on ritlecitinib"	The Company requests that EAG include the source of the parameter, which is ALLEGRO 2b/3.	Some relevant data is omitted.	Amended to say "A rate ratio applied to the incidence of TEAEs on ritlecitinib estimated from ALLEGRO 2b/3[3]"
Section 5.2.5.4, page 99: "Each transition matrix contained 3-month probabilities which were assumed equivalent to 12-week probabilities."	The Company requests that the EAG rephrase this sentence. The matrix is 12-week probabilities, assumed equivalent to 3-months for the purpose of the model.	Misrepresentation of the data.	The model uses 12-week rather than 3-month cycles so the company's statement cannot be true. The reporting of response outcomes

			from ALLEGRO-LT is at 12, 15, 18, 21 and 24 months and these are the time periods used to describe the transition matrices in the model. The company stated in response to clarification question B1 that “A twelve-week block was considered as three months in the ALLEGRO-LT study protocol therefore, Month 3 and Week 12 are interchangeable” which doesn’t really clarify the issue. Whatever time periods these data were estimated from, they were assumed equivalent to 12-week probabilities. Therefore, the EAG

			has not made the suggested amendment but would welcome further clarity on this from the company.
Section 5.2.5.10, page 105: "Given the company's concern regarding the direct measures of utility included in the trial outcomes, and the lack of data in the literature which they considered acceptable for use in the model, the company conducted a vignette study to inform the economic analysis".	The Company requests that the EAG rephrase this sentence. "Given the company's concern regarding the direct measures of utility included in the trial outcomes, and **only limited, unacceptable** ~~the lack of~~ data **available** in the literature ~~which they considered acceptable~~ for use in the model due to the limitations in EQ-5D, the company conducted a vignette study to inform the economic analysis".	Rephrased for clarity.	Not a matter of factual accuracy

Section 5.2.5.10, page 105: "The first phase was to develop draft vignettes informed by qualitative semi-structured interviews with patients (N=3 adults; N=3 adolescents) and carers (N=5), a detailed literature review and a retrospective analysis of the ALLEGRO 2b/3 trial data."	The Company notes that this phase also included a review of AAPPO.	Misrepresentation of the approach.	The EAG does not know what "a review of AAPPO" refers to specifically. The company analysed HADS and AAPPO data from the ALLEGRO 2b/3 trial. The EAG described this as, *"a retrospective analysis of the ALLEGRO 2b/3 trial data".* The EAG has now amended the text to say *"a retrospective analysis of the HADS and AAPPO outcomes from the ALLEGRO 2b/3 trial,"* and hope this addresses the company's concern.

Section 5.2.5.11, page 108: "TEAEs were assumed to be managed through admission rather than primary care, thus the unit costs were sourced from the National Schedule of NHS Costs 2020-21, as clarified in the Company's response to clarification question B11.[5]"	The Company encourages the EAG to review the choice of reference 5, as it is not linked.		This should have referred to the company's clarification response and has been amended accordingly.
Section 5.3.4.3, page 117: "Clinical evidence used when implementing the interim stopping rule" – entire section	The Company is raising this regarding clarification of the response.	Thank you for highlighting the confusion surround derivation the number of patients who pass the stopping rule at Week 24 ([redacted]) and the number of patients carried forward for assessment of response at later time points ([redacted]). The Company appreciates the nuance of implementing this approach in the economic model raised by the EAG and will work to resolve this during technical engagement. To clarify the confusion for the EAG, the Company would like to highlight that the response to B5 at clarification contains a typo. At Week 24, there were [redacted] patients missing due to	The EAG has modified this section to account for the additional information provided by the company in their FAC response. However, the main thrust of the EAG's argument in this section still holds. This was that the [redacted] patients missing due to reasons other than COVID-19 should not have been assumed to be steady/improved for the purposes of the interim stopping

		COVID-19 and [redacted] patients missing due to reasons other than COVID-19, assumed missing at random. We hope that this resolves the confusion on the derivation of patient numbers.	rule at 24 weeks but missing at random for the purposes of the response rates after 24 weeks. The EAG hopes this can be further explored and satisfactorily resolved by the company at technical engagement.
Section 5.3.4.7, page 122 (Figure 13)	The company request that the ERG revise the figure or remove it to present a relevant comparison of data.	The Company have concerns regarding Figure 13 considering the drop to 45% on treatment occurs at the final stopping rule. Given that there is a stopping rule in the model for patients to discontinue treatment if SALT score is >20 at Week 48, the output of the model data and the KM for patients whose SALT score remains below SALT 20 are not comparable.	The EAG was attempting to demonstrate in Figure how the slopes for discontinuation compare between the model and the K-M data from the trial. However, it acknowledges that the fact it was plotting the proportion of the whole population

			starting treatment rather than the proportion of the group responding at 48 weeks made the comparison less clear. The EAG has updated Figure 13 so that it uses information from the 'Data after discontinuations' column and the proportions are presented as a proportion of those on responding at 48 weeks.
Section 5.3.4.10, page 129: "Finally, the EAG believes that the vignettes lack face validity when compared with the quantitative data the Company has used to inform them. For example, the vignette for SALT>50 (A3) states, [redacted] whereas the data from the ALLEGRO 2b/3 trial (Table 8, Appendix H) suggest that the most common response to,	Sentence(s) to be removed.	The Company believes that it is not plausible for the EAG to make this comparison. The purpose of the vignette development was to provide greater level of nuance and granularity to patients with AA's lived experience, and what truly matters to them. Best practise was followed in	This is not a matter of factual accuracy. If the vignettes have been developed using quantitative data on the experience of patients it is reasonable to consider how the

"over the past week, how often did you feel embarrassed about your hair loss?" in patients with a SALT score of 50-100 was never ([redacted]), followed by rarely ([redacted]). In fact, only [redacted] and [redacted] responded sometimes, often, and always respectively. Similarly, [redacted] of patients with a SALT score of 50 to 100 responded 'not at all' when asked, "over the past week, how much did you limit your interactions with others because of your hair loss?" This contrasts with a description of, "[redacted]," for the relevant vignette (A3). The draft vignette for SALT>50 used the wording, "[redacted]," which seems less definite than the wording in the final vignette. The impact of patient and clinician feedback on the vignettes, as summarised in Appendix H Table 12, shows an increase in frequency for many of the domains with the CS stating that "there was consensus that the severity of the impacts was understated across all health states." Although it is best practice to have draft vignettes validated by patients and clinicians, this does in this case appear to have resulted in quite a discrepancy between the quantitative data used to inform the vignettes and the final		terms of the vignette development, as described by Matza *et al.* (2021).[4] To have prioritised what the trial said instead of refining based on an additional detailed literature review (including background to the validation of AAPPO, and the qualitative feedback from AA patient interviews and clinician experts to inform and test the vignettes) would undermine the feedback we received and ignore the opportunity to accurately capture the vignette. ALLEGRO 2b/3 had a select population. For example; selective enrolment (i.e., patients with psychological involvement were not eligible for enrolment) in ALLEGRO 2b/3 means that patients are unlikely to experience this response to hair loss. Baseline characteristics (i.e., time since first diagnosis) in ALLEGRO 2b/3 means that	final vignettes compare to the quantitative data on which they are based.

wording. On this basis, the EAG would question whether the utility values estimated accurately reflect the average experience of patients with the relevant SALT score.”		patients may have shown adaptation to their AA; therefore, they are unlikely to experience the same level of psychological response to recent hair loss. Given this, a more comprehensive approach was taken to inform the vignettes. We argue this more accurately reflects the lived experience of AA from a UK-based cohort.	
Section 5.3.4.10, page 130: “For example, severe hidradenitis suppurativa is associated with significant pain which is not present in severe AA”.	Sentence(s) to be removed.	To contextualise the burden of severe AA to patients in the UK, the CS (B.1.3.2) discussed concepts from published literature and confirmed the relevance via qualitative interviews with representatives from the UK AA community. Ten representatives from six Patient Advocacy Groups (PAG) focused exclusively on patients with AA with ≥50%	This is not a matter of factual accuracy. The EAG checked with their clinical advisors that the degree of pain experienced in severe hidradenitis suppurativa is not comparable with the degree of pain experienced in severe AA.

		hair loss (i.e., severe AA) living within the UK. Figure 5 (CS B.1.3.2) outlines the physical impacts of alopecia areata which includes pain. Therefore, comparison of the utility values obtained for patients with AA in the vignette study with those of hidradenitis suppurativa in the literature is appropriate due to overlapping characteristics of the burden of disease.	
Section 5.4.2.11, page 135: **"**We have also conducted exploratory analyses in which we halve the number of appointments for psychological support and the number of wigs received in both arms"	Sentence and subsequent analyses to be removed.	To conduct this scenario is implausible as the modelled resource use should reflect the access to resources that patients have access to, which was informed by clinical opinion based on their experience of managing patients with AA in the UK.	This is not a matter of factual accuracy.
Section 5.4.3.1, page 140 (Table 29): EAG base case applying analyses 1-9 (Probabilistic)* Ritlecitinib	Extreme outlier ICER, the Company encourages the EAG to review this figure [redacted]	The Company encourages the EAG to review this figure as this reflects the probabilistic ICER likely being skewed from the deterministic ICER. This	The EAG has reviewed this and does not believe that the discrepancy

ICER [redacted]		highlights a potential error in the EAGs revision to the model.	between the PSA and deterministic results indicates an error introduced by the EAG but if the company wishes to identify something specific for the EAG to consider then this could be done at technical engagement.
Section 7.2, Page 144: "The EAG also considered that the utility values generated from the vignette study should be treated with caution given their concerns regarding the lack of correspondence between the quantitative data used to inform the vignettes and the wording in the final vignettes"	The Company requests that the EAG remove this sentence.	The Company used validated AAPPO to generate the initial draft vignettes. These were then validated with patients with AA to ensure the lived experience of patients by hair loss severity was accurately captured. The EAG acknowledges the methods of the vignette are appropriate and in line with good practice as outlined by Matza *et al.* (2021).[4]	This is not a matter of factual inaccuracy.

Issue 8 Typographical errors

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Throughout the EAG report	To update the text as follows: "Ritlectinib" should be replaced with "Ritlecitinib" throughout the report.	Typographical error (spelling).	This spelling has been corrected throughout the report.
Throughout EAG report	To update the text as follows: "120/50 mg" should be replaced with "**200/50** mg" throughout the report. For example, "Although the ~~120/50 mg~~ **200/50** mg placebo arm was comparable [redacted]"	Typographical error (wrong number).	"120/50 mg" has been replaced with "200/50 mg" throughout the report.
Throughout EAG report	To update the text as follows: "200/40 mg" should be replaced with "200/**50** mg" throughout the report. For example, "withdrawal by participant appeared to be more prevalent in the ~~"200/40 mg"~~ 200/**50** mg"	Typographical error (wrong number).	"200/40 mg" has been replaced with "200/50 mg" throughout the report.

Throughout EAG report.	To update the text as follows: "Daily" should be replaced with **"once** daily**"** when referring to the dose of ritlecitinib. For example, "Ritlecitinib 50 mg **once** daily (n=130)".	Misspecification.	"Daily" has been replaced with "once daily" throughout the report when referring to the dose of ritlecitinib.
Section 4.2.1.1, Page 31-35 (Table 5): The key exclusion criteria for the ALLEGRO 2a proof of concept study has not been included.	The Company requests that the EAG detail the exclusion criteria for the ALLEGRO 2a proof of concept study.	Typographical error (omission). Exclusion criteria for the ALLEGRO 2a proof of concept is included in CSR provided as part of the original company submission.	The exclusion criteria for the ALLEGRO 2a proof of concept study as reported in the CSR have been added to Table 5.
Section 4.2.1.1, Page 36: "The CSRs for the ALLEGRO 2b/3 study, the ALLEGRO-LT study, the ALLEGRO 2a proof of concept study, and the ALLEGRO-2a safety study do not report any detail on how patients were identified and recruited (in the case of ALLEGRO-LT, this relates to the de novo patients). The company's clarification response (question A18) outlined eligibility criteria, but	The Company requests that this is removed from the EAG report.	During the clarification questions, the Company described the data that they were going to provide in response to this question. Unfortunately, only one member of the EAG was able to attend. The Company appreciates that there was a limited response available at this point and confusion may have arisen on the required information to be provided.	This is not a matter of factual inaccuracy. The company did not provide this information that the EAG requested at clarification. The information that the EAG required regarding the identification and recruitment of patients was not in the CSRs, as described at clarification response. The EAG needs to know details on how the

did not clarify how participants were identified and recruited. Therefore, the EAG cannot assess whether the process of recruitment may have introduced selection bias into these studies, nor whether a representative sample of patients with severe AA is likely to have been recruited to each.”		The Company would like to take the opportunity to clarify the approach to identifying and recruiting participants for these studies in the hope that it will allay the EAG’s concerns.	patients recruited into the trials were identified, for instance, were all patients attending a specific clinic screened for recruitment into the study or were certain patients selected by trials clinicians based on clinicians’ opinion about their prognosis? Did every patient who met the inclusion criteria stand an equal chance of being recruited? The EAG views this as important because it could potentially impact on the trial results.
Section 4.2.1.1, Page 36: “Participants in the ALLEGRO-2a safety study were also required to be aged ≥18 (and ≤50) years, and were could have less severe AA than specified in the final NICE scope”	To update the text as follows: "Participants in the ALLEGRO-2a safety study were also required to be aged ≥18 (and ≤50) years and ~~were~~ could have less severe AA than specified in the final NICE scope."	Typographical error (grammar).	The text has been amended as suggested by the company.

Section 4.2.1.1, page 36: "Table 6 in the Company's clarification response (question A23) clarifies that [redacted] patients ([redacted]% of 603 roll-over patients; [redacted]% of 1050 treated patients)"	To update the text as follows: "Table 6 in the Company's clarification response (question A23) clarifies that [redacted] patients ([redacted]% of 603 roll-over patients; [redacted]% of 1050 treated patients)"	Typographical error (wrong number).	The text has been amended as suggested by the company.
Section 4.2.1.1, page 38: "The 50 g placebo arm"	To update the text as follows: "The ~~50 g~~ 50 **mg** placebo arm"	Typographical error (spelling).	The text has been amended as suggested by the company.
Section 4.2.1.5, page 48: "The EAG considers the design of PEDFIC2 to be open to potential biases such as attrition bias, natural recovery and regression to the mean (particularly in relation to efficacy), due to being open-label and single-arm."	To update the text as follows: "The EAG considers the design of ~~PEDFIC2~~ **ALLEGRO-LT** to be open to potential biases such as attrition bias, natural recovery and regression to the mean (particularly in relation to efficacy), due to being open-label and single-arm." The EAG should also consider the inclusion of "attrition bias, natural recovery and regression to the mean".	Typographical error.	The text has been amended as suggested by the company.

Section 4.2.4.1, page 61 (Figure 4):	Update the reference.	The reference on this figure doesn't match with the EAG report references.	The reference has been updated.
Section 4.2.4.1, page 63 (Figure 6):	Update the reference.	The reference on this figure doesn't match with the EAG report references.	The reference has been updated.
Section 4.2.4.1, page 64: [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.1, page 65: "28.30 to 20.02"	To update the text as follows: **"32.33 to 50.02"**	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.1, page 65: "20.55 to 68.42"	To update the text as follows: **"50.55 to 68.42"**	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.2, page 69: "In the ALLEGRO-LT study, the proportion of patients with an ELA response (≥2-grade improvement from baseline/score of 3) at Week 24 (Month 6) was greater than those in the ALLEGRO 2b/3 study in both the de novo	To update the text as follows: "In the ALLEGRO-LT study, the proportion of patients with an ELA response (≥2-grade improvement from baseline/score of 3) at Week 24 (Month 6) was greater than those in the ALLEGRO 2b/3 study in both the de novo	Typographical error (wrong number).	This text has been amended.

[redacted] and roll-over [redacted] patients."	[redacted] and roll-over patients."		
Section 4.2.4.2, page 69: "By Week 48 (Month 12), the proportion of patients with an EBA response had reached [redacted] in the de novo cohort and [redacted] in the roll-over cohort."	To update the text as follows: "By Week 48 (Month 12), the proportion of patients with an ~~EBA~~ **ELA** response had reached [redacted] in the de novo cohort and [redacted] in the roll-over cohort."	Typographical error (wrong outcome).	This text has been amended.
Section 4.2.4.3, page 73: "By Week 48 (Month 12), the proportion of patients with an EBA response had reached [redacted] in the de novo cohort and [redacted] in the roll-over cohort."	To update the text as follows: "By Week 48 (Month 12), the proportion of patients with an ~~EBA~~ **PGI-C** response had reached [redacted] in the de novo cohort and [redacted] in the roll-over cohort."	Typographical error (wrong outcome).	This text has been amended.
Section 4.2.4.4, page 74 (Table 15, ALLEGRO 2b/3, ritlecitinib, 10 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.4, page 74 (Table 15, ALLEGRO 2b/3, ritlecitinib, 10 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.

Section 4.2.4.4, page 74 (Table 15, ALLEGRO 2b/3, ritlecitinib, 30 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.4, page 74 (Table 15, ALLEGRO 2b/3, ritlecitinib, 30 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.4, page 74 (Table 15, ALLEGRO 2b/3, ritlecitinib, 10 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.4, page 75 (Table 15, ALLEGRO 2b/3, ritlecitinib, 30 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.4, page 75 (Table 15, ALLEGRO 2b/3, ritlecitinib, 10 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 4.2.4.4, page 77 (Table 15, ALLEGRO-LT,	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.

ritlecitinib, roll-over cohort 50 mg): [redacted]			
Section 4.2.4.5, page 84 (Table 16, ALLEGRO 2b/3, ritlecitinib, 10 mg): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	This text has been amended.
Section 5.2.1, page 93 Throughout the text	The Company encourages the EAG to consider updating the relevant figures to two decimal places.	Typographical error.	Not a matter of factual accuracy
Section 5.2.5.4, page 100 (Table 20, from SALT ≤ 10 to SALT score 21-49, transitions from month 21 to month 24, patients discontinue treatment and transition to BSC): [redacted]	To update the text as follows: [redacted]	Typographical error (wrong number).	The proportion in the model is actually 0.746% which rounds to 0.7% to rounding to one d.p. No change made.
Section 5.2.7, page 108: "[redacted] QALYs"	The Company notes that [redacted] QALYs is [redacted] QALYs to three decimal places. This therefore means that its [redacted] QALYs to two decimal places.	Inconsistent presentation of data.	The EAG is reporting the PSA results based on its re-run of the company's PSA which are reported as [redacted] QALYs in Table 24

			so this is not inconsistent. No change made.
Section 5.3.4.10, page 128: "HAD"	To update the text as follows: "HADS"	Typographical error (spelling).	Thanks. Corrected as requested.
Section 5.3.4.14, page 132 "The EAG was unable to corroborate the disutility of 0.07 applied for acne, folliculitis and urticaria which the Company claims was based on the disutility for severe redness reported by Stein et al. (2018)."	"The EAG was ~~unable~~ able to corroborate the disutility of 0.07 applied for acne, folliculitis and urticaria which the Company states were based on the disutility for severe redness reported by Stein et al. (2018)."	The value of 0.07 can be found in Table 6 in the Stein *et al* journal, referenced to the row 'Severe redness/skin peeling', column 'Coefficient'.	Table 6 provides regression coefficients not the disutility for those with this adverse event. The disutility for severe redness/skin peeling is reported in Table 7 of Stein et al. (2018) as 0.060.
Section 7.1, page 143: "At Week 24, [redacted] and [redacted] of de novo and rollover participants, respectively, had a SALT score of ≤20 at Week 24"	To update the text as follows: "At Week 24, [redacted] and [redacted] of de novo and rollover participants, respectively, had a SALT score of ~~≤20~~ ≤**10** at Week 24"	Typographical error (wrong number).	Thanks. Corrected as requested
Section 5.2.7, Page 109, Table 25: Incremental acquisition costs: £[redacted]	£[redacted]	Typographical error (wrong number).	Apologies, corrected as requested.

Issue 9 Confidentiality marking

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG response
Section 4.2.1.1, page 40: "Baseline characteristics are similar to the *de novo* cohort, with [redacted] female and [redacted]% white"	Baseline characteristics of the ALLEGRO-LT study have been marked as academic in confidence in the company submission and therefore should remain confidential.	"Baseline characteristics are similar to the de novo cohort, with [redacted] female and [redacted] white"	This AIC marking has been added.
Section 5.2.1, page 93: "[redacted]%"	Baseline characteristics in the economic model have been marked as academic in confidence in the company submission and therefore should remain confidential.	[redacted]	Apologies, corrected as requested.
Section 7.1, page 143: "[redacted]% and [redacted]%"	Results of the ALLEGRO 2b/3 study have been marked as academic in confidence in the company submission and therefore should remain confidential.	"[redacted] and [redacted]%"	Apologies, corrected as requested.

References

1. Pfizer data on file. UK interviews with dermatologists with a specialist interest in hair disorders for the use of ritlecitinib for patients with AA. 2022.
2. Harries M, Macbeth A, Holmes S, *et al.* The epidemiology of alopecia areata: a population-based cohort study in UK primary care. *British Journal of Dermatology* 2021.
3. Xu L, Liu KX & Senna MM. A practical approach to the diagnosis and management of hair loss in children and adolescents. *Front Med (Lausanne)* 2017. 4: 112.
4. Matza LS, Stewart KD, Lloyd AJ, *et al.* Vignette-Based Utilities: Usefulness, Limitations, and Methodological Recommendations. *Value in Health* 2021. 24: 812–821.

1. Harries MJ, Sun J, Paus R, King LE, Jr. Management of alopecia areata. *BMJ* 2010;341:c3671.
2. Xu L, Liu KX, Senna MM. A practical approach to the diagnosis and management of hair loss in children and adolescents. *Front Med (Lausanne)* 2017;4:112.
3. Pfizer. A Phase 2b/3 Randomized, Double-Blind, Placebo-Controlled, Dose-Ranging Study to Investigate the Efficacy and Safety of PF-06651600 in Adult and Adolescent Alopecia Areata (AA) Subjects With 50% or Greater Scalp Hair Loss; 2022.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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 **Monday 5 June 2023**. 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

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	Tommy Price
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Pfizer Ltd
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state the name of the company, amount, and purpose of funding.	
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry	

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

<table>
<tr><th>Key issue</th><th>Does this response contain new evidence, data or analyses?</th><th>Response</th></tr>
<tr><td>(1)The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup</td><td>Yes~~/No~~</td><td>The company has conducted a subgroup analysis for the alopecia totalis (AT)/ alopecia universalis (AU) subgroup with the appropriate inputs as suggested by the EAG. Three scenario analyses have been conducted investigating the cost-effectiveness of ritlecitinib compared to best supportive care (BSC) when used for patients:

<table>
<tr><th>Scenario</th><th>ICER</th></tr>
<tr><td>Base Case</td><td>£14,290</td></tr>
<tr><td>1) Who are classified as AT/AU prior to treatment,</td><td>£16,625</td></tr>
<tr><td>2) Who are not classified as AT/AU prior to treatment.</td><td>£13,304</td></tr>
<tr><td>3) According to the expected distribution of patients with AT/AU and non-AT/AU prior to starting treatment with ritlecitinib amongst those who are eligible for treatment with ritlecitinib. The ICERs in the AT/AU and non-</td><td>£13,620</td></tr>
</table>
</td></tr>
</table>

		AT/AU populations were weighted based on the proportion of patients with severe alopecia areata (AA) who are classified as AT/AU and non-AT/AU.

The proportion of patients with severe AA who have AT/AU used to calculate the weighted ICER was estimated to be 9.52%. This was estimated by dividing the percentage of patients with AT/AU as a subpopulation of the total population of patients with AA (4%) by the percentage of patients with AA who present with severe AA at first presentation (42%) as follows: $\frac{4\%}{42\%} = 9.52\%$.[1,2] Further details regarding these estimates can be found in Section B.1.3.1.3 of Document B.

This proportion was preferred to the proportion of patients in the ALLEGRO 2b/3 trial with AT/AU (46.0%) because the ALLEGRO 2b/3 trial was enriched with AT/AU patients.[3] The ALLEGRO 2b/3 study was enriched to ensure the trial included a sufficient sample of patients who are classified as AT/AU to be able to assess the effectiveness of ritlecitinib in this subgroup.

All scenarios remain cost-effective in the company's updated base case with ICERs of £16,625, £13,304 and £13,620 respectively.

The estimated ICER for patients with AT/AU prior to treatment is higher than the estimated ICER for the other two scenarios described because patients with AT/AU have a higher baseline SALT (95-100) compared to patients with severe AA without AT/AU (SALT 50-95), making it less likely for these patients to achieve a SALT score of ≤20 within 48 weeks of beginning treatment with ritlecitinib. However, the increase in the ICER is not substantial and treatment with ritlecitinib remains cost-effective. Moreover, given that patients with AT/AU represents a small subset of patients with severe AA, the ICER according to the expected distribution of patients with AT/AU and non-AT/AU prior to starting treatment with ritlecitinib is lower than the ICER in the base case.

(2)ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	~~Yes~~/No	The company disagrees that the ICER for the whole population should be based on a weighted average of outcomes for adults and adolescents. In response to clarification questions from the EAG, the company provided results from a longitudinal exposure response analysis of clinical efficacy data from the ALLEGRO 2b/3 study which concluded that age does not modify the treatment effect of ritlecitinib compared to BSC [Clarification question company response; A12]. The only potential treatment effect modifier identified was AA severity status, which demonstrates that there is no need for age weighted ICERs. Dermatologists with a specialist interest in hair disorders have confirmed (3/3 clinicians, 100%) that they would not expect to see a treatment effect difference between adults and adolescents.[4] Furthermore, the weighted ICER approach is more flawed than the original company approach. In the economic model provided by the company, the scenarios for adult and adolescent, adult only, and adolescent only, are informed by the following data from the ALLEGRO 2b/3 and ALLEGRO-LT trials, respectively: 1) Adult and adolescent (base case); clinical efficacy data is informed by all patients and demographic characteristics are in line with the whole cohort. 2) Adult only (scenario); clinical efficacy data is informed by adult data only and demographic characteristics are in line with the adult cohort only. 3) Adolescent only (scenario); clinical efficacy data is informed by both adult and adolescent data and demographic characteristics are in line with the adolescent cohort only. The adolescent only scenario (3) considers the clinical effectiveness of adults and adolescents due to the sample size of adolescents in the ALLEGRO 2b/3 study (14.6% of the full cohort) being too small to provide meaningful estimates of cost-effectiveness, reducing the accuracy in expected outcomes for this subgroup. In taking a weighted average approach of the adult and adolescent populations, further imbalances would be created within the data by over-engineering an already

		balanced cohort. This imbalance creates clinical and therefore economic uncertainty in comparison to using the population as a whole.
(3)Assumption of no treatment waning based on limited long-term evidence	Yes/~~No~~	The company's base case assumes that patients who have responded to treatment for ~ 2 years (96 weeks) will have a stable SALT score (stay in state) until they discontinue treatment. The ALLEGRO-LT trial is an ongoing trial that includes a number of cohorts rolled over from ALLEGRO 2b/3 alongside newly recruited (*de novo*) patients.[5] The evidence from the ALLEGRO-LT trial detailed below supports an assumption of no treatment waning [Document B 3.2.3.4]. **Additional follow-up data from ALLEGRO 2b/3 and ALLEGRO LT supports the assumption of no treatment waning** The EAG has suggested additional follow up data would provide further evidence to support this assumption. An updated (Interim) analysis providing additional follow-up data is included below.[6] The following patient cohorts have been selected to inform the updated analysis as they align with EAG comments on removing those patients who received 30mg with or without loading dose including matching placebo: 1. Patients who were treated with a 50 mg dose for 48 weeks in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial. 2. Patients who were treated with a 200 mg loading dose (4 weeks) followed by a 50 mg dose (20 weeks) in the ALLEGRO 2b/3 trial, followed by a 50 mg dose in the ALLEGRO-LT trial. 3. Patients who began on placebo (24 weeks) and transitioned to a 50 mg dose (24 weeks) in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial.

4. Patients who began on placebo (24 weeks) and transitioned to a 200 mg loading dose (4 weeks) followed by a 50 mg dose (20 weeks) in the ALLEGRO 2b/3 trial followed by a 50 mg dose in the ALLEGRO-LT trial.
5. Patients entering the ALLEGRO-LT trial *de novo* who were treated with a 200 mg loading dose (4 weeks) followed by a 50 mg dose of ritlecitinib and excluded those participants with known androgenetic alopecia and participants with a screening or baseline SALT score ≤50. This modified *de novo* group were considered to ensure alignment with the proposed population eligible to receive ritlecitinib.

Figure 1 and Figure (SALT≤20), and Figure 2 and Figure 3 (SALT≤10) below show the updated data across selected groups. The 200/50mg cohorts include groups 2,4,5 (above). The 50mg cohorts include groups 1 and 3. Data tables are provided as a reference (tables 51a 3.1.1 and 51a 3.2.1).[6]

Figure 1: ALLEGRO-LT: Response Based on SALT ≤ 20 up to Month 24 (Interim Analysis Selected cohorts, 200/50 mg dose)

Abbreviations: CI, confidence interval; QD, once daily; SALT, Severity of Alopecia Tool.

		Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤10. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints. Source: Pfizer data on file. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT) - Interim Analysis. 2023.[6] **Figure 2: ALLEGRO-LT: Response Based on SALT ≤ 20 up to Month 24 (Interim Analysis Selected cohorts, 50 mg dose).**  *Abbreviations: CI, confidence interval; QD, once daily; SALT, Severity of Alopecia Tool.* *Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤10. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints. Source: Pfizer data on file. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT) - Interim Analysis. 2023.[6]* **Figure 2: ALLEGRO-LT: Response Based on SALT ≤ 10 up to Month 24 (Interim Analysis Selected cohorts, 200/50 mg dose).**

		[redacted] *Abbreviations: CI, confidence interval; QD, once daily; SALT, Severity of Alopecia Tool.* *Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤10. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.* *Source: Pfizer data on file. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT) - Interim Analysis. 2023.*[6] **Figure 3: ALLEGRO-LT: Response Based on SALT ≤ 10 up to Month 24 (Interim Analysis Selected cohorts, 50 mg dose).** [redacted]

		Abbreviations: CI, confidence interval; QD, once daily; SALT, Severity of Alopecia Tool. *Note: n/N indicated for each timepoint: N = number of patients with observed data, n = number of patients achieving SALT ≤10. The ALLEGRO-LT trial is ongoing; therefore, a lower number of patients appear at later timepoints.* *Source: Pfizer data on file. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT) - Interim Analysis. 2023.*[6] Of those still participating in the study at each time point, all but one patient had follow-up data at 24 months in the 200/50mg cohort and all patients in the 50mg cohort had follow up data at 24 months. Therefore, all but one patient provided data at 24 months across all combined cohorts. The data supports stabilisation of the proportion of patients with SALT≤20 and SALT≤10 after two years of continuous treatment regardless of whether a loading dose of 200 mg was administered during the first four weeks of treatment. **Feedback from clinical experts support no waning** Dermatologists with a specialist interest in hair disorders (3/3 clinicians, 100%) support the assumption of no treatment waning in their clinical experience. That is, if patients have responded for ~2 years there is an expectation that they would continue to respond moving forward until they discontinue. One dermatologist referenced their own ritlecitinib trial patient experience. One dermatologist referenced the long-term data from baricitinib which also supports the assumption of no treatment waning in severe AA.[7] **Uncensored analyses provides additional evidence of no waning** The uncensored analysis of discontinuation which is provided as a scenario analysis in the economic model is consistent with the assumption of no treatment waning. Patients may discontinue ritlecitinib treatment for any reason from health states. This is derived by extrapolating time on treatment amongst patients with a SALT score ≤20 after 48 weeks of treatment. A high time on treatment, as shown by the

		uncensored analysis, is consistent with a lack of waning effect of treatment i.e., that patients who respond tend to persist with treatment. Taken together, this data further supports the assumption of no treatment waning and there is no apparent alternative data to challenge this assumption. We therefore believe that the assumption of stay in state where patients who continue to respond at 2 years remain in their current SALT state until they discontinue is justified.
(4)Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	Yes/~~No~~	In the original company submission, all patients who had rolled over into the ALLEGRO-LT study from ALLEGRO 2b/3 were included to ensure all relevant data was considered. However, the EAG stated a preference in the EAG report of using data only from patients who have previously received a 50mg dose of ritlecitinib to estimate the long-term transition matrices. The company agrees with this recommended change and has updated the base case accordingly. In doing so, data from patients who had switched from a 30mg dose in the ALLEGRO 2b/3 study to a 50mg dose at the start of the ALLEGRO-LT study are excluded from the long-term extrapolation. Previously, as transition matrices included data from 48 weeks of ritlecitinib treatment, data for patients who switched from a 30mg dose to a 50mg dose started being collected at the point at which dose changed. Therefore, there may have been an increase in response for these patients which would not reflect continued treatment with the 50mg dose. Data for patients who had received a 200mg loading dose in the ALLEGRO 2b/3 study before switching to the 50mg maintenance dose are still included as the patients were deemed sufficiently similar to those who had a 50mg dose from the start of treatment. Moreover, the change from 200mg to 50mg occurred after 4 weeks of treatment with ritlecitinib, so patients would have been on a stable dose of 50mg for 44 weeks before the inclusion of their clinical data in the extrapolation transition matrices. Therefore, it is not expected that a ‘jump’ in efficacy could occur at the point of Week 48 for these patients.

(5)Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	Yes/~~No~~	In severe AA, spontaneous remission is thought to be rare. While the number of patients who can go on to experience spontaneous remission varies, it is not durable and the proportion of patients with severe AA who maintain spontaneous remission is very low.[8,9] This is further validated by the EAG (EAG report; Section 5.3.4.6) in the following statement: *"The EAG's clinical experts advised that spontaneous remission was extremely unlikely in their experience in patients with severe AA and became less likely beyond the first 6 months to a year"*. This is consistent with company advice from dermatologists with a specialist interest in hair disorders. In the original company submission, the percentage of patients on BSC assumed to achieve spontaneous remission from Week 24 and assumed to be constant from then on was 1.54% (equal to the percentage of adult and adolescent patients in the placebo arm of the ALLEGRO 2b/3 trial who had a SALT score ≤10 at Week 24). This is a conservative assumption, considering the rate of spontaneous remission is rare (as validated by clinical experts consulted both by the company and by the EAG), and spontaneous remission would more appropriately be measured as a SALT score of zero which is inherently rarer than a SALT score of 0-10. In the EAG's base case, spontaneous remission is applied to patients on BSC but not to patients who have discontinued treatment with ritlecitinib because the EAG argue that spontaneous remission *"would have been incorporated within the group of patients deemed to have responded to ritlecitinib and therefore would have already been captured within the modelled outcomes for 48 weeks and the status of these patients as continued responders in the ALLEGRO-LT study"*. However, there is no evidence to support that treatment with ritlecitinib will impact the likelihood of a patient achieving spontaneous remission over time. If a patient does respond to treatment and then discontinues treatment, they may well go on to experience spontaneous remission in the future. This was supported by clinical advice from dermatologists with a specialist interest in hair disorders. This assumption was

		accounted for in the original company submission, that a constant proportion of patients have spontaneous remission to account for patients gaining and losing spontaneous remission over time. Therefore, exclusion of the assumption that patients previously treated with ritlecitinib could not experience spontaneous remission is biased against ritlecitinib. Furthermore, if as the EAG prefer, a proportion of ritlecitinib responders are always in spontaneous remission then this contradicts the points raised on expected duration of treatment. The original assumption to apply spontaneous remission to discontinuers of ritlecitinib was justified with clinicians who agreed it was relevant to include for patients treated with ritlecitinib.[10] This is because some patients will discontinue treatment with ritlecitinib following response to treatment, which may be indefinitely sustained in an unknown number of patients. Therefore, the EAG's approach to spontaneous remission is not appropriate. Additionally, given the EAG's clinical experts advised that spontaneous remission is extremely unlikely in patients who would be eligible for treatment with ritlecitinib, this indicates that the assumption that the rate of spontaneous remission is equal to the percentage of patients in the placebo arm of the ALLEGRO 2b/3 trial who had a SALT score ≤10 at Week 24 is also not appropriate. Discussion with the EAG at the technical engagement meeting led to a conclusion that removing spontaneous remission would be the most appropriate approach given the points raised above due to simplicity and the limited impact of the assumption. Therefore, the company has updated the base case of their economic model to remove spontaneous remission for all patients. In the company's updated base case, after Week 24, all BSC patients are assumed to return to the SALT 50-100 health state and no patients who discontinue treatment with ritlecitinib remain in the SALT 0-10 health state indefinitely. BSC patients are assumed to have a loss of response in the same way as patients who discontinue treatment with ritlecitinib, such that they remain in the same health state for one cycle

		after the final data point for BSC patients in the ALLEGRO 2b/3 trial (Week 24 to Week 36), before transitioning sequentially through the health states with a greater SALT score each cycle until reaching 'BSC SALT ≥50'. In the updated base case, the option previously applied in the economic model to revert patients on BSC with SALT 21-49 to SALT 50-100 is switched off as these patients will return to SALT 50-100 through the assumed loss of response. The exclusion of spontaneous remission has a minimal impact on the new base case ICER (£14,290), which is increased to £14,297 when spontaneous remission is applied to patients on BSC and patients who discontinue treatment with ritlecitinib.
(6)Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	Yes~~/No~~	Given the EAG's concerns about the discontinuation analysis for reasons other than a loss of response, the company has updated the discontinuation analysis to censor patients at the time they have a SALT score ≥20 (at which point their discontinuation of treatment would be triggered by the stopping rule) rather than excluding patients who stop responding at any time point from the entire analysis. This is directly in line with the EAG's preferred approach for the discontinuation analysis. The method used to extrapolate the Kaplan Meier (KM) data was based on the guidelines presented in the NICE Decision Support Unit technical support document 14.[11] Dermatologists with a specialist interest in hair disorders validated that it is appropriate to extrapolate data from the ALLEGRO-LT trial to estimate the long-term rate of treatment discontinuation (3/3 clinicians, 100%).[10] The plot of hazard of discontinuation over time presented in Figure 4 was used to assess whether a proportional hazards model or an accelerated failure time (AFT) model should be used. As shown in Figure 4, the hazards over time are not constant, indicating that an AFT model should be used, such as the Weibull, log-logistic, log-normal or generalised gamma models. The change in hazards at 1.4 years is thought to be driven by a reduction in patient numbers. Of the [redacted] patients included in the discontinuation analysis, only [redacted] patients remain at risk of discontinuation after 1.4 years of ritlecitinib treatment and, amongst these patients, [redacted] discontinued.

		Conversely, amongst the [redacted] patients whose exposure to ritlecitinib was less than 1.4 years, there were [redacted] discontinuation events. Therefore, once further follow-up is available it is expected that the hazards would remain constant. This supports an expected plateau in the hazards of discontinuation following any initial discontinuation due to AEs/tolerability. A high retention to treatment with ritlecitinib amongst patients who are responding to treatment was validated by dermatologists with a specialist interest in hair disorders based on their clinical experience.[4] It was suggested this could be because patients with AA are typically physically well, increasing the likelihood of patients remaining on treatment with JAK inhibitors for longer compared to patients treated with JAK inhibitors for other conditions. Given this, a constant hazards model may be appropriate despite what the data indicates. Opting for an AFT offers a more conservative approach on treatment duration for the base case. **Figure 4: Hazards plot of the ALLEGRO-LT discontinuation data** 

The Weibull curve has been selected for the updated base case given that, of the AFT models, it has the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) scores as shown in Table 1. According to the Weibull curve shown in Figure 5, approximately [redacted]% of patients would be expected to remain on treatment with ritlecitinib four years after achieving SALT score ≤20 following 48 weeks of ritlecitinib treatment. The visual fit of the parametric distributions compared to the KM is displayed in Figure 5 and Figure 6. The sensitivity of the cost-effectiveness of ritlecitinib to the choice of parametric curve is presented in Table 4.

Table 1: AIC and BIC statistics for parametric distributions fit to ALLEGRO-LT discontinuation

Distribution	AIC	BIC
Exponential	[redacted]	[redacted]
Weibull		
Gompertz		
Log-logistic		
Lognormal		
Generalised Gamma		

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion

Figure 5: Updated parametric distributions fit to ALLEGRO-LT discontinuation

[redacted]

Figure 6: Updated parametric distributions fit to ALLEGRO-LT discontinuation (two-year time horizon to assess visual fit to the KM data)

The original discontinuation analysis, presented in Figure 7, was estimated by excluding patients with a SALT score >20 at any time point from the entire analysis. In the base case, the exponential curve was chosen based on statistical fit which resulted in a slightly more pessimistic estimate of [redacted]% of patients expected to remain on treatment with ritlecitinib four years after achieving SALT score ≤20 following 48

		weeks of ritlecitinib treatment. When applying the old base case for the survival analysis in the economic model according to the new base case, the ICER of ritlecitinib compared to BSC was £14,420 which is slightly higher than the current base case (£14,290), demonstrating the minimal impact of the update to the discontinuation analysis. **Figure 7: Parametric distributions fit to ALLEGRO-LT discontinuation prior to technical engagement (Figure 29 of Document B)**  The company has also rerun the scenario analysis for the scenario in which the final stopping rule is SALT score ≤20 at Week 48 but patients are not required to discontinue treatment if their SALT score is >20 after Week 48 as an error was made filtering the data for this scenario in the original analysis. The duration of follow-up available in the corrected scenario is equal to the other scenarios ( ).

		The company has also corrected the time axis of the KM curve (Figure 5 and Figure 6) such that it is aligned with the survival data.
(7)Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	Yes/No	Utility measures in AA must capture the full burden of the disease to enable effective decision-making on the cost-effectiveness of potential new treatments. The company recognises the EQ-5D as the NICE reference case and the key elements: patient responses with societal tariffs applied from published value sets. The company approach has been to align in principle to this by reflecting the patient experience of AA with a link to societal preferences. However, whilst EQ-5D can capture the depression and mental health of patients, other concepts of HRQoL that are important to patients with AA are missing. These gaps were in the domains of social functioning, relationships, emotional, physical, appearance and financial.[12] This is acknowledged by the EAG and in the recent draft guidance for Baricitinib in severe AA.[13] Whilst the EQ-5D might be adequate in other diseases, the lack of patient centricity with the lived burden of AA as clearly articulated by patients and dermatologists leads to uncaptured value. This is potentially due to the complex nature of the disease, its heterogeneity and psychosocial impact (Doc B 1.3.2, 3.4.1 & 3.4.4). Therefore, cost-effectiveness analysis in AA that utilise QALY estimates based on EQ-5D sources likely undervalue any meaningful treatment benefits to HRQoL. This risks inequitable and inefficient decision-making for medicines that clearly support an important unmet need in an underserved population of patients with AA. The following sub-headings outline in more detail a number of potential reasons why EQ-5D based utilities lack face validity. Firstly, how literature-based methods fail to align with the severity of the condition and risk introducing bias. We also cover in more detail the case for EQ-5D lacking sensitivity and content validity in AA; this

		aligns with our own qualitative and quantitative research, which is supported by patient advisory groups (PAG) and the British Association of Dermatologists (BAD). Finally, the company believes that patient-centred methods that capture societal preferences (i.e., the submitted vignette study) is the most appropriate approach to obtaining health state utilities in AA using recommended methods that reduce bias. The vignette-based approach was directly informed by empiric qualitative and quantitative evidence to ensure that utility estimates accurately reflect the full burden and voice of patients living with AA. The vignettes were developed through a targeted literature review, AA-specific clinical trial analyses of validated PROs, and consultation with patients and family members. The direct TTO valuation of the vignette-based health states ensure that societal preferences were incorporated per NICE guidance. This is analogous to the EQ-5D reference case (that is, collect patient responses from trials then translate those to societal preferences from published EQ-5D value sets). **EQ-5D Utilities from the trial lack face validity and should not be included in the economic analysis.** The EAG "*acknowledges that there may be some underestimation of QALY gains when using EQ-5D based utility estimates obtained directly from ALLEGRO-2b/3*". (Page 13 EAG report) This is in line with the company submission which outlines a number of potential reasons including selected population recruited and limited follow up data, as suggested by the EAG, but also that the EQ-5D fails to capture all HRQoL domains that are relevant to patients with AA. Therefore, the company does not agree with the EAG suggestion to include EQ-5D based utilities from the trial as the company base case when both company and EAG agree (regardless of reason) that it

		underestimates QALY gains. The challenges of capturing AA treatment-related HRQoL changes in large clinical trials has been acknowledged in recent draft guidance for Baricitinib for the treatment of severe AA.[13] The committee concluded that there is uncertainty and may result in uncaptured value of treatment. The company maintains that the ALLEGRO 2b/3 EQ-5D results lack face validity and should not be used in the economic analysis as it will lead to underestimation of the HRQoL benefits of treatment to patients with AA. **Published EQ-5D measures of severity are not aligned to SALT scores and are at risk of bias** There is a potential disconnect between clinician and patient assessment of severity. Therefore, patient centric development of utility estimates from multiple sources to develop vignettes more accurately captures the burden of the condition. The EAG reference two papers as evidence for EQ-5D sensitivity in the AA literature.[14,15] Whilst this evidence might be perceived as having more face validity, this is based on a subjective assessment of AA severity (mild, moderate, severe) and is subject to bias. Neither study included SALT scores to grade patients, instead grading severity was based on clinician judgment and may not correlate exactly with patient ratings of severity, and their perception of hair loss. Studies suggest a moderate to low patient-physician alignment.[14,16] This may be due to the visible nature of AA, particularly where the impact is on the head and face. This may result in physicians making quick assessments of the impact and extent of hair loss but fails to capture the severity as rated by patients.[14] This might include the impact of hair loss on patients' self-image leading to a distorted view of how bad the hair loss was or even an inability to describe the impact which is an associated

		condition in patients with AA (alexithymia).[17] It is possible that this subjective assessment from different clinicians encompassing more than hair loss alone, produced bias. With no formalised structure in place, we cannot ascertain how other factors could have informed their grading. **The company has provided evidence that EQ-5D lacks sensitivity and content validity in AA (Doc B 3.4.1 and EQ-5D Manuscript)**[18] The EQ-5D only includes five questions covering five dimensions of health and there has been an ongoing discussion regarding whether the EQ-5D misses important areas of health for certain diseases.[19] The EuroQol Group have supported a research programme to develop and test "bolt-on" dimensions in disease areas including respiratory disease,[20] psoriasis,[21,22] as well as vision, hearing and tiredness.[23] If the five dimensions of the EQ-5D do not capture the impact or burden of a disease, then it is likely that cost-effectiveness analyses of treatments in these diseases will underestimate the true value of a treatment. Furthermore, if the trial-based EQ-5D utilities underestimate the impact on HRQoL then it is inconsistent to suggest that EQ-5D from the literature resolves the issue of whether EQ-5D captures all of the burden of the condition. Particularly, as the EAG "… *accepts that the company has provided some evidence supporting a lack of content validity for the EQ-5D in this population…"* which includes both qualitative and quantitative data alongside PAG and dermatologists with a specialist interest in hair disorders feedback. There is a reasonable perspective, whereby the EAG consider *"that other sources of data from the literature would support the construct validity of the EQ-5D in differentiating between patients with different severities of AA".* (Page 120, Table 26 EAG report). However, the literature does not support the assumption that the EQ-5D

		is capturing all of the burden of AA and we advise caution when comparing with other skin conditions which may not be appropriate. The EAG already state in their report that comparisons of utility estimates from other skin conditions are inappropriate so we feel these statements potentially contradict one another (Page 129 EAG report). The EAG reference a review from 2015 but we do not agree this covers a range of skin conditions but rather 12/16 being in patients with Psoriasis (Page 132 EAG report). The appropriateness of the EQ-5D in other dermatological conditions has also been questioned previously because important HRQoL concepts are missing.[22,24] For example, NICE recently commented that "the EQ-5D often fails to capture quality-of-life improvements for people with skin conditions".[24] More recently, guidance on Baricitinib for the treatment of severe AA also found that EQ-5D-5L may not be capturing important aspects of the condition and the impact on HRQoL is uncertain.[13] Empirical evidence also suggests the EQ-5D may be insensitive to the full impact of AA.[15, 18,25] Whether EQ-5D from the literature captures all of the burden of AA is a subjective assessment. We argue the observed (narrow) range and varied values across severity in the literature show a failure to capture all of the burden and therefore lacks sensitivity and validity in AA. This is supported by PAGs and dermatologists with a specialist interest in hair disorders and is covered in more detail in the company submission. This includes a systematic literature review, targeted literature review, quantitative research along with qualitive research with patient advocacy groups, and dermatologists with a specialist interest in hair disorders (Doc B 3.4). Key elements of HRQoL for patients with AA are omitted from the EQ-5D. These gaps were in the domains of social functioning, relationships, emotional, physical, appearance and financial (Figure 44) Doc B 3.4.4). In the omission of these elements of HRQoL, which are important to patients with AA, the EQ-5D lacks content validity.

		The EAG report acknowledges the difficulties in aligning these broad domains to the relevant domains associated with AA *….."Whilst the purpose of the EQ-5D is to cover the broad domains of health that are common across many conditions, the EAG recognises that patients may not see how these domains are relevant in the context of their specific condition and this is reflected in the company's qualitative research with patient advocacy group representatives (n=9)".* (Page 130 EAG report). [redacted] Therefore, in patients treated for AA, some of the gain in HRQoL is not reflected in the QALY estimates and the magnitude is unknown. EQ-5D captures some, but fails to capture all, of the important and relevant elements of HRQoL for patients with AA. If we follow this logic through, any source (literature or PRO mapping i.e., DLQI) that uses EQ-5D potentially underestimates the HRQoL impact, thereby undervaluing treatments for AA. NICE recommend that alternative methods for utility generation are used in such circumstances where the EQ-5D has been shown to be inappropriate.[26] One method outlined is the time-trade off (TTO) approach, whereby utilities are elicited from vignette descriptions of health states. Alternative approaches to EQ-5D are not unusual, a substantial proportion of cost-utility analyses submitted to NICE use utilities that deviate from the recommended EQ-5D approach, including utilities based on vignettes.[18,27,28]

		In conclusion, the company feels it has provided enough evidence to deviate from the reference case and include as its base case an alternative approach in line with NICE guidance.[29] The company also believe the EAG has not provided enough evidence to support the claim that EQ-5D fully captures the HRQoL impact for patients with AA. The company believe the vignette approach, which includes capturing societal preference through the TTO, attempts to identify the missing gain in QoL and capture the full burden of AA into the utility estimate. The vignettes were based on; 1. Data from standardised, validated PRO measures (AAPPO, HADs) and 2. Data from in depth qualitative research with patients, family members and clinicians. Therefore, we suggest that the vignettes do reflect the full burden of AA for patients. **Vignette Study Methodology follows best practice and any deviation risks introducing bias.** *"The EAG has concerns regarding the face validity of the final vignettes in comparison to the quantitative data used to develop them and therefore believes they should be treated with caution*" (Page 13 EAG report) The vignettes were developed from multiple credible sources of information. The company has followed the methodology as outlined by Matza et al,[30] which the EAG has acknowledged as reputable guidance. Any deviations from the methods used risks introducing bias. All of the suggested good practices outlined by Matza et al. were followed including patient and clinician interviews, conceptual model development and validation work, to ensure the research stayed "true" to the patient voice, covering concepts that are relevant to an AA patients lived experience.[31] **<u>Additional points raised by the EAG</u>**

		"the vignettes may not be applicable to patients who have similar SALT scores but are currently managing their alopecia without receiving or seeking to receive systemic treatments" (Page 129 EAG Report). Our focus is (and should be) on those AA patients who are interested in receiving treatment. The company argue that the EAG suggested approach is inconsistent with the population we are trying to ascertain utility values for. *The EAG believes that the vignettes lack face validity when compared with the quantitative data the company has used to inform them. For example, the vignette for SALT>50 (A3) states, "you frequently feel self-conscious or embarrassed about your hair loss," whereas the data from the ALLEGRO 2b/3 trial (Table 8, Appendix H) suggest that the most common response to, "over the past week, how often did you feel embarrassed about your hair loss?" in patients with a SALT score of 50-100 was never (34%), followed by rarely (25%). In fact, only 21%, 12% and 9% responded sometimes, often, and always respectively.* This suggests that one source of data has precedence over multiple sources. This risks introducing bias into the final vignettes. Multiple sources of information should inform the vignette development if best practice is followed, as recommended by Matza et al.[30] Therefore, the final vignettes would be unlikely to align with any one source of information given the best practice steps that are followed. Rather, its aim is to reflect the lived burden of the patient with AA across multiple sources to reduce potential bias. To further support this we describe below how the vignettes were developed and the reasoning behind the assimilation of those multiple sources to arrive at the final vignettes.[31] The clinical trial data were used as the primary basis for developing the draft descriptions of impacts on self-consciousness and embarrassment. While 'never'

		was among the most frequently reported response options for both items highlighted by the EAG, there was no clear modal value that could be considered representative of the typical patient experience. Therefore, we considered the distribution of responses to determine an 'average'. In this case, 'sometimes' was selected for each item to represent the distribution of responses in the draft vignettes. Data from Pfizer's preference study (B7981048, B7981072 analysed in the vignette study)[31,32] were also considered (data provided in report appendix) when drafting the description. Responses to the AAPPO in the preference study were slightly more severe than the clinical trial data, with [redacted] 'sometimes' feeling self-conscious and [redacted] 'often' feeling self-conscious. A similar pattern was observed for the embarrassment item ([redacted] 'sometimes' feeling embarrassed, [redacted] often feeling embarrassed). The draft vignettes were refined based on exploratory interviews with patients, debrief interviews with patients and HCPs and evidence from the literature. All patients and HCPs agreed that the impact on self-consciousness and embarrassment was understated. In debrief interviews, some patients described 'regularly' or 'always' feeling self-conscious. HCPs also suggested revising, with suggestions ranging from 'often' to 'constantly' feeling self-conscious or embarrassed. Exploratory interviews also supported frequent impacts, with patients stating that they felt self-conscious in public spaces and when meeting new people. Based on the evidence described above, the draft descriptions were revised, giving additional weight to the evidence gathered from the interviews. It should also be considered that the clinical trial excluded people with psychiatric conditions (e.g., suicidal ideation or behaviour in the past year, or clinically significant depression), so these data may underestimate the impact of AA. *Similarly, [redacted] of patients with a SALT score of 50 to 100 responded 'not at all' when asked, "over the past week, how much did you limit your interactions with others*

		because of your hair loss?" This contrasts with a description of, "you limit your social interactions with family, friends and colleagues because of your hair loss a lot and you often find it difficult to meet new people," for the relevant vignette (A3). The draft vignette for SALT>50 used the wording, "you limit your social interactions because of your hair loss quite a bit," which seems less definite than the wording in the final vignette. The draft description of the social impacts was primarily based on the clinical trial data and the distribution of responses for the social interactions item. However, evidence from the literature and the interviews suggested that the impact may be more severe, and this was considered when revising the vignettes. In exploratory interviews, all patients described avoiding social situations. There was also consensus in debrief interviews between patients and HCPs that the qualifier 'quite a bit' understated the impact on social activities. Suggested rephrasing for the SALT score ≥50 vignette included 'frequently', 'very frequently' 'significantly', 'a lot' and 'quite a lot'. The patients interviewed also described a wide-ranging impact on social activities and relationships, including those with friends, family, partners and colleagues. Difficulties with meeting new people, dating/intimate relationships were also frequently mentioned across interviews. This seems to be consistent with the existing literature, with one study describing that a third of patients experience a relationship ending because of AA.[33] Based on the evidence described above and recommendations from patients and KOLs, the frequency of the social impact was increased in the final vignette, and context was added to highlight the full impact in the description.
(8)Carer disutility based on a vignette for a carer of an adolescent with severe	Yes~~/No~~	The company has accepted the EAG's suggestion and applied caregiver disutility to carers of adolescents only. However, feedback we have obtained from PAGs and dermatologists with a specialist interest in hair disorders, support the evidence presented in the company

alopecia areata has been applied at all ages		submission, that the disutility applied to carers and caregivers are relevant to both adults and dolescents.[10, 12,34] For instance, one cohort study of 229 family members of patients with AA found that 69.9% of family members of adults with AA (aged 17+) experienced some HRQoL impairment and 14% experienced a very large or extremely large effect (measured using the Family Dermatology Life Quality Index [FDLQI]).[34] However, the burden is greatest and most widespread for caregivers of adolescents and this is supported by feedback from dermatologists with a specialist interest in hair disorders.[34] The company also accept the comment by the EAG regarding the vignette methodology that did not specifically account for caregivers of adults who suffer from severe AA. More research is required on the specific QoL impact on caregivers of adults with severe AA. Therefore, we agree to only apply the carer or caregiver disutility to the adolescent population. In addition, rather than assuming that all caregiver disutility resolves when an adolescent's SALT score reaches below SALT50, we agree with the EAG suggestion to provide an estimate for the mild to moderate health states. There is evidence that the psychosocial strain on caregivers increases with severity of disease, as described in Section B.1.3.2.4 of Document B.[35] In a prospective study conducted in the US of 153 paediatric patients with AA, significant mild-to-moderate negative correlations were found between SALT scores and both FDLQI and Quality of Life in a Child's Chronic Disease Questionnaire (QLCCDQ) scores.[35] Dermatologists with a specialist interest in hair disorders (3/3 clinicians, 100%) have also confirmed it is reasonable to expect the disutility of a carer or caregiver of adolescents to also apply to the mild to moderate health states. One dermatologist noted that they would not apply a disutility to the SALT≤10 health states. We discussed several options with the EAG during the technical engagement meeting and we have presented these as scenarios. We believe a linear approach has less potential for bias and therefore present these as scenario A & B. We have also provided a third scenario C, using a relative change to the absolute utility

estimates obtained from the vignette study. The three scenarios lead to only small increases in the ICER with estimates of £14,293, £14,310 and £14,296, respectively, compared to the updated base case ICER of £14,290.

Scenario	ICER (£/QALY)
Base Case	**14,290**
A) Assuming a linear relationship in caregiver disutility across health states, extending the disutility to the SALT 21-49 health state by applying 50% of the caregiver disutility applied to the SALT 50-100 health state to the SALT 21-49 health state.	14,293
B) Assuming a linear relationship in caregiver disutility across health states, extending the disutility to all health states by applying: i. 75% of the caregiver disutility applied to the SALT 50-100 health state to the SALT 21-49 health state, ii. 50% of the caregiver disutility applied to the SALT 50-100 health state to the SALT 11-20 health state, and iii. 25% of the caregiver disutility applied to the SALT 50-100 health state to the SALT 0-10 health state.	14,310
C) Anchoring the magnitude of caregiver disutility in each health state proportionally to the decrement	14,347

		in utility for each health state compared to the utility of patients with a SALT score of 0-10.

The third scenario (C) has been estimated by calculating the proportional decrements to health state utility for patients with AA compared to the SALT 0-10 health state. The proportion of the maximum disutility for each health state was then applied to the disutility of caregivers of patients with a SALT score of 50-100 to estimate the caregiver disutility for each health state. The resultant caregiver disutilities are shown in Table 2.

Table 2: Caregiver disutility applied in scenario analysis

State	SALT 50-100	SALT 21-49	SALT 11-20	SALT 0-10
SALT 0-10 health state utility at baseline: 0.9190				
Disutility of patients with AA compared to the SALT 0-10 health state	[redacted]	[redacted]	[redacted]	[redacted]
Proportion of maximum disutility	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutility	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SALT, Severity of Alopecia Tool

(9)Utility values are not age-adjusted	Yes/~~No~~	In the original company submission, age-adjusted utilities were not included as the relationship between age and utility amongst patients with AA is not clear. However, to present a conservative analysis, the company has accepted the EAG's suggestion and applied age-adjusted utilities in the economic model. The health state utilities

		have been adjusted using average age-dependent utilities for the general population of the UK, calculated using Ara and Brazier (2010).[36] A multiplicative approach has been taken such that the health state utility is multiplied by the relative value of the age-specific utility to the age-specific utility at the baseline age in the model. The inclusion of age-adjusted utilities has a minimal impact on the new base case ICER (£14,290), which is reduced to £13,820 when age-adjusted utilities are excluded, as shown in the Sensitivity analyses around revised base case.

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: Missing patients in the clinical evidence when implementing the interim stopping rule	Section 5.3.4.3, p.122-123 Section 5.4.2.13, p.141	~~Yes/~~No	The company has accepted the EAG's correction of carrying forwards the [redacted] patients who were missing due to non-COVID-19 reasons at Week 24, assumed to be steady/improved, to contribute to the SALT >50 health state in Week 36 and Week 48. In the original company submission, these patients were excluded from contributing to the response rates after Week 24. The EAG's correction has resolved this issue and the company have applied it to their updated base case.
Additional issue 2: The EAG's probabilistic ICER was not close to the EAG's deterministic ICER.	Section 1.7, p.15	~~Yes/~~No	The company has rerun the probabilistic sensitivity analysis in the updated base case which does not include spontaneous remission and the mean probabilistic ICER (£14,482) is close to the deterministic ICER (£14,290) indicating robustness of the model. The driver of the issue which caused a large difference between the EAG's probabilistic ICER (£89,888) and deterministic ICER (£66,674) was the application of spontaneous remission to patients on BSC but not to patients treated with ritlecitinib.

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)
Key issue 4: Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	Transition probabilities for patients on ritlecitinib from 48 weeks were estimated from all patients in the ALLEGRO-LT trial who would pass the stopping rule.	The company has accepted the EAG's suggestion and used estimated transition probabilities only from patients on the 50 mg dose (regardless of whether there was a loading dose) to inform the matrices applied from 48 weeks in its base case.	Company base case ICER: £13,179 Updated ICER: £13,294
Key issue 5: Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	The rate of spontaneous remission was assumed equal for BSC patients and patients who discontinued treatment with ritlecitinib.	Spontaneous remission is no longer applied. Patients on BSC are now assumed to regress to the SALT 50-100 health state in the same way as patients who have discontinued treatment with	Company base case ICER: £13,179 Updated ICER: £13,294

		ritlecitinib, such that patients remain in the same health state for one cycle after the final data point (Week 24 to Week 36), after which point they gradually lose any prior improvement in SALT score by transitioning sequentially through the health states with a greater SALT score until reaching 'BSC SALT ≥50'.	
Key issue 6: Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	Time to discontinuation was calculated by excluding all patients with a SALT >20 after Week 48 to avoid double counting and the exponential curve is applied.	Patients with SALT >20 after Week 48 are included in the discontinuation analysis and then censored at the time point when their SALT score is first greater than 20 and the Weibull curve is applied.	Company base case ICER: £13,179 Updated ICER: £13,023
Key issue 8: Carer disutility based on a vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages	Utility value for carers was applied to carers of adolescents and adults.	Utility values for carers are applied to carers of adolescents only.	Company base case ICER: £13,179 Updated ICER: £14,191
Key issue 9: age adjusted utilities	Utilities were not adjusted for age.	Utilities are adjusted for age.	Company base case ICER: £13,179 Updated ICER: £13,348

Patients missing from numbers used to calculate baseline characteristics	[redacted] male patients who were ≥18 years were reported from the ALLEGRO 2b/3 trial.	The company has accepted the EAG's correction and amended the number of male patients who are ≥18 to [redacted] from the ALLEGRO 2b/3 trial.	Company base case ICER: £13,179 Updated ICER: £13,179
All-cause mortality in first 48 weeks	All-cause mortality was not included in the first 48 weeks.	The company has accepted the EAG's base case assumption that all-cause mortality is included in the first 48 weeks.	Company base case ICER: £13,179 Updated ICER: £13,140
Resource use for psychological support across treatments	The number of appointments for psychological support varied across treatments.	The company has accepted the EAG's base case assumption that the resource use for psychological support is equal across treatments.	Company base case ICER: £13,179 Updated ICER: £13,170
Resource use for adverse events	TEAEs included in the model were assumed to result in hospital admission.	The company has accepted the EAG's base case assumption that TEAEs are assumed to result in GP appointment.	Company base case ICER: £13,179 Updated ICER: £12,976
Additional issue 1: Missing patient assumptions	When the interim stopping rule was applied, patients who were missing due to reasons other than COVID-19 did not contribute to the response rates thereafter.	The company has accepted the EAG's correction and patients who were missing due to reasons other than COVID-19 contribute to the response rates thereafter.	Company base case ICER: £13,179 Updated ICER: £13,334

Capping the maximum of the caregiver disutility by zero	Maximum disutility was not capped by zero.	The company has accepted the EAG's correction and capped the caregiver disutility by zero.	No impact on base case
Inclusion of the uncertainty in rate ratio for AEs associated with BSC in the PSA	Per cycle probability for AEs associated with BSC when running the PSA was not calculated using the probabilistic value.	The company has accepted the EAG's correction and applied the probabilistic value for rate ratio in calculating the per cycle probability for AEs associated with BSC when running the PSA.	No impact on base case
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [redacted]	Incremental costs: £[redacted]	Company base case ICER: £13,179 Updated ICER: £14,290

Sensitivity analyses around revised base case

Probabilistic sensitivity analyses results for the company revised base-case for ritlecitinib versus BSC are presented in Table 3. Ritlecitinib is associated with 1.297 additional QALYs and £18,742 additional costs, with a corresponding ICER of £14,450 per QALY gained. The results demonstrate that the analysis is robust to parameter uncertainty with the probabilistic results lying close to the deterministic result (Table 4).

Table 3: Probabilistic results: ritlecitinib vs. BSC

Treatment	Total costs (£)	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£)
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	14,450

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; Inc. – incremental; QALY, quality-adjusted life year

As illustrated in the ICEP ([redacted]), ritlecitinib was more costly and more effective than BSC in the majority of iterations.

[redacted]

The CEAC is displayed in [redacted] to illustrate the probability of ritlecitinib being cost-effective compared to BSC, at various willingness to pay thresholds. At willingness to pay thresholds above [redacted], ritlecitinib is likely to be more cost-effective than BSC.

[redacted]

Thorough sensitivity analysis of the model was performed, with results presented in Table 4. In all scenarios, the ICER of ritlecitinib relative to BSC is less than £20,000.

Table 4: Scenario analyses of the updated base case of the model

Model setting tests	Base case assumption	Scenario assumptions	ICER of ritlecitinib relative to BSC (£)
Base case	-	-	14,290
Perspective	Payer	Societal	10,174
Time horizon	Lifetime	5 years	18,795

Model death in first 48 weeks?	Yes	No	14,338
Age group	≥12 years	≥18 years	15,312
Age group	≥12 years	≥12 to <18 years	13,773
Stopping rule criteria	Interim+Final	Final Only	9,877
Final SALT score	SALT ≤ 20	SALT ≤ 10	14,637
Final stopping rule time point	48 weeks	36 weeks	14,509
Discontinue patients based on SALT score after 48 weeks	SALT ≤ 20	No	16,085
Extrapolation of LT data after 24 months	Stay in state	Last observation carried forwards	16,207
Extrapolation of LT data after 24 months	Stay in state	Average	16,980
Treatment discontinuation rate curve	Weibull	Exponential	14,358
Treatment discontinuation rate curve	Weibull	Gompertz	15,089
Treatment discontinuation rate curve	Weibull	Log-logistic	14,123

Treatment discontinuation rate curve	Weibull	Lognormal	13,935
SALT >50 HCRU assumption	SALT 50-99	SALT 100	14,470
Utility weight source	TTO Analysis	TTO Analysis (SA)	13,563
Include carer disutility	Yes	No	14,347
Disutility weight source	TTO Analysis	TTO Analysis (SA)	14,290
Disutility weight source	TTO Analysis	TTO Analysis (TE 1)	14,293
Disutility weight source	TTO Analysis	TTO Analysis (TE 2)	14,310
Source of AE cost	GP	NHS Reference Costs	14,516
Caregiver disutility population	≥12 to <18 years	≥12 years	13,240
AA type: AT/AU	SALT ≥50	AT/AU only	16,625
AA type: non-AT/AU	SALT ≥50	Non-AT/AU only	13,304
Adjust AA utility weights via population norms?	Yes	No	13,820
Include wig fitting cost	Yes	No	14,343

Abbreviations: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LT, long term; SA, sensitivity analysis; SALT, Severity of Alopecia Tool; TTO, time-trade off

References

1. Fricke VAC & Miteva M. Epidemiology and burden of alopecia areata: a systematic review. *Clin Cosmet Investig Dermatol* 2015. 8: 397–403.
2. Pfizer data on file. Summary of SÆfetyWorks Analysis Methods and Results. 2018.
3. Pfizer. *A Phase 2b/3 randomised, double blind, placebo-controlled,dose ranging study to investigate the efficacy and safety of PF-06651600 in adult and adolescent alopecia areata (AA) subjects with 50% or greater scalp hair loss.* (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT03732807>
4. Pfizer data on file. Clinical validation meeting. 2023.
5. Pfizer. *Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT)*. (clinicaltrials.gov, 2022). at <https://clinicaltrials.gov/ct2/show/NCT04006457>
6. Pfizer data on file. Long-Term PF-06651600 for the Treatment of Alopecia Areata (ALLEGRO-LT) - Interim Analysis. 2023.
7. Senna M, Mostaghimi A, Ohyama M, *et al.* Long-Term Efficacy of Baricitinib in Alopecia Areata: 104-Week Results From BRAVE-AA1 and BRAVE-AA2. 2023.
8. Tosti A, Bellavista S & Iorizzo M. Alopecia areata: A long term follow-up study of 191 patients. *J Am Acad Dermatol* 2006. 55: 438–441.
9. Pratt CH, King LE Jr, Messenger AG, *et al.* Alopecia areata. *Nat Rev Primer* 2017. 3: 17011.
10. Pfizer data on file. UK interviews with dermatologists with a specialist interest in hair disorders for the use of ritlecitinib for patients with AA. 2022.
11. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. *Rep Decis Support Unit* 2011.
12. Pfizer data on file. Qualitative research in alopecia. Targeted literature review and patient advocacy group representative interviews in the United Kingdom. 2022.
13. NICE 2023. Final draft guidance: Baricitinib for treating severe alopecia areata. 2023. Available at: https://www.nice.org.uk/guidance/gid-ta10941/documents/final-appraisal-determination-document.
14. 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.
15. Bewley A, Galvan S, Johansson E, *et al.* Measuring the Burden of Alopecia Areata with the European Quality of Life-5 Dimensions (EQ-5D): Results from a real-world survey in 5 European countries. *Value Health*

16. Reid EE, Haley AC, Borovicka JH, *et al.* Clinical severity does not reliably predict quality of life in women with alopecia areata, telogen effluvium, or androgenic alopecia. *J Am Acad Dermatol* 2012. 66: e97-102.
17. Sellami R, Masmoudi J, Ouali U, *et al.* The relationship between alopecia areata and alexithymia, anxiety and depression: a case-control study. *Indian J Dermatol* 2014. 59: 421.
18. Pfizer data on file. Does the EQ-5D measure the full impact of alopecia areata on patients' quality of life? - updated manuscript. 2023.
19. Rowen D, Brazier J, Ara R, *et al.* The Role of Condition-Specific Preference-Based Measures in Health Technology Assessment. *PharmacoEconomics* 2017. 35: 33–41.
20. Hoogendoorn M, Oppe M, Boland MRS, *et al.* Exploring the Impact of Adding a Respiratory Dimension to the EQ-5D-5L. *Med Decis Mak Int J Soc Med Decis Mak* 2019. 39: 393–404.
21. Swinburn P, Lloyd A, Boye KS, *et al.* Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value Health J Int Soc Pharmacoeconomics Outcomes Res* 2013. 16: 1156–1162.
22. Rencz F, Mukuria C, Bató A, *et al.* A qualitative investigation of the relevance of skin irritation and self-confidence bolt-ons and their conceptual overlap with the EQ-5D in patients with psoriasis. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 2022. 31: 3049–3060.
23. Yang Y, Rowen D, Brazier J, *et al.* An exploratory study to test the impact on three 'bolt-on' items to the EQ-5D. *Value Health J Int Soc Pharmacoeconomics Outcomes Res* 2015. 18: 52–60.
24. NICE. 2021. Barcitinib for treating moderate to severe atopic dermatitis: technology appraisal guidance. Available at: https://www.nice.org.uk/guidance/ta681/resources/baricitinib-for-treating-moderate-to-severe-atopic-dermatitis-pdf-82609375014853.
25. Burge R, Anderson P, Austin J, *et al.* The patient-reported burden of alopecia areata by current severity: a real-world study in the United States [Poster 26158]. in (2021).
26. Rowen D, Brazier J, Wong R, *et al.* Measuring and valuing health-related quality of life when sufficient EQ-5D data is not available. *Measuring and valuing health-related quality of life when sufficient EQ-5D data is not available* 2020. at <http://nicedsu.org.uk/wp-content/uploads/2021/07/DSU-hierarchy-of-evidence-report-310720-Final-for-website-1.pdf>
27. Tosh JC, Longworth LJ & George E. Utility values in National Institute for Health and Clinical Excellence (NICE) Technology Appraisals. *Value Health J Int Soc Pharmacoeconomics Outcomes Res* 2011. 14: 102–109.

28. Beale RC, Wickstead RM, Chen G, *et al.* No EQ-5D? Analysis of Alternative Utility Value Sources Used in Nice Appraisals for Oncology Indications. *Value Health* 2017. 20: A448.
29. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. 2022. at <https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741>
30. Matza LS, Stewart KD, Lloyd AJ, *et al.* Vignette-Based Utilities: Usefulness, Limitations, and Methodological Recommendations. *Value Health* 2021. 24: 812–821.
31. Pfizer data on file. Vignette study for utility estimation in Alopecia Areata. 2022.
32. Pfizer data on file. Clinical study report. Study B7981048: alopecia areata benefit-risk trade-off study. 2022.
33. 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.
34. Liu LY, King BA & Craiglow BG. Alopecia areata is associated with impaired health-related quality of life: a survey of affected adults and children and their families. *J Am Acad Dermatol* 2018. 79: 556-558. e1.
35. Putterman E, Patel DP, Andrade G, *et al.* Severity of disease and quality of life in parents of children with alopecia areata, totalis, and universalis: a prospective, cross-sectional study. *J Am Acad Dermatol* 2019. 80: 1389–1394.
36. Ara R & Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010. 13: 509–518.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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.

NICE National Institute for Health and Care Excellence

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 **Monday 5 June 2023.** 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.

Clinical expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

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.

Clinical expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

Part 1: Treating 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 and Norwich University Hospitals NHS Trust and on behalf of British Association of Dermatologists
3. Job title or position	Consultant Dermatologist
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 alopecia areata? ☐ A specialist in the clinical evidence base for 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.	Nil.
8. What is the main aim of treatment for alopecia areata?	

(For example, to stop progression, to cure the condition, or prevent progression or disability)	
9a. What do you consider a clinically significant treatment response? (For example, a reduction in disease activity by a certain amount) **9b. Can SALT score be indicative of hair regrowth in parts of the body other than the scalp?**	
10. In your view, is there an unmet need for patients and healthcare professionals in alopecia areata?	
11. How is 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? If yes, please explain how the technology is expected to lead to this improvement (for example due to improvement in symptoms causing pain or reduced anxiety).	
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the overall 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 • Which health-related quality of life measures best capture changes in quality of life for people with alopecia areata (for adolescents and adults)?	
18. How does caring for someone with alopecia areata impact the carers quality of life? • Is the impact on quality of life different for carers of adolescents and adults with alopecia areata?	
19. 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?	
20. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
21. 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?	
22. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	
23. How do data on real-world experience compare with the trial data?	
24. 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.	

NICE National Institute for Health and Care Excellence

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

The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup	I agree that a subgroups analysis would be very informative as in clinical practice, those with AU/AT are less likely to spontaneously regrow and are also more likely to have greater difficulty achieving complete regrowth from treatments. Whilst not strictly immediately relevant to the cost efficacy issue, I fear that the selection of SALT 20 as an endpoint may have missed clinically meaningful response in some who could have started with SALT 100 (AT/AU) and regrown 75% of scalp hair to the point of not needing to wear a wig, but be counted as a non-responder. In clinical practice, not needing to wear a wig, or having “satisfactory” hair growth, as assessed by patients themselves, would be the most clinically important end point.
ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	Outside of my field and unable to comment.

Assumption of no treatment waning based on limited long-term evidence	Allegro-LT data will be helpful to inform the longer-term response to treatment. I have no personal experience of long-term patient treatment with JAK inhibitors and so am unable to comment on this further. Long-term treatment with oral immunosuppression for severe alopecia areata tends to lead to a static efficacy level at around 12 months, however some patients will require a complete 12 months of therapy with methotrexate for example, in order to reach maximal benefit. Following this time, any additive regrowth is unlikely and is usually only a small further increase in surface area. A maintenance dose is usually set following this and any dose reduction or cessation of treatment is with careful, informed clinician-patient discussions, balancing the risk of fall of hair when stopped.
Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	Agree with EAR that any long-term treatment effect data is best to be observed from the planned treatment dose only.
Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	Those with severe or very severe alopecia will have a lower likelihood of spontaneous regrowth than those with mild or moderate alopecia areata. As all participants were defined as having severe alopecia (>50%), the likelihood of spontaneous regrowth in participants will be lower than the general population of patients with alopecia areata. Those with alopecia totalis/universalis (100% hair loss) will have a further significantly reduced chance of spontaneous regrowth. We often quote in the order of 1% likelihood. The chance of spontaneous regrowth is also impacted by duration of current episode, hence the chance of spontaneous regrowth in a very severe group stopping ritlecitinib (assuming 1 year of treatment) would be very small. It is conceivable that there will be a cohort of patients with SALT closer to 50 that will have been included in the studies that will retain the ability to spontaneously regrow when moving to the BSC, even if they

	have not responded to ritlecitinib. It is extremely difficult to estimate whether treatment with a JAK inhibitor influences future chances of regrowth when the treatment is stopped, and whether this is spontaneous or a delayed treatment effect. I would suggest that only those with SALT in the lower end of the bracket 50-100 would be likely to spontaneously regrow and this effect may be lost after several years. I would suggest that spontaneous regrowth should be removed for any AU/AT subgroup analysis but retained for "milder, shorter" disease but any numerical thresholds would be arbitrary, but could be estimated based on historic epidemiological data.
Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	I am unable to comment on discontinuation rates for JAK inhibitors, as I work in NHS practice only and have not yet prescribed. Perhaps analogies can be drawn from treatment of Eczema in real world data, as I do not feel that other non-dermatological disease areas, such as Rheumatoid arthritis for example, would predict discontinuation rates in alopecia areata. In Methotrexate or Contact immunotherapy (DCPC) treatment, we explain to patients that usually treatment can take 12 months to have an effect. If no treatment effect is seen with DCPC at 12 months, we can continue to 18 months (based on sparse evidence) but patients in this category would often opt to stop due to the time commitment and risk/benefit balance. The same discussion can occur at 12 months with methotrexate, but the risk benefit discussion usually tips in favour of stopping at 12 months. These discussions are infrequent, even in tertiary practice.
Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	I am not able to comment on the aspect of using vignettes to evaluate cost-efficacy. I do, however, agree with the company that EQ5D does not appear to capture well the impact of alopecia areata on the persons affected. Data on the fourfold increase in the domain of anxiety and depression within EQ5D appears in published work, also demonstrating that the other 4 domains are largely unaffected.(Br J Dermatol. 2017: 176; 1170-1178)

	Within my consultations, patients frequently describe that they have conversations with family and friends who are worried that they "may have cancer" as they have lost their hair. Patients describe feelings of guilt and a need to convince others that they are well and healthy. I do wonder whether this issue, that is very specific to Alopecia areata (in particular totalis and universalis), gives a non-representative view of impact of the condition on health, as measured by self-reported EQ5D questionnaires, due to a feeling of a need to convince others, including loved ones, that they are well. This should of course be explored with the patient representatives themselves, but such discussions frequently occur in my secondary and tertiary consultations. I would support use of alternative methods to assess health utilities, other than EQ5D, but I am aware of the need for consistency across NICE processes.
Carer disutility based on a vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages	The impact of alopecia of any severity is significant on young people (12-17 years.) Often the impact on the individual, young person or adult, can be highly variable and not always related to the perceived severity of the condition, as defined by clinician area and severity scores. Consequently, the impact on caregivers and household members of those with alopecia will be variable too. Frequently, adults with alopecia areata will attend their consultations with family members or partners. Those accompanying individuals can frequently express the impact of the condition on the family home as whole. Other family members, such as siblings of young adults, and not just carers, are likely to be impacted also. I feel that the carer disutility is unlikely to be a linear relationship based on age or surface area of hair loss, and more an impact of individual circumstances and the noticeability of hair lost. This is based on anecdotal evidence from clinical practice. I advise gathering additional evidence of carer disutility by vignettes but including adults and differing severities of disease also.

Utility values are not age-adjusted	Outside of my field and unable to comment.
Are there any important issues that have been missed in EAR?	

NICE National Institute for Health and Care Excellence

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Alopecia areata is a complex and unpredictable condition with significant mental health impact, impact on social functioning, potential for loss of earnings and significant expense of self-treatment and prostheses/camouflage.
- There is a significant unmet need in this condition for licensed and effective treatments for severe alopecia areata.
- The treatment of Alopecia areata in young people (12-17 years) is extremely difficult due to the lack of available, safe, effective, licensed treatments and the balance between long term risks of systemic immunosuppression, or commitment to attending weekly for contact immunotherapy, with the impact of significant hair loss on adolescent development and self-identity.
- The efficacy data for ritlecitinib is encouraging and I am pleased to see the inclusion of adolescents in the clinical trials for this drug.
- EQ5D does not capture the full impact of alopecia areata on persons affected and other health utility measures should be explored.

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.

NICE National Institute for Health and Care Excellence

For more information about how we process your personal data please see our privacy notice.

Clinical expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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.

NICE National Institute for Health and Care Excellence

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 **Monday 5 June 2023.** 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.

Clinical expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

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.

Clinical expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

Part 1: Treating 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	Dr Nekma Meah
2. Name of organisation	British Association of Dermatologists British Hair & Nail Society
3. Job title or position	Consultant Dermatologist
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 alopecia areata? ☒ A specialist in the clinical evidence base for 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.	None
8. What is the main aim of treatment for alopecia areata?	The goal of treatment for alopecia areata (AA) is to suppress disease activity, prevent further hair loss and favourably impact quality of life.

(For example, to stop progression, to cure the condition, or prevent progression or disability)	
9a. What do you consider a clinically significant treatment response? (For example, a reduction in disease activity by a certain amount) **9b. Can SALT score be indicative of hair regrowth in parts of the body other than the scalp?**	A 50% reduction in the baseline SALT score and an improvement in the QoL measure Severity of Alopecia Tool (SALT) is a measure of scalp hair loss only. (0%-100%, no scalp hair loss to complete scalp hair loss)
10. In your view, is there an unmet need for patients and healthcare professionals in alopecia areata?	There is a recognised unmet need for safe, effective, accessible long-term treatments for AA in adults and children. Current systemic treatments in AA are limited *Meah N et al. The Alopecia Areata Consensus of Experts (ACE) study: Results of an international expert opinion on treatments for alopecia areata*. https://doi.org/10.1016/j.jaad.2020.03.004 and not supported by robust evidence. *Lai VWY et al. Systemic treatments for alopecia areata: A systematic review*. https://doi.org/10.1111/ajd.12913
11. How is alopecia areata currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which?	Patients with limited disease are usually treated with potent topical corticosteroids and patients with more extensive disease e.g., alopecia universalis/alopecia totalis may receive diphencyprone (DCP). NB: this is only available in certain centres. Systemic therapies e.g., immunosuppressive therapies may be initiated for extensive disease in specialist hair clinics. Some hair clinics may be supported by clinical psychology, and this may also form part of the patient's treatment journey. The treatment of alopecia areata in the NHS is based on the 2012 alopecia areata guidelines *Messenger AG et al. British Association of Dermatologists' guidelines for the management of alopecia areata 2012.* https://doi.org/10.1111/j.1365-2133.2012.10955.

• 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 is a geographical discrepancy as only a few specialist hair centres offer diphencyprone (DCP) for patients with severe AA. Access to treatment is therefore impacted by availability of local services. Patients with extensive AA in primary are referred to secondary care general dermatology for optimisation.
• What impact would the technology have on the current pathway of care?	The technology appraisal will enable patients with severe AA to consider timely effective treatment
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	The technology will be used for patients with clinically significant AA (e.g., SALT 50%). These patients are currently offered potent topical corticosteroids, contact immunotherapy or systemic immunosuppressive therapies.
• How does healthcare resource use differ between the technology and current care?	Current care for patients with severe AA includes potent topical corticosteroids, contact immunotherapy or systemic immunosuppressive therapies. Serial SALT assessment/QoL assessment, are undertaken for patients offered topical corticosteroids, contact immunotherapy, whereas serial SALT assessment/QoL assessment and monitoring bloods are undertaken for patients on systemic immunosuppressive therapies. The frequency of the blood monitoring varies for each systemic immunosuppressive therapy.
• In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic)	Secondary care dermatology setting.
•	
• What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	SALT training for some clinicians.
13. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. Results from ALLEGRO 2b/3 study has demonstrated clinically meaningful SALT scores of <20% at week 24 for patients on ritlecitinib 50mg OD. Best
•	

• 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? If yes, please explain how the technology is expected to lead to this improvement (for example due to improvement in symptoms causing pain or reduced anxiety). •	comparator DCP offers to treat only scalp hair loss, whereas ritlecitinib has been shown to encourage scalp, eyelash and eyebrow regrowth. Whilst alopecia areata is known to impact psychological wellbeing, it is not associated with impact on life expectancy. Yes. ALLEGRO 2b/3 results demonstrated an increase in health-related quality of life captured in the PGI-C Patient Global Impression of Change (PGI-C) as moderate to greatly improved. Current care e.g., DCP is site specific, whereas ritlecitinib has been shown to encourage scalp, eyelash and eyebrow regrowth.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the overall population?	AU/AT patients with longer duration of AA (i.e., >10 years), may be less responsive treatment. *Meah N 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*. https://doi.org/10.1016/j.jaad.2020.09.028
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)	I would think this would be reasonable transition. JAK inhibitors are currently being prescribe in the NHS for other dermatological indications e.g. atopic dermatitis and therefore most will be familiar with pre-screen investigations and monitoring requirements. Training on recording of SALT scores may be needed.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Other indications to consider treatment aside from patients with severe disease as aforementioned include visibility significant hair loss (difficult to conceal) with high psychological impact.

	Interruptions to treatment will occur due to primary or secondary failure. Other reasons to stop treatment would include family planning, patient choice, development of severe/intolerable side effects.
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 • Which health-related quality of life measures best capture changes in quality of life for people with alopecia areata (for adolescents and adults)?	
18. How does caring for someone with alopecia areata impact the carers quality of life? • Is the impact on quality of life different for carers of adolescents and adults with alopecia areata?	There is evidence to support that carers of AA patients are impacted. Family members can be impacted, particularly families of children compared to families of adults. *Liu LY et al. Alopecia areata is associated with impaired health-related quality of life: A survey of affected adults and children and their families.* https://doi.org/10.1016/j.jaad.2018.01.048
19. 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? •	Yes, as approval would mean timely access, timely management with effective treatment for patients with severe AA as currently DCP is restricted to only a few hair centres in the UK. Yes. Currently no systemic treatment is licensed in the UK for severe alopecia areata.

• Does the use of the technology address any particular unmet need of the patient population?	Improved QoL, reduced disease burden and greater career aspirations. Studies have shown increase rates of work absenteeism, unemployment, impact on relationships, education and career choice. *Muntyanu A et al. The burden of alopecia areata: A scoping review focusing on quality of life, mental health and work productivity*. https://doi.org/10.1111/jdv.18926
20. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The management of side effects experienced by the patient will be no different to the management of side effects experienced by any other immunosuppressive systemic agents. The most common reported side effects from ALLEGRO 2b/3 with ritlecitinib include: upper respiratory tract infection, nasopharyngitis and headache. Most adverse effects were mild or moderate in severity.
21. 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?	 The primary endpoint for the study was response based on an absolute SALT ≤20 at Week 24. Response based on a 75% improvement in SALT score from baseline (SALT 75) up to Week 48. Capturing the PGI-C response Whilst data from open label ALLEGRO-LT may be helpful (duration of the study 36 months), real-world long-term registry data will be much more informative. However, in the absence of long-term clinical trial data/real world registry data, extrapolating data at 24 months may be appropriate to predict long-term clinical outcomes. Not that I am aware of

22. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	Not that I am aware of
23. How do data on real-world experience compare with the trial data?	I do not have real world clinical experience clinical experience with ritlecitinib. I have prescribed JAK inhibitors in Australia for alopecia areata. My experience overall, has been that patients often tolerate this well with mild- moderate side effects.
24. 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.	Patients with longer duration of disease >10 years may be discriminated. Patients with special site involvement e.g., extensive bearded AA with significant cultural and religious importance may be discriminated in the appraisal. Similarly, patients with low SALT score, with significant eyebrow loss may also be affected if the technology appraisal requires patients with a minimum SALT score.

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.	

NICE National Institute for Health and Care Excellence

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

The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup	Patients with alopecia totalis/ alopecia universalis (AU/AT) have severe disease phenotype and respond poorly to treatment. However, it is noteworthy that nearly half the patients recruited to the ALLEGRO 2b/3 study had a baseline SALT of 100 (AU/AT), hence I do not think the additional subgroup analysis is imperative.
ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	I would agree it would be useful for the company to stratify for adults/adolescent subgroup. It is important to note that adolescent recruitment to ALLEGRO 2b/3 was less than 15% and therefore clarity on whether statistically meaningful data was obtained for the adolescent group specifically, would be important. This additional data would be helpful for dermatologists treating adolescents with alopecia areata (AA).
Assumption of no treatment waning	Alopecia areata is considered a chronic disease with unpredictable relapses despite ongoing maintenance treatment. For example, triggers such as infection may destabilise patients with a steady

based on limited long-term evidence	Severity Alopecia Tool (SALT) score. Whilst data from open label ALLEGRO-LT may be helpful (duration of the study 36 months), real-world long-term registry data will be much more informative. However, in the absence of long-term clinical trial data/real world registry data, extrapolating data at 24 months rather than 12 months may be reasonable to support the company's claim of no treatment waning in the long term.
Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	Patients entering ALLEGRO LT had previously participated in either the ALLEGRO 2a or 2b/3 study and 350 additional 'de novo' adult and adolescent patients not previously enrolled in either study. NB: participants from the Phase 2a & 2b/3 studies were eligible to enrol within >30 days between their last dose in their prior study and first visit in ALLEGRO-LT. Some of these patients will have been on ritlecitinib 10mg or ritlecitinib 30mg OD. Whilst the half-life of retilicitinib is understood to be short, a 30-day wash out period may be insufficient when compared to the wash out period needed in similar trials. I note that for ALLEGRO 2A ***a 12 week*** wash out period was required for patients to enter if they had used an oral or topical JAK inhibitor previously. (*King B 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*. https://doi.org/10.1016/j.jaad.2021.03.050
Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	I agree with the EAG on this point. It is plausible that spontaneous remission can occur both in the treatment and non-treatment arm. However, spontaneous remission is unlikely to occur in the severe disease AU/AT subgroup.
Company's estimate of discontinuation leads to an	AA is a chronic disease. *Cranwell WC et al. Treatment of alopecia areata: an Australian expert consensus statement*. https://doi.org/10.1111/ajd.12941 and as with other chronic dermatological

unrealistically high mean duration on treatment	conditions, patients will be required to be on treatment long term. Unlike other systemic therapies utilised in AA, reducing to lowest possible dose to maintain durable remission will not be feasible with ritlecitinib as only the 50mg dose is being considered for licensing. I agree with the EAG recommendations to repeat the survival analysis; excluding patients who stopped responding to ritlecitinib at any point during the study rather than only those at 48 weeks.
Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	I believe the EQ5D is an unsuitable health quality measure for AA patients and data from this does not capture the full impact of AA. Domains such as mobility, self-care and pain/discomfort are not relevant for patients presenting with extensive alopecia. AA carries a significant psychological burden with increase rates of work absenteeism, unemployment, impact on relationships, education and career choice. *Muntyanu A et al. The burden of alopecia areata: A scoping review focusing on quality of life, mental health and work productivity*. https://doi.org/10.1111/jdv.18926 Furthermore, it is likely that AA patients when asked on the EQ5D 'how good or bad your health is today? are likely to rate this favourably as they may not necessarily feel physically unwell/unhealthy
Carer disutility based on a vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages	I support recommendations to provide additional analysis for carer disutility. Family members can be impacted, particularly families of children compared to families of adults. *Liu LY et al. Alopecia areata is associated with impaired health-related quality of life: A survey of affected adults and children and their families.* https://doi.org/10.1016/j.jaad.2018.01.048
Utility values are not age-adjusted	I agree with the EAG's recommendations that the company should provide a model for adolescent and adult subgroups.
Are there any important issues that have been missed in EAR?	

NICE National Institute for Health and Care Excellence

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Alopecia areata is a chronic relapsing remitting disease that carries a significant psychological burden, including impact on employment, relationships, education, career choice and family members
- There is an unmet urgent need for robust effective treatment for alopecia areata in both children and adults.
- The ALLEGRO study has demonstrated promising outcomes for scalp, eyebrow and eyelash AA.
- Longer term data from ALLEGRO-LT will also be beneficial

Click or tap here to enter text.

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

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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 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 alopecia areata or caring for a patient with 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.

NICE National Institute for Health and Care Excellence

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.

Patient expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

The deadline for your response is **5pm** on **Monday 5 June 2023.** 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.

Patient expert statement

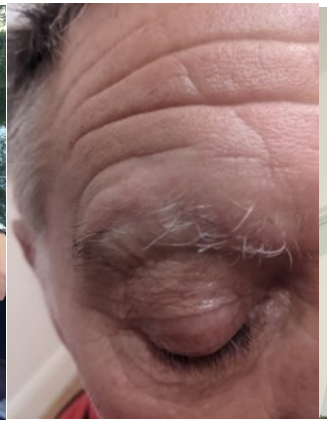
Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

Part 1: Living with this condition or caring for a patient with alopecia areata

Table 1 About you, alopecia areata, current treatments and equality

1. Your name	Andrew James Anthony (aka Tony) Ferguson
2. Are you (please tick all that apply)	☒ A patient with 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 alopecia areata? **If you are a carer (for someone with alopecia areata) please share your experience of caring for them.** **How is caring for an adolescent with alopecia areata different to caring for an adult with alopecia areata?**	I retired in Feb 2019 aged 64. I first noticed hair loss in Sept 2019 and was diagnosed with AA in Dec 2019. My hair up to that point had been very strong, still mostly dark with grey at the temples. My first diagnosis was that 80% of sufferers would see their hair fully return within 6-12 months. More spaces gradually appeared, occasionally being replaced with a white, fluffy growth. I was then told by Dermatology in Amersham that my hair should return within 12-18 months (basically an extension of six months on their original prognosis). About 18 months after original diagnosis, they told me that it could take 24 months though the bare patch at the back of my skull was “not a good sign.” It was only then that I realised that AA was my permanent condition. By that stage, I was losing my eyebrows and eyelashes which had a devastating effect on my appearance. In some ways, Lockdown was lucky timing, allowing me to become more socially introverted without drawing attention to my AA condition. By the time Lockdown ended, I was shaving the hair remnants on my head and was completely bald. 2019 -2021 -2023   

	The pictures above depict my journey over the past four years. I have lost my self-confidence and avoid meeting people who have not seen me since pre-AA. I attended my cousin's funeral back in Ireland in 2021- it was harrowing to see the look of shock on the faces of old acquaintances when they were told who I was, having at first not recognised me. Previously quite gregarious, I now restrict myself to a small circle of family and friends. I have turned down several offers of employment as I feel unable to face the wider world. To describe AA as a "cosmetic" condition is both inaccurate and gratuitously insulting. The impact on my life has been devastating. And this is an old man speaking, whose race is nearly run. It must be exponentially more difficult for younger people, adolescents and women. The scale of their suffering is indescribable and beyond my imagination.
How does caring for someone with alopecia areata impact the carers quality of life? Is the impact on quality of life different for carers of adolescents and adults with alopecia areata?	I experienced a significant impact on my partner and particularly on our sex life. My partner has supported me steadfastly throughout. However, when one loses patches of hair all over your body including pubic hair, you simply do not feel attractive. Desire and drive diminish. Psychologically, I have struggled in this respect.
7a. What do you think of the current treatments and care available for 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?**	I tried a wide variety of treatments, both NHS and Private sector, all of which proved futile. I also tried alternative remedies like acupuncture, tea tree oil, rosemary oil, heavy iron et al. I also took to self- prescribing and tried things like Minoxydil. Basically nothing has worked for me. It seems obvious that we need drugs that are designed to combat AA, not repurposed dermatology treatments. I must also confess disappointment with Amersham Dermatology dept who have been consistently uncaring.

8. If there are disadvantages for patients of current NHS treatments for alopecia areata (for example, how they are given or taken, side effects of treatment, and any others) please describe these	Current NHS treatments are simply not fit for purpose. Doctors do not know what causes AA (vague references to stress or family history, neither of which were applicable in my case). Nor do doctors know how to cure it. Remedies are “hit & miss” and usually deploying things that have been designed for different conditions like psoriasis and vitiligo.
9a. If there are advantages of ritlecitinib 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 ritlecitinib 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**	I understand there are no licensed treatments for severe AA in the UK which I find scandalous. The biggest advantage of Ritlecitnib is that it has been designed to tackle the problem. Whether it can do so effectively and without serious, adverse side effects I am not qualified to say. But it would give hope to all sufferers like me, who have lost all hope of finding an effective treatment. It seems obvious to a layman that we need treatments that are designed specifically for AA, which can be tuned through the collection of performance data.
10. If there are disadvantages of ritlecitinib over current treatments on the NHS please describe these. For example, are there any risks with ritlecitinib? If you are concerned about any potential side effects you have heard about, please describe them and explain why	
11. Are there any groups of patients who might benefit more from ritlecitinib 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	

12. Are there any potential equality issues that should be taken into account when considering alopecia areata and ritlecitinib? 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.	
13. Are there any other issues that you would like the committee to consider?	A big issue is hope. I complete these documents as an old man who has no hope of ever being cured of AA. I have tried everything I can imagine to get cured. I have spent £7k-£10k on a variety of treatments including microblading eyebrows and eyelashes, acupuncture etc. I have turned down employment and deprived the economy of tax revenue because of AA. I have exhausted the NHS. For me, there is no hope. Not having a licensed treatment in UK, specifically designed to tackle AA, adds to the darkness. Whether the answer is JAK inhibitors or some other form of medical science, sufferers need hope. Hope that they will be cured from this awful affliction. Hope that the NHS cares.

NICE National Institute for Health and Care Excellence

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

The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup	
ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	
Assumption of no treatment waning based on limited long-term evidence	
Long-term extrapolation based on data from patients receiving	

doses other than the anticipated licensed dose	
Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	
Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	
Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	EQ-5D does not fit AA and causes a distortion. As a model, it minimises the effects of AA
Carer disutility based on a vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages • We consider patient perspectives may particularly help to address this issue • Please describe if the impact on quality of life for carers of adolescents and adults with alopecia areata is similar.	In my limited experience, I would say the impacts on carers are different and also severe. In Part 1, I referred to the impact on our sex life caused by me not feeling attractive due to the effects of AA. This caused extra strain on our relationship and will have been difficult for my partner.

Utility values are not age-adjusted	
Are there any important issues that have been missed in EAR?	

NICE National Institute for Health and Care Excellence

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Even for an old man like me, AA is a devastating affliction, causing loss of confidence, anguish and severe anxiety
- People must be given hope. Hope that a cure is coming. This is especially applicable to young adults and females.
- Not having a licensed treatment in the UK is a disgrace to the NHS.
- To describe AA as a “cosmetic condition” is like describing shellshock as a “headache”. It is gratuitously insulting.
- We need treatments that have been specifically designed to tackle AA, not re-purposed dermatology afterthoughts.

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

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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 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 alopecia areata or caring for a patient with 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.

NICE National Institute for Health and Care Excellence

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.

Patient expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

The deadline for your response is **5pm** on **Monday 5 June 2023.** 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.

Patient expert statement

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

NICE National Institute for Health and Care Excellence

Part 1: Living with this condition or caring for a patient with alopecia areata

Table 1 About you, alopecia areata, current treatments and equality

1. Your name	Lynn Wilks
2. Are you (please tick all that apply)	☒ A patient with 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: Work with Alopecia UK I see FB group entries & have contacted people taking JAKs on behalf of AUK ☐ 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 alopecia areata? **If you are a carer (for someone with alopecia areata) please share your experience of caring for them.** **How is caring for an adolescent with alopecia areata different to caring for an adult with alopecia areata?**	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. I had occasionally suffered with patchy alopecia 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, I felt I lost my identity. Even with a wig I had visible differences which people did stare at and ask about. I will not go out in public without a wig on. 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. I still feel sad that I have no hair and often I am just putting on a brave face. It hurts when friends talk about their hair, a bad hair day or the therapy they feel from a good hair cut. I just have to accept and live with my visible difference. Whilst I had a good GP who did a range of blood tests and referred me to dermatology – the long waitlist meant I paid for a private appointment. It is frustrating that the NHS doesn't seem to take alopecia seriously and that the limited effective treatments are given out by a postcode lottery system

How does caring for someone with alopecia areata impact the carers quality of life? Is the impact on quality of life different for carers of adolescents and adults with alopecia areata?	From the local alopecia UK support group I attend I hear the challenges faced by parents of adolescents and young adults in caring for a person with alopecia. What can a parent do for a young person who has anxiety, depression, does not want to attend school or go out socially or is even suicidal. They themselves feel sad, frustrated there are no treatments or a cure, angry at the NHS for being unsympathetic and useless in the fact they cannot say or do anything to help. My husband comments that he hated to see me upset, depressed, and withdrawn. He was worried about my mental health. While our relationship was strong we know of people whose relationships have broken down because of the burden of alopecia
7a. What do you think of the current treatments and care available for 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 to me 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. It was frustrating to hear discussions in the STA for baricitinib (public committee meetings) about 'watch and wait' or 'best supportive care' There is no best supportive care for alopecia areata, as up until now, there were no licensed and effective treatments! 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. For many people, including me, we self fund 'best supportive care' for our lifetime with alopecia areata e.g. microblading of eyebrows every 2 years, wigs, counselling

8. If there are disadvantages for patients of current NHS treatments for 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 ritlecitinib 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 ritlecitinib 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 ritlecitinib 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 and the immense improvement in quality of life and overcoming psychosocial impacts 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 and the great improvement to my quality of life - so I could act as 'normal' and not have a visible difference 9c: Trial and word of mouth results suggest that ritlecitinib 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 and is available for adolescents
10. If there are disadvantages of ritlecitinib over current treatments on the NHS please describe these. For example, are there any risks with ritlecitinib? If you are concerned about any potential side effects you have heard about, please describe them and explain why	There are very few treatments available from the NHS and only baricitinib at time of typing this, licensed for alopecia areata. My understanding is that there are some possible side effects with ritlecitinib 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 ritlecitinib than cyclosporin and methotrexate. Also, ritlecitinib will be licensed for alopecia.

11. Are there any groups of patients who might benefit more from ritlecitinib 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	It is interesting that in your notes to the left you highlight physical impairments only With alopecia you need to consider the considerable mental health and psychosocial impacts. Anxiety, depression, isolation and even suicide. I would ask that ritlecitinib be made available on the NHS for those patients who are really suffering from these impacts and hence severe decreases in quality of life 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. For men, they are expected to put on a brave face as many suffer from male pattern baldness, yet we know they suffer the same psychosocial issues and decreases in quality of life so they could benefit more than at present.
12. Are there any potential equality issues that should be taken into account when considering alopecia areata and ritlecitinib? 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.	There is data which shows that alopecia areata may be more common in those with lower socioeconomic status. There is an equity concern that those may be hardest hit with other out of pocket costs e.g. wigs, microblading, hats, counselling. As answered above in question 11. Please consider ethnic populations, where the stigma of hair loss may be greater. Severe alopecia areata is associated with 'severe physical disability' which is classed as a disability by the UK disability and equality act 2010. 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 ritlecitinib can provide.

NICE National Institute for Health and Care Excellence

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

The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup	As a patient I feel sad that this is so focussed on assumptions & numbers and not real life data or people. I believe the EAG thought the ICER for these groups of people with less response would be higher, but in-fact it was lower. As a person with alopecia, I am just a lay person who doesn't understand all your models. I do understand the difference I would feel if I could have my hair back.
ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	Again, the focus is on one set of 'fictional' (as opposed to real life) assumptions v another. With the CIS meeting cost per qualy levels and then the EAG doesn't.
Assumption of no treatment waning based on limited long-term evidence	No comment

Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	No comment
Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	The EAG wants to exclude any cases of spontaneous remission and thinks these are already counted. We understand that spontaneous remission can occur in up to 10% people at any time, and do hear anecdotally of such cases. I will never give up hope, but then it will be great when I can access one of the two licenced, effective treatment from the NHS
Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	Think you are just getting lost in the numbers and assumptions
Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	NICE seems obsessed with EQ-5D as the 'standard' QoL tool. We heard the case against EQ-5D for alopecia areata in the baricitinib HTA and here we go again – with NICE not listening! Perhaps the committee can reflect on the clinical expert opinion from the baricitinib HTA, as I believe this is the same committee for ritlecitinib I am concerned it is noted in the ritlecitinib trials – patients were excluded a) suicidal ideation or b) severe anxiety/depression – those are the patients who could most benefit from ritlecitinib and hair regrowth. We also commented that patients recruited into a trial may be more positive from the hope of being treated. It is sad that using the EAG idea of 'EQ-5D data from the literature' that the ICER increases dramatically above the £30K threshold. PLEASE consider patients not just numbers. PLEASE consider more appropriate and specific QoL measures
Carer disutility based on a vignette for a carer of an adolescent with severe	Oh – and this is the one question where you want patient opinion. Well, I think the EAG assumptions here are nonsense!

alopecia areata has been applied at all ages • We consider patient perspectives may particularly help to address this issue • Please describe if the impact on quality of life for carers of adolescents and adults with alopecia areata is similar.	1) So you want to take out any care giver disutility for adults – ridiculous! Partners especially, as well as family and friends certainly suffer pain and have costs in caring for a person with alopecia. This may be worry so that their own mental health suffers, they have to do more tasks, if the patient stops going out socially, will not go out to do the shopping etc There may even be more financial responsibility for the carer if the adult patient cannot go to work, which I have heard of in many instances – male and female patients 2) Is the impact on QoL for carers of adolescents and adults with alopecia similar? Yes, from my personal experience and what I see and hear on the social media forums and from friends with alopecia or as carers, YES the quality of life impacts in terms of worry, support given, tasks done, response to support etc are ALL similar. Even financial burdens in terms of time away from work, time to attend hospital appointments. Of course, I will add that it is tragic to hear of a young person losing out on his/hers future, with alopecia meaning disruption from school/college and I was saddened to hear of two young adults who had disengaged from school and contemplated/tried suicide – I can't imagine the stress or 'disutility' on those parents!, That having access to ritlecitinib from the NHS could give those youngsters their life back
Utility values are not age-adjusted	Are NICE allowed to be ageist? – Just enjoyed a conversation with a very active 99yr old today
Are there any important issues that have been missed in EAR?	Oh, if only there could be a blood test or simple questionnaire, to identify those patients who have such horrendous psychosocial impacts from alopecia, and who we know would 'get their life back' and have such amazing improvements in their QoL from having hair regrowth. Please consider the patient stories, the real life data from the clinical experts. The JAKs are the first licensed, effective treatment for alopecia areata. The fact they are available for patients with AD, a similar non life threatening, chronic condition – just because they managed to have a) more appropriate QoL tools & measure and b)had a clear comparator of a similar medicine of similar costs. I ask the EAG and committee to look beyond just the numbers

NICE National Institute for Health and Care Excellence

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 ritlecitinib offers in improvement in quality of life. Consider an appropriate Qol tool in your assessment.
- I want to be offered NHS care, treatments and support for alopecia. ‘Best supportive care’ is currently very little, but that is because there are no real effective comparator treatments and little support. While this may mean ‘uncertainty in the ICER’ why should ritlecitinib not be approved because there is no real, current, expensive comparator to make it cost effective

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.

NICE National Institute for Health and Care Excellence

☐ **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

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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 **Monday 5 June 2023**. 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)	Alopecia UK
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state the name of the company, amount, and purpose of funding.	**Author – None** **Alopecia UK - Pfizer - Unrestricted research grant of £50,026.28 received in 2022. Research around the psychological impact and economic burden of alopecia areata** **Eli Lilly - £20,000 corporate sponsorship to Alopecia UK**
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
The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup	~~Yes~~/No	Presumably the company can provide this information. However, it is noteworthy that the ICER in the EAG's estimate of cost-effectiveness for the AT/AU groups is lower despite reporting poorer response rates in this group. This highlights the potential of Rictlecitinib in severe alopecia. All cost effectiveness measures in alopecia are complicated because of the lack of a standard care pathway which is acknowledged by the EAG. However, BSC in this context is considered to be non-pharmacological interventions including wigs. The reality of BSC for people with alopecia is that it is rarely at the expense of the health service, but rather comes at significant cost to the individual and their families. Therefore, the BSC cost in this scenario is minimal when the reality for those living with alopecia is often quite different. It is not fair that patients with severe alopecia areata can be denied access on the NHS to JAK inhibitors, because they are not cost effective as there is no real and costly BSC. Please consider how a JAK inhibitor can be

		recommended for a condition like AD; which has other licenced treatments as a comparator. Please do not reject ritlecitinib just because there is no easy medicinal comparator and no good standard of care pathway
ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	~~Yes/~~No	Please see comments below regarding the difficulties of only applying carer disutility to adolescents for a condition with readily recognised psychological impacts.
Assumption of no treatment waning based on limited long-term evidence	~~Yes/~~No	No further comments
Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	~~Yes/~~No	Worth noting that participants were moving from 30 mg to 50 mg and not to a dose that is higher than the anticipated licensed dose. Response in these instances may be linked to dose, but that further supports the use of a 50 mg dose in alopecia.
Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	~~Yes~~/No	As noted by the EAG, cases of spontaneous remission in severe alopecia are rare especially when hairloss has been established for a long period of time. A figure of 10% is used, but in AT/AU this is likely to be much less and remission is rarely sustained. It seems reasonable to assume that if spontaneous remission can occur in those who are not receiving treatment, then it may also occur in those who have had, and discontinued, treatment, especially when this is followed over time. Recurrence of alopecia upon cessation of a Jak inhibitor is estimated to be in the region of 50-70% (Yan et al. 2022). Given the low incidence of remission and the high relapse rate, the number of rictlecitinib responders who sustain response on BSC are likely to be negligible.

Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	~~Yes~~/No	These calculations are again complicated by the lack of a standard care pathway for alopecia. Dosing and duration of many medications differ with the indexed disease. It is likely that this will be similar in the case of JAK inhibitors in severe alopecia versus other diseases where remission is more likely.
Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	Yes/~~No~~	EQ-5D is not an appropriate instrument for measuring health utilities in alopecia when considered as a single indexed disease. Unless accompanied by co-morbid conditions, alopecia is only likely to affect two of the five domains ('Anxiety/Depression' and 'Usual Activities') for most people. Therefore, even someone who is profoundly depressed and has limited social activities as a result of their alopecia will score reasonably well based on EQ-5D. It is noteworthy that the current trial excluded those with severe depression, which further exacerbates the use of EQ-5D in this assessment. It is also worth noting that the study by Bewley and colleagues, which is cited by the EAG, found significant levels of anxiety/depression in those with severe alopecia, which contributed to the overall reduction in QoL. It is not possible to apply data from the literature which include those with anxiety and depression to a study, which excluded these individuals. Similarly, work from Edson-Heredia et al. 2022 confirm worse anxiety/depression scores on EQ-5D for those with severe alopecia compared with those who had mild or moderate alopecia. However, this study further used the WPAI and found significant impairments in work and activity in those with sever alopecia compared with those who had mild or moderate disease. Importantly, this study used a disease-specific patient reported outcome tool, Skindex-16 AA, which measured the psychosocial

		and physical effects of AA. Skindex-16 is formed of three scales covering patient symptoms (four items), emotions (seven items), and functioning (five items),22 and patients can select one of seven answers that lie on a Likert-type scale ranging from “never bothered” to “always bothered”, with scores varying from 0 (no effect) to 100 (effect experienced all the time). All scales were significantly worse in those with severe alopecia when compared with mild and moderate disease. EQ-5D as a stand-alone tool for measuring health utilities in severe alopecia is not appropriate.
Carer disutility based on a vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages	Yes~~/No~~	Carer disutility has been applied at all ages in the current analysis. Although several publications exist measuring carer disutility in adolescents with severe alopecia, data is difficult to obtain for adults with severe alopecia. However, alopecia is readily recognised as a condition that is accompanied by psychological challenges including anxiety and depression. Carer disutility for adults with depression has been studied. Prosser et al. 2015 reported significant carer disutility associated with adult depression. Furthermore, utility spill-over from patient and carer reported outcomes was comparable between adults with dementia and depression. Application of carer disutility to adults with severe alopecia warrants consideration.
Utility values are not age-adjusted	~~Yes~~/No	Utility values are not age-adjusted, but are applied in the same way in the intervention and placebo arms. Although the lack of age-adjustment may influence ICER, the change/gain in QALY between the intervention and placebo groups will be valid.

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
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

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.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
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

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]

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 **Monday 5 June 2023**. 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

NICE National Institute for Health and Care Excellence

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)	British Association of Dermatologists
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state the name of the company, amount, and purpose of funding.	**None**
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
The company has not provided a cost-effectiveness analysis for the alopecia totalis/ alopecia universalis subgroup	~~Yes~~/No	Whilst from a clinical perspective we agree that a cost-effectiveness analysis for the alopecia totalis/alopecia universalis (AT/AU) subgroup would be useful (as such patients respond poorly to treatment), it is noted that nearly half the patients involved in the ALLEGRO trial were patients with AT/AU(baseline SALT 100). Therefore, further subgroup cost-effectiveness analysis is considered not essential.
ICERs for the whole population should be based on a weighted average of outcomes for adults and adolescents	~~Yes~~/No	The adolescent population included in the ALLEGRO trial was exceptionally small (<15%), and therefore, we would support the EAG recommendations to stratify for adults and adolescents.
Assumption of no treatment waning based on limited long-term evidence	~~Yes~~/No	Alopecia areata (AA) is a chronic disease with unpredictable exacerbations. There is no guarantee of long-term sustained remission; nevertheless, patients still benefit from active treatment. We agree with the EAG that further follow-up data from ALLEGRO-LT may substantiate the company's assertion of no treatment waning in the long term. Acquiring real-world data from patients on treatment as part of clinical practice/registry data is likely to

		more informative. A national alopecia areata safety and effectiveness register is currently being set up in the UK.
Long-term extrapolation based on data from patients receiving doses other than the anticipated licensed dose	~~Yes~~/No	We would agree with the EAG in that we are not aware of any additional analyses that would help to resolve this issue. It is plausible that increasing the dose from 30 mg to 50 mg may have underestimated the results; however, if a sufficient washout period between the two doses has been applied (N.B. ritlecitinib has a systemic half-life of 2 hours), then the effects of this would be considered marginal.
Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	~~Yes~~/No	We would agree with the EAG on this issue.
Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment	~~Yes~~/No	AA is a chronic inflammatory disease with unpredictable exacerbations, and it is therefore anticipated that most patients with AA will be on treatment long term. Treatment discontinuation may occur on a case-by-case basis, e.g. family planning, non-responders, or patient choice. There is a recognised unmet need for more effective treatments for AA and the emphasis for the treating clinician is to achieve the target SALT score, rather than when treatment should be discontinued. We do not consider it appropriate to compare data of the duration of use of JAK inhibitors for other indications as hair growth as a result of active treatment takes noticeably longer. The bias highlighted by the EAG should be addressed; supporting recommendations for the company to repeat their survival analysis and censoring patients at the time they stop responding rather than excluding patients who stop responding at any future time point. **N.B.** In AA, particularly in severe disease, most hair follicles are miniaturised, and many are in telogen. This will inevitably take time to

		recover and is likely to vary between hair follicles depending on their stage in the hair cycle when the inflammation is suppressed. In a normal scalp, hair follicles can remain in telogen (exogen + kenogen) for 3-6 months. Once hair follicles are in anagen the linear rate of hair growth will not exceed 1 cm/month even if inflammation is fully suppressed. This could take 2-3 months to register on a SALT score, possibly longer. To make an analogy with corticosteroids with reference to other inflammatory skin conditions, if dermatitis is treated with a potent topical corticosteroid the response can usually be seen within a day or two; if AA is injected with corticosteroid, it would take 6-8 weeks to see the response and likely longer before this would register on a SALT score.
Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO) instead of study EQ-5D outcomes	~~Yes~~/No	Whilst we understand that EQ-5D provides a standardised measure across different technology appraisals, we would argue that EQ-5D is not an appropriate health quality measure for AA. It underestimates and inaccurately reflects the impact of AA on a patients' quality of life, for example domains such as **mobility** and **self-care** do not apply to AA patients. There were strong concerns expressed amongst the experts about the utility of EQ5D as a health quality measure for AA in this and the baricitinib appraisals. Furthermore, it was noted that patients with clinically severe depression were excluded from the ALLEGRO trial, which may have skewed the final results. Finally, the EAG has cited a few articles supporting EQ-5D as a responsive measure in different AA severities as a basis of using these data in the model. However, the Adelphi studies that were cited and supported by Lilly were put forward as a more representative utility value in the appraisal for baricitinib but were rejected by the committee, presumably due to their

		relatively poor methodology compared with the EQ-5D collected during the BRAVE studies.
Carer disutility based on a vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages	Yes/~~No~~	Studies have demonstrated that carers of AA can be affected (*Aschenbeck KA-O et al. Importance of Group Therapeutic Support for Family Members of Children with Alopecia Areata: A Cross-Sectional Survey Study*; *Liu LY et al. Alopecia areata is associated with impaired health-related quality of life: A survey of affected adults and children and their families*. Therefore, we would support recommendations to provide carer disutility before and after a response to treatment.
Utility values are not age-adjusted	~~Yes~~/No	We agree with the EAG's conclusion that the company should provide a model that includes an age adjustment for utility values (e.g. a separate analysis for adolescent and adult subgroups).

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: Company's decision to conclude that the only comparator of interest is best supportive care, defined as 'non-pharmacological therapy.' p.53 company submission	The company has developed a therapeutic pathway for AA (p. 42 company submission) where BSC is considered the preferred approach in AA management and has asserted that: • 'typically dermatologists outside of a specialist centre would not prescribe systemic treatment and patients would go straight to being treated with best supportive care (BSC)' p.43 • 'Only dermatologists with a specialist interest in hair disorders would consider prescribing systemic treatment.' p.43	~~Yes~~/No	as expressed during the Baricitinib Consultation, clinical experts strongly disagree that BSC, defined exclusively as non-pharmacological therapy, is considered the only viable comparator when assessing the efficacy and cost-effectiveness of ritlecitinib and of other emerging therapies. We would have preferred if the company aligned its comparator in line with NICE's reference case (i.e. "established clinical management without ritlecitinib") and would have presented economic modelling based on this (see our previous response for baricitinib on the results of a BHNS survey), especially because access through the NHS to wigs and orthotics is also conditional, varies by region and therefore, is not equally/universally available. Furthermore, it is likely that some sub-population groups (i.e. adolescents, males) might not be comfortable wearing wigs. Notably, the company states that the study participants could use wigs alongside pharmacological interventions, if wished, *as an add-on* and not as an alternative. *ALLEGRO 2b/3 study – 'in which placebo patients were permitted to use non-pharmacological management such as wigs.'- company submission p.21*

			'non-pharmacological treatment may have been used alongside investigational treatment and it would be an add on to pharmacological treatment rather than an alternative to it.' (p.3 -clarification questions- company submission) *'the ALLEGRO 2b/3 study, patients were able to continue using non-pharmacological clinical management such as wigs alongside the investigational treatments (ritlecitinib or placebo).' (p.3 -clarification questions- company submission)*

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.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
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

Ritlecitinib for treating severe alopecia areata in people 12 years and over. A Single Technology Appraisal

Addendum: EAG comments of the company's technical engagement response

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Emma Hock, Senior Lecturer, ScHARR, University of Sheffield, Sheffield, UK Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK Andrew Metry, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Sarah Ren, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK Gill Rooney, Programme Manager, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Emma Hock, Senior Lecturer, ScHARR, University of Sheffield, Sheffield, UK
Date completed	17/06/2023

1. Introduction

In June 2023, the company submitted their response to technical engagement (TE) for the appraisal of ritlecitinib for treating severe alopecia areata (AA) in people 12 years and over.[1] The company's TE response includes a written response form which presents a brief discussion of each of the key issues identified in the External Assessment Group (EAG) report.

The company's TE response includes updated (interim) data on SALT response from the ALLEGRO-LT study and an updated survival analysis for time to treatment discontinuation. The TE response also includes a new version of the model which had been used to generate updated cost-effectiveness estimates.

This addendum provides a brief commentary on the company's TE response, and should be read in conjunction with the EAG report. Section 2 provides a summary of the company's response and the EAG's critique of these points; whist Section 3 presents a fuller description of the EAG's critique on the company's response to particular issues and new analyses presented by the company. Section 4 provides a brief description of the changes in the updated model submitted by the company. Section 5 presents the methods for additional exploratory analyses undertaken by the EAG. Section 6 presents the results of additional exploratory analyses undertaken by the EAG.

All results presented in this document include the Patient Access Scheme (PAS) discount which reduces the cost per pack of 30 capsules from a list price of £[redacted] to £[redacted]. This is unchanged from the discount offered at the time of the original company submission (CS).[2]

2. Summary of company's TE response and EAG comments

The main points discussed in the company's TE response and the EAG's comments are summarised in Table 1. Where further critique was considered necessary, this is provided in Section 3.

Table 1: Summary of company's TE response and EAG comments

Key issue	Headline points in company's TE response	EAG comments
Key issue 1: The company has not provided a cost-effectiveness analysis for the alopecia totalis (AT)/ alopecia universalis (AU) subgroup	• The company has now provided subgroup analyses within the cost-effectiveness modelling for the population with AT/AU prior to treatment and the population without AT/AU prior to treatment. • The ICERs for patients with AT/AU are higher than the ICERs for patients without AT/AU as a greater absolute change in SALT score would be required to achieve a SALT score ≤ 20 and the probability of achieving a treatment response was therefore lower for the AT/AU subgroup (EAG report Section 4.2.4.6). • The company estimated the average ICER across both groups using a weighted average approach and using the proportion of patients expected to	The EAG is broadly happy with the updates to the cost-effectiveness model to provide ICERs for the AT/AU and non-AT/AU subgroups. The efficacy data during the first 48 weeks and the long-term transition matrices have been populated with subgroup specific data for the AT/AU and non-AT/AU subgroups. AT/AU status was a stratification factor for randomisation in ALLEGRO 2b/3 and therefore populating the model with data from these subgroups is a valid approach. However, other inputs, such as the discontinuation analysis, utilities and resource use, have not been updated to be specific to AT/AU status. The EAG notes that the subgroup analysis for AT/AU status has only been populated for the scenario where a SALT score of ≤20 is used to determine response, and is also only available when selecting age ≥12. However, data are provided for implementing only the final stopping rule (response at 48 weeks) or both the interim and final stopping rules (no worsening at 24 weeks and response at 48 weeks), with the latter being the company's base case. EAG noted that there was one AT/AU patient on ritlecitinib showing in health state SALT 21-49 at 48 weeks in the analysis without the interim stopping rule, who was showing in health state SALT ≥50 in the analysis with both the interim and final stopping rule. Given that none of the AT/AU patients failed the interim stopping rule, the reason for this discrepancy is unclear to the

Key issue	Headline points in company's TE response	EAG comments
	have AT/AU in clinical practice (9.52%). This was lower than the proportion with AT/AU in ALLEGRO 2b/3 ([redacted]%) as this group was intentionally oversampled during recruitment to provide a sufficiently large group to assess efficacy within the AT/AU subgroup. • The weighted average ICER across the two subgroups is lower than the ICER when populating the model with data pooled across both groups due to the lower proportion of patients with AT/AU.	EAG. The EAG has therefore explored the impact of setting the data equal to that showing for the final stopping rule to see if this single discrepancy is likely to be important. This had minimal impact on the ICERs.
Key issue 2: ICERs for the whole population should be based on a weighted average of outcomes for	• The company argues that age at treatment is not a treatment effect modifier and therefore it is not necessary to conduct subgroup analyses by starting age. • The company notes that the efficacy inputs for the adolescent subgroup	If the company believes that there is no difference in efficacy for adults and adolescents then it would have been appropriate to have applied the data pooled across both age subgroups in both the subgroup analysis for adults and the subgroup analysis for adolescents. The EAG believes that this would have led to the ICERs being similar for the weighted average approach and the company's preferred approach. This is because the company's approach already takes a weighted average for carer disutility, which is the only other major difference between the age-based subgroups.

Key issue	Headline points in company's TE response	EAG comments
adults and adolescents	are based on pooled data across all ages, whilst the efficacy inputs for the adult subgroup are specific to the adult cohort, meaning that the trial results for adults are over-represented in the modelling when using a weighted average approach. • The company prefers to model the population as a whole, using average baseline characteristics and applying efficacy estimates from data obtained across the whole ALLEGRO 2b/3 cohort.	The EAG notes that the company chose not to apply the same efficacy data across both subgroups. Instead, the company has populated the model with subgroup data specific for adults in the subgroup analysis for adults. However, the company has not chosen to populate the model with data specific to adolescents due to the small sample size. Had these data been provided, and used to populate the model, it would better reflect the uncertainty in cost-effectiveness in the adolescent population due to the lower number of adolescent patients included in the ALLEGRO-2b/3 trial. The EAG considers that their weighted average approach makes the best use of the data available to the EAG but presents it base case scenario using both approaches for the committee to consider.
Key issue 3: Assumption of no treatment waning based on limited long-term evidence	• The company provided an updated (interim) analysis of SALT response (SALT ≤20 and SALT ≤10) data from the ALLEGRO-LT study for all patients who had received the 50 mg dose of ritlecitinib (presented separately for combined cohorts who	The data from the updated (interim) analysis are more complete than previously presented data from the ALLEGRO-LT study, and include only those with severe AA (SALT ≥50 at baseline) and those treated with a 50 mg dose for the majority of their treatment. Whilst the EAG agrees that there is no evidence of treatment waning up to 24 months in the data presented in the company's TE response, the EAG would argue that the high proportion of missing data at 24 months, and the fact that the data are much less complete beyond 24 months, makes it difficult to conclude that there will be no treatment waning beyond 24 months, based on the evidence presented.

Key issue	Headline points in company's TE response	EAG comments
	did and did not receive a 200 mg loading dose). • The company claims that the updated (interim) analysis data "*supports stabilisation of the proportion of patients with SALT≤20 and SALT≤10 after two years of continuous treatment regardless of whether a loading dose of 200 mg was administered during the first four weeks of treatment*". • The company claims that the uncensored analyses of treatment discontinuation provides additional evidence of no waning	The EAG does not consider the treatment discontinuation analysis to be relevant for determining whether patients who have responded (i.e. achieved a SALT score ≤20) at 48 weeks will maintain a SALT score ≤20 beyond 24 months. A more detailed EAG critique of the company's TE response to this issue is presented in Section 3 of this addendum.
Key issue 4: Long-term extrapolation based on data from patients	• The company has accepted the EAG's preference to exclude patients who transitioned from a 30mg to a 50mg dose at the start of the ALLEGRO-LT study from the	The EAG is satisfied with this amendment to the company's base case to exclude patients transitioning from the 30mg dose when estimating the long-term transition matrices. The company has not explicitly commented on whether it accepts the EAG's approach for estimating the average transition matrix from the second-year data (EA1). However, this does not affect the company's base case as the company has maintained their preference for assuming

Key issue	Headline points in company's TE response	EAG comments
receiving doses other than the anticipated licensed dose	analysis used to predict transitions beyond 48 weeks.	no further transitions (i.e. a steady state) beyond 48 weeks. The EAG has confirmed that in the company's scenario analysis which implements the average transition matrices instead of assuming a steady state (Table 4 of company's TE response), the company has used the EAG's preferred approach, both in terms of excluding the patients transitioning from 30mg (EA2) and in terms of how the average transition matrix has been calculated (EA1). It should be noted that the data used in the long-term extrapolation appears not to have been updated using the latest data cut from ALLEGRO-LT which the company provided in response to issue 3.
Key issue 5: Spontaneous remission applied when patients switch from ritlecitinib to best supportive care	• The company accepts that spontaneous remission is rare especially beyond 6 months • The company maintains that the EAG were wrong to include spontaneous remission in the BSC arm but remove it from patients switching from ritlecitinib to BSC. The reason given is that ritlecitinib treatment would not be expected to alter the likelihood of patients experiencing spontaneous remission	Overall, the EAG is satisfied with the modifications the company has made to the model in response to this issue. The EAG would agree that cases of spontaneous remission are rare in those who have not experienced hair regrowth for at least 6 months, which was the target population for ALLEGRO-2b/3. The EAG notes that [redacted] patients who achieved a SALT score ≤ 10 at week 24 in the BSC arm of ALLEGRO-2b/3, had [redacted], making it unclear to the EAG how [redacted] met the inclusion criteria (EAG report pages 61 & 63). In addition, the EAG agrees that achieving a SALT score ≤ 10 at week 24 in the BSC arm of ALLEGRO-2b/3 does not necessarily indicate that the patient has experienced a spontaneous remission that will remain durable in the long-term, as assumed previously in the

Key issue	Headline points in company's TE response	EAG comments
	• Th company accepts that a SALT score ≤ 10 at week 24 in the BSC arm of ALLEGRO 2b/3 is not a good estimate of the rate of spontaneous remission which would require a SALT score of 0 to be achieved • The company has removed spontaneous remission from the model • In the updated model, patients on BSC achieving a SALT score <50 during ALLEGRO 2b/3 are assumed to stay in that state for one cycle and then gradually revert to a SALT score >50 by transitioning through one state each cycle. This is equivalent to the assumption previously applied to ritlecitinib discontinuers.	company model. For these reasons, the EAG is satisfied that cases of treatment responders in the BSC arm are no longer treated as spontaneous remissions. The EAG is satisfied that the model now makes equivalent assumptions about what happens to SALT scores in those who achieved a response by 24 weeks during the trial period on BSC and those who are modelled to switch from ritlecitinib to BSC after discontinuing treatment. Given that the ALLEGRO-2b/3 trial was restricted to patients with no evidence of regrowth within the previous 6 months, the committee may wish to consider whether the trial results and cost-effectiveness analysis are applicable to patients with symptom onset of less than 6 months or where there has been evidence of recent regrowth, as spontaneous remission may be more likely in these patients.
Key issue 6: Company's	• The company conducted a new survival analysis for time to	Based on the model selection information provided by the company the EAG does not consider that there is reason to select one parametric model over any other on the basis of fit to the

Key issue	Headline points in company's TE response	EAG comments
estimate of discontinuation leads to an unrealistically high mean duration on treatment	discontinuation, censoring patients at the time they have a SALT score ≥ 20 rather than excluding these patients from the entire analysis. • An estimate of the observed hazard indicated an increase in hazard from 1.4 years, however, as a small number of patients remain at risk at this time, the estimate is highly uncertain. • A Weibull model was selected by the company based on statistical measures of fit to the observed data, together with the decision to adopt an accelerated failure time (AFT) model	observed data. There is a substantial difference in the extrapolations. The EAG have selected the exponential distribution for its preferred base case scenario and has tested a range of extrapolations as the discontinuation rate is considered uncertain. A more detailed EAG critique of the company's TE response to this issue is presented in Section 3 of this addendum.
Key issue 7: Utilities in the model are based on vignettes which have been valued using	• The company states that EQ-5D utilities from the trial lack face validity and should not be included in the economic analysis • The company claims that published EQ-5D measures of severity are not	Overall, the EAG's position on the choice of utility data has not changed. The EAG still considers that the high baseline EQ-5D scores and the lack of change in utility during the ALELGRO 2b/3 trial, in patients whose AA severity was reduced, may be related to trial inclusion criteria or the limited duration of follow-up, and does not necessarily indicate that EQ-5D is inappropriate for measuring utility in patients with AA. In addition, the EAG maintains that the published outcomes reported from the Adelphi AA database suggest that the EQ-5D can

Key issue	Headline points in company's TE response	EAG comments
time-trade-off (TTO) instead of study EQ-5D outcomes	aligned to SALT scores and are at risk of bias • The company claims that it has provided evidence that EQ-5D lacks sensitivity and content validity in AA • The company states that the vignette study methodology follows best practice and any deviation risks introducing bias. • The company states that discrepancies between PRO outcomes reported in ALLEGRO 2b/3 and vignette descriptions are reasonable as the latter is informed by multiple sources of evidence. • The company has maintained their preference to use utility data from vignettes.	distinguish between patients with different levels of AA severity. No additional evidence has been provided that impacts on either of these EAG conclusions with the company mainly referring to evidence already referenced in the EAG report. In particular, the EAG does not consider the fact that severity categories in the Adelphi AA database were physician assessed rather than been based on SALT scores, as being sufficient to discount the estimates from the Adelphi AA database. The EAG still considers that the data from the vignette study should be treated with caution. A more detailed EAG critique of the company's TE response to this issue is presented in Section 3 of this addendum.
Key issue 8: Carer disutility based on a	• The company has included the EAG's preference that carer disutility is applied only to carers of	Overall, the EAG is satisfied that the company's additional analyses have explored the potential impact of assuming some disutility for carers of adolescents with a SALT scores ≥ 50, and have found that this is not likely to be a significant driver of the ICER. However, it notes that all of

Key issue	Headline points in company's TE response	EAG comments
vignette for a carer of an adolescent with severe alopecia areata has been applied at all ages	adolescent patients in their updated model. • The company states that whilst there is evidence of a caregiver burden in a proportion of people caring for adults with AA, the burden is greatest in carers of adolescents. • The company also accepts the EAG's point that the vignette study did not seek to estimate carer disutility for carers of adult patients. • The company also accepts that caregiver disutility should apply to states other than the SALT ≥ 50 health state and has explored three alternative approaches to estimating these using the estimate of disutility for SALT ≥ 50 and various alternative assumptions.	the scenarios provided by the company are based solely on assumptions regarding the proportion of disutility that will be resolved by various improvements in SALT scores. The EAG still considers this to be an area of significant uncertainty because the company has not measured utility in caregivers before and after a response to treatment. A more detailed EAG critique of the company's TE response to this issue is presented in Section 3 of this addendum. The EAG notes that the company did not examine the ICER specifically in the adolescent subgroup when exploring the impact of the three approaches to estimate carer disutility for SALT >50 health states and has provided scenario analyses exploring this using the EAG's preferred base case assumptions.

Key issue	Headline points in company's TE response	EAG comments
	• These alternative scenarios did not have a large impact on the ICER in the company's scenario analysis.	
Key issue 9: Utility values are not age-adjusted	• The company has included an adjustment in the model to account for the expected decline in utility values in the general population with age, based on data from Ara and Brazier (2010).[3]	The EAG is satisfied with the company's inclusion of an age-adjustment for utility.

Abbreviations: AA - alopecia areata;AT - alopecia totalis; AU - alopecia universalis; BSC - best supportive care; EA - exploratory analysis; EQ-5D - EuroQOL quality of life measure, 5 Dimensions; ICER - incremental cost-effectiveness ratio; PRO - patients reported outcomes; SALT - Severity of Alopecia Tool.

3. EAG's critique on key issues 3, 6, 7, and 8.

The EAG has already made brief comments on issues 1, 2, 4, 5 and 9 in Table 1 and does not consider it necessary to provide further commentary on these issues. However, additional critique is provided below on the company's responses to issues 3, 6, 7 and 8.

Key issue 3: Assumption of no treatment waning based on limited long-term evidence

In their TE Response,[1] the company have presented new data from an updated (interim) analysis of the ALLEGRO-LT study (data cut-off date [redacted]). The company has reported this for two groups, referred to as "Ritlecitinib 200/50 mg" and "Ritlecitinib 50 mg", and the company's results are reproduced in ***[redacted]2. The first group included all patients who received a 200 mg loading dose of ritlecitinib followed by 50 mg ritlecitinib (including those treated with a 200/50 mg regime and those who began on placebo and then transitioned to a 200/50 mg regime in the ALLEGRO 2b/3 study, and *de novo* patients entering the ALLEGRO-LT study (with a baseline SALT score of ≥50) who received a 200/50 mg ritlecitinib regime). The second group included all patients who received a 50 mg dose of ritlecitinib (including those treated with a 50 mg dose in the ALLEGRO 2b/3 study and placebo in the ALLEGRO 2b/3 study and were treated with a 50 mg dose of ritlecitinib in the ALLEGRO-LT study) (see ***[redacted]2). These patient cohorts align with the EAG's preference for removing patients who received the 30 mg dose (with or without a 200 mg loading dose); thus the updated (interim) data is provided for patients who have received the anticipated licensed dose of ritlecitinib for the majority of their treatment. According to this data, [redacted] had reached Month 24 ([redacted]) as of the data cut-off date.

The updated data presented in the company's TE Response[1] is more complete than the data presented in and alongside the CS,[2] due to the later data cut. It is also considered by the EAG to be more representative of the anticipated licensed dosing regimen for the following reasons. Firstly, patients transitioning from the 30 mg dose to the 50 mg dose have been excluded, and secondly, the cohort has been restricted to those with a SALT ≥ 50 at baseline.

However, the EAG does note that there are a substantial proportion of patients with missing data at 24 months in both groups ([redacted] of the [redacted] patients started on the 200/50mg dosing regimen; [redacted] of the [redacted] started on the 50mg dosing regimen). Furthermore, the reason for these patients having missing data up to 24 months (in all except [redacted] case) is not due to this being an ongoing study. It is unclear how missing data have been dealt with in this interim analysis. In previous analyses presented in the CS, the company assumed that patients with data missing due to COVID-19 were missing at random, but those with data missing for other reasons were assumed not to have responded. It is unclear what assumptions the company has applied in this case, although it appears to the EAG that all persons with missing data

have been assumed to be missing at random based on the data presented in ***[redacted]2. Nevertheless, the EAG notes that when calculated as a proportion of the overall sample size, although percentages of patients in each pooled treatment group attaining a SALT score of ≤ 20 are lower than those reported in ***[redacted]2, they remain stable over time, with no evidence of treatment waning up to 24 months (see ***[redacted]1). However, the less complete data beyond 24 months (see ***[redacted]2) means that the proportion of responders falls after 24 months, particularly in the 200/50 mg cohort, when assuming that those with missing data are non-responders, as shown in ***[redacted]1.

[redacted]1

[redacted]

The company have stated that the data from the updated (interim) analysis of the ALLEGRO-LT study support the assumption of no treatment waning. Whilst the EAG agrees that there is no evidence of treatment waning up to 24 months in the data presented in the company's TE response, the EAG would argue that the high proportion of missing data at 24 months, and the fact that the data are much less complete beyond 24 months, makes it difficult to conclude that there will be no treatment waning beyond 24 months. It also notes that the assumption of no treatment waning is only applied beyond 24 months in the economic analysis and is then applied indefinitely provided patients do not discontinue treatment for other reasons. The EAG therefore maintains its preference for using the average transition matrix calculated using data from the second year of follow-up, using the EAG's pooled data approach, in its base case analysis.

The company claims that their uncensored analysis of discontinuation provides additional evidence of no waning as it shows a high time on treatment. However, a patient's choice to continue treatment cannot be used as direct evidence of the patient maintaining a SALT score ≤20. The EAG therefore does not consider this evidence to be relevant to the question of whether there will be treatment waning after 24 months.

[redacted]

	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

[redacted]

Key issue 6: Company's estimate of discontinuation leads to an unrealistically high mean duration on treatment

The company has conducted an updated survival analysis for time to discontinuation, censoring patients at the time they have a SALT score ≥20 rather than excluding these patients from the entire analysis, which was their previous approach. The company states that the methods used are in line with NICE TSD 14 recommendations however there is some misinterpretation of the guidance as outlined below. The company present an estimate of the observed hazard rate over time, which is reproduced as [redacted]2 below. Although this is a useful technique when selecting an appropriate parametric model, choice of AFT/PH model is not relevant in this situation. [redacted]2 indicates an apparent change in hazard at around 1.4 years, however, as pointed out by the company, there are only a small number of patients remaining at risk of discontinuation at this time (n=[redacted] out of an initial [redacted] at risk). Therefore, the hazard estimate is highly uncertain towards the end of the time period and extreme estimates are not unusual in this situation. The company has not stated the method used to generate this plot. Including further details and providing confidence intervals would offer further clarity. Other diagnostic plots recommended to inform model choice (e.g. complimentary log-log plot) in NICE TSD 14 were not provided.

[redacted]2 [redacted]

[redacted]

Weibull was selected by the company based on statistical measures of fit to the observed data (AIC/BIC as shown in Table 3), together with the decision to adopt an AFT model (which is not necessary in this case). However, the EAG notes that there is little difference in the AIC/BIC statistics between the models and the AIC/BIC figures for the exponential model are lower. There is also not much difference in visual fit to the KM data (see * [redacted] 3). The EAG therefore consider that there is no reason to select one model over any other on the basis of fit to the observed data. Whilst the company argues that the hazards are not constant based on the hazards plot provided, there does not appear to be much variation in the hazards up to 1.4 years, the point at which the company states that the data are less reliable due to a reduction in patient numbers. There is, however, a substantial difference in the extrapolations (**

[redacted] 4). The EAG prefers to use an exponential risk of discontinuation, based on this curve having the lowest AIC/BIC figures and the hazards appear relatively stable up to 1.4 years. However, the EAG notes that the apparent increase in discontinuations after 1.4 years is not adequately explained by the company. The company has also not provided the number at risk at each time point other than stating that there were [redacted] patients with exposure less than 1.4 years and [redacted] remaining at risk after 1.4 years. A table showing the timing of both the discontinuation events and the censoring events would have been helpful in interpreting the hazard plots in this case. From the information provided by the company, the EAG estimates that [redacted] patients were censored before 1.4 years ([redacted]). The EAG attempted to reconstruct the survival data from the information provided by the company. From this it believes that there were high numbers censored around 24 weeks and 40 weeks but the exact timing of the censoring is uncertain due to the limited details provided by the company. However, this would be in keeping with the large drop-off in SALT score follow-up data between 15 and 18 months (as time in the survival analysis is measured as time since 48 weeks, so 24 weeks in the survival analysis is equivalent to a total follow-up time of ~17 months) and the low numbers with follow up at 2 years, as previously discussed in EAG report Section 5.3.4.4. In addition, the EAG believes that the survival analysis for discontinuation has not been conducted using the latest data cut-off from ALLEGRO-LT, as presented by the company in response to issue 3, as this would be expected to provide more complete follow-up than the analysis presented in the previous model. Instead, the updated analysis provides a change in hazard occurring at ~1.4 years which is consistent with the data provided in the post clarification model, suggesting it is based on the same duration of follow-up.

Given the EAG's concerns regarding the high degree of censoring around 40 weeks, and the variation in long-term discontinuation rates predicted by the various parametric models, which all provide a reasonable fit to the observed data, the EAG considers the true long-term discontinuation rate to be uncertain. It has therefore selected the exponential curve, as this has the lowest AIC/BIC and has explored the impact of applying both the Gompertz curve which has the highest rate of discontinuation and the lognormal curve which has the lowest.

Table 3: AIC and BIC statistics for parametric distributions fit to ALLEGRO-LT discontinuation (reproduced from company's TE response to issue 6, Table 1)

Distribution	AIC	BIC
Exponential	[redacted]	[redacted]
Weibull	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]
Lognormal	[redacted]	[redacted]
Generalised Gamma	[redacted]	[redacted]

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion

[redacted]

[redacted]

[redacted]

[redacted]

Key issue 7: Utilities in the model are based on vignettes which have been valued using time-trade-off (TTO)

Many of the comments provided by the company on this issue relate to evidence previous submitted within the CS, which the EAG has already considered and critiqued.[4-12] The EAG has tried to focus its

comments in this section on any new evidence cited by the company or additional clarification provided by the company on the evidence already submitted.

The company states, "*it has provided evidence that EQ-5D lacks sensitivity and content validity in AA (Doc B 3.4.1 and EQ-5D Manuscript).*"[2, 8] In terms of content validity, the EAG does not believe that the company has provided any new evidence and it relies predominantly on the evidence previously submitted which included literature reviews, qualitative research with patient advocacy groups and clinicians, and quantitative evidence on the EQ-5D from ALLEGRO 2b/3.[6, 8, 12] The EAG's discussion of the company's evidence on the content validity, construct validity and responsiveness of the EQ-5D based on the evidence previously submitted can be found in Section 5.3.4.10 of the EAG report. In terms of new evidence, the company's TE response discusses the fact that the EuroQol group are exploring possible bolt-on dimensions to the EQ-5D for a range of conditions, with psoriasis being the only skin condition mentioned,[13, 14] and the other clinical areas being respiratory disease, vision, hearing and tiredness.[15, 16] Whilst this evidence provides some broader context, it does not specifically demonstrate that the EQ-5D lacks face validity in AA.

As part of its assessment of the appropriateness the EQ-5D, the EAG cited a review by Yang et al, which reported the EQ-5D as having good validity and responsiveness in people with a range of skin conditions, with 12 of the 16 papers being in patients with psoriasis.[17] The company's TE response comments that this should not be characterised as covering a range of skin conditions because of the predominance of psoriasis and advises caution in using studies from other skin conditions to draw conclusions regarding the appropriateness of the EQ-5D in AA. The EAG would agree that the most relevant information to answer the question of whether the EQ-5D is appropriate in AA would come from studies in patients with AA. However, it notes that the company's TE response describes research to develop "bolt-on" dimensions for respiratory disease, vision, hearing and tiredness when discussing the potential inappropriateness of the EQ-5D in severe AA.[15, 16] The EAG considers that these are less relevant to the question of whether the EQ-5D is appropriate in severe AA than evidence from other skin conditions. The company's TE response also draws on conclusions made in the appraisal of baricitinib for atopic dermatitis,[18] and references two papers discussing adaptations to the EQ-5D for patients with psoriasis.[13, 14] The EAG would argue that for consistency both of these should be considered irrelevant if the review by Yang *et al.*[17] is to be discounted because it considered other skin conditions.

With regards to the sensitivity of the EQ-5D, the company states, "*empirical evidence also suggests the EQ-5D may be insensitive to the full impact of AA,*" and cites, three sources of evidence,[7, 8, 10] one of which is an analysis of the ALLEGRO 2b/3 study reported in the unpublished company sponsored study by Lloyd *et al.*[8] Whilst the conclusion of this paper is accurately summarised in the company's

statement, the remaining two sources cited do not support this conclusion.[7, 10] Burge *et al.* reports that *"for the EQ5D-5L patients reported a lower quality of life with increasing physical-rated severity of AA",*[7] and Bewley *et al.* concludes that "*Patients with severe AA reported lower QoL and higher anxiety and/or depression than patients with mild and moderate AA, as measured by EQ-5D".*[10] One key difference between these sources is that the studies reported by Burge *et al.* and Bewley *et al.* were cross-sectional, whereas, the comparison reported by Lloyd *et al.* used EQ-5D scores obtained during ALLEGRO 2b/3 study in which all patients had severe AA at baseline.[7, 8, 10] It is possible that EQ-5D differences exist between patients with different levels of AA severity, but these simply did not become apparent within the timeframe of the ALLEGRO 2b/3 study because patients were still adapting to their new AA severity level as previously suggested in the EAG report.

The company also cites statements from the British Association of Dermatologists (BAD) and Alopecia UK from the baricitinib appraisal committee meetings, which the EAG is unable to corroborate, and so has not reproduced here. However the EAG notes that the TE responses from both the BAD and Alopecia UK state that they do not consider EQ-5D to be an appropriate measure of HRQoL in AA.[19, 20] In addition both BAD and Alopecia UK noted that the clinical trials excluded patients with significant depression, and both commented that the data from the Adelphi AA data base, reported by Bewley *et al.*, may be more representative.[19, 20]

As part of its assessment of construct validity, the EAG considered three studies reporting outcomes from the Adelphi AA database, which the EAG considers do provide evidence of construct validity for the EQ-5D specifically in AA.[7, 10, 11] Furthermore, the EAG considers that the data reported by Bewley *et al.* for the European cohort are an acceptable source of utility values for the economic model.[10] On this point, the company states that, *"published EQ-5D measures of severity are not aligned to SALT scores and are at risk of bias*". The company's TE response then goes on to argue that the studies by Bewley *et al.* and Edson-Heredia (the latter of which reported outcomes from the same database as Bewley but for a Japanese cohort) are subject to bias because severity was based on clinical judgement.[10, 11] The company states, "*studies suggest a moderate to low patient-physician alignment in hair loss severity*," and cites two studies in support of this statement.[11, 21] The first study cited by the company is the study by Edson-Heredia *et al.* which reports outcomes from the Adelphi AA database for the Japanese cohort [11] The EAG disagrees that this study shows moderate to low physician and patient alignment for severity, as the study reports a moderate level of agreement (kappa 0.60, $p<0.001$).[11] The second study cited by the company, by Reid *et al.*, reports that physician and patient reported hair loss severity are highly correlated (Spearman correlation = 0.47, $P < 0.0001$), but the average severity is higher for patients than physicians.[21] The EAG considers that this second study is of limited relevance as there were only a small number of patients with AA (n=23 out of 104 in total).[21] In addition, the scale used by the patients and physicians were not equivalent. The patients were asked

to rate severity on a 5-point scale, whereas clinicians were required to use the SALT score to determine severity (SALT<25=1, SALT 25-49=2, SALT 50-74=3, SALT 75-99=4 and SALT 100=5). As the authors state, one possible reason for the discrepancy in severity is that, "*patients are not aware of the full spectrum of possible disease outcomes when rating severity*".[21] The severity scales used for the patients with skin conditions other than AA who made up the remainder of the cohort in this paper, were also different, making it hard to see how the study provides a conclusion specific to AA. Overall, the EAG does not consider that the two papers cited support the company's conclusion that there is poor agreement on severity between patients and clinicians. Furthermore, this is not really relevant, as the key question is whether the physician rated severity correlates strongly to the SALT score, as the SALT score has been used to define the health states in the model for which utility scores are required. As previously discussed in the EAG report (page 126), Edson-Heredia *et al.* report that the mean hair loss was 8.2%, 26.2% and 72% in patients with mild, moderate, and severe AA, suggesting that physician assessed severity was strongly correlated with hair loss in the Japanese cohort of the Adelphi AA database.[11] The authors also report that the degree of hair loss from the scalp was described as the main indicator of severity for physicians (85% of physicians), followed by the patient's level of distress (10% of physicians).[11] Whilst fewer details are available for the abstracts by Bewley *et al.* and Burge *et al.* on the method of severity assessment, both report outcomes from the same database (Adelphi AA) and describe severity as being based on clinical judgement or physician assessment.[7, 10] These factors were previously considered by the EAG when selecting the estimate from Bewley *et al.* for use in their basecase and the EAG does not consider that the issues raised by the company in their TE response are sufficient to warrant ignoring this data source.

The company argues that the EQ-5D results from the trial lack face validity and should not be included in the economic analysis. As previously stated in the EAG report (Issue 7 summary, p13), the EAG accepts that the ALLEGRO 2b/3 trial outcomes may not have captured the full potential benefit of treatment, but considers that this was potentially related to the limited length of trial follow-up and the selective population which excluded patients with significant anxiety or depression. The company is claiming that the FAD of baricitinib for the treatment of severe AA supports their decision not to use the EQ-5D data because it acknowledges that the EQ-5D may not be capturing important aspects of the condition, but the EAG does not believe the company has captured the context of the committee's conclusions accurately. The FAD states, "*The committee concluded that severe alopecia areata can have a profound impact on quality of life that is not shown in the overall baseline EQ-5D-5L scores for people taking part in the BRAVE trials. It considered that this could be because the EQ-5D-5L may not be picking up important aspects of the condition or people in the trials may not be representative of people with severe alopecia areata being treated in the NHS in terms of anxiety and depression."*[22] The EAG believes that the committee's consideration of the BRAVE trial was in-line with the EAG's interpretation of the results from ALLEGRO 2b/3, in that the EAG considers that the high baseline

values could be due to the population rather than being conclusive that the EQ-5D is not appropriate. In addition, the EAG notes that the committee concluded to consider both the EQ-5D outcomes from the BRAVE trial and the EQ-5D outcomes from the Adelphi AA database in their decision making in the appraisal of baricitinib.[22] The committee did not conclude that the EQ-5D was inappropriate for decision making in severe AA, although the EAG acknowledges that the committee was not provided in that appraisal with data from a vignette study as an alternative data source. The EAG maintains that presenting an economic analysis using data from the trial would be in keeping with the NICE reference case and the company should have provided this to the committee, before going on to present a non-reference case scenario. This is supported by the fact that the committee in the baricitinib appraisal decided to consider both the trial-based EQ-5D outcomes and the estimates from the Adelphi AA database.[22] Overall, the EAG considers that it is up to the company to persuade the committee that it has sufficient evidence to demonstrate that the EQ-5D is not appropriate specifically in AA. The EAG does not believe that the company has provided any substantive new evidence at TE on this point. It also does not consider the fact that the severity states in the Adelphi AA database were physician assessed to be sufficient to discount these literature-based estimates of EQ-5D, especially given that the information from Edson-Heredia *et al.* suggests that the physician assessed severity was mainly based on SALT scores. It therefore maintains its preference to use the data from Bewley *et al.* in its base case.

In the EAG report, it was highlighted that the patients involved in the qualitative research to inform the vignettes were all required to have had specific treatments previously or be interested in receiving systemic treatment. The EAG therefore questioned whether the vignettes would be applicable to patients who have similar SALT scores, but who are currently managing their severe AA without receiving or seeking a systematic treatment (EAG report, p137). In their TE response the company states that their focus is on patients seeking treatment, which is the population for which the vignette study was attempting to ascertain utility values. However, the EAG notes that the proposed marketing authorisation [redacted]. Therefore, it is important for the committee to consider how applicable the QALY gains estimated from the company's model would be to the broader group of patients who would not have been eligible to participate in the research informing the vignette study, as these patients may also be offered treatment if a positive recommendation is made.

The EAG's report (p137) questioned the face validity of the vignette descriptions due to apparent inconsistencies between the language in the vignettes and the AAPPO outcomes reported by SALT score from the ALLEGRO 2b/3 which was used to inform the vignettes. The company's TE response argues that this places too much emphasis on this single source of evidence and that the vignette development involved many other sources including interviews with patients and clinicians. The

company states in its TE response that data from the AAPPO preference study were also considered and that these showed outcomes that were slightly more severe than the outcomes from ALLEGRO 2b/3.[5, 6] The EAG notes that the company previously stated (CS, Appendix H, p52) that outcomes from the preference study were "*broadly consistent with the trial data*," which seems inconsistent with the company's TE response. In addition, the AAPPO preference study only included outcomes from a single time point for patients with SALT ≥50 and therefore the company stated previously that the outcomes from the trial were the primary source of information for developing the draft vignettes (CS, Appendix H, p52). Even when considering the information from the AAPPO preference study, the EAG is unsure how statements such as "[redacted]"[2] are supported when only [redacted]% of adults responded 'often' and [redacted]% responded 'sometimes' for the question of how often they feel self-conscious.[5] In addition, the AAPPO preference study reports that [redacted]% of adults responded 'not at all' to the question on how often they limit their interactions with others.[5] This is lower than the [redacted]% who responded 'not at all' to this question in the ALLEGRO 2b/3 study for SALT ≥50.[2] However, the EAG still considers that the statement in the vignette (A3, describing SALT 50-100) which says, "[redacted]"[2] is stronger than would be supported by either evidence source. In particular it doesn't say [redacted] and the frequency element of the statement only applies when it says "[redacted]". The company's TE response also describes how the interviews with patients contributed to the language being modified between the draft and final vignettes and states that based on the various sources of evidence, including the interviews with patients and clinicians, "[redacted]." The EAG accept that multiple sources of evidence should be brought to bear when developing vignettes, including the experiences of patients with the condition and those clinicians who treat them. However, it still considers that the vignettes lack face validity when compared to the quantitative data used to derive them, even when considering the alternative data source of the AAPPO preference study discussed in the company's TE response. None of the additional clarification provided by the company at TE on the conduct of the vignette study has modified the EAG's opinion that the utility estimates from the vignettes should be treated with caution.

Key issue 8: Carer disutility based on a vignette for a caregiver of an adolescent with severe alopecia areata has been applied at all ages

As discussed previously in Table 1, the company has accepted the EAG's preference that caregiver disutility is only applied during a patient's adolescent years because this corresponds with the description of the health state valued in the vignette study. The company has also updated their analysis to address one of the EAG's secondary concerns on caregiver disutility. In the company's original analysis, caregiver disutility was only applied to carers of patients with a SALT score ≥50. However, the EAG did not consider it reasonable to assume that all caregiver disutility would resolve if the patient's SALT score decreased below 50. In response to this, the company has provided three scenarios exploring alternative assumptions for the caregiver disutility for patients with a SALT score <50. These are summarised in Table 4. None of these are included in the company's TE base case.

Table 4: Company's scenario analysis exploring the impact of including carer disutility across all heath states for carers of adolescents

Scenario description	SALT 50-100	SALT 21-49	SALT 11-20	SALT 0-10	ICER (≥12 years) *
Company's base case	[redacted]	[redacted]	[redacted]	[redacted]	14,290
A) SALT 21-49 has 50% of disutility of SALT 50-100	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
B) Assuming 75%, 50% and 25% of the disutility for SALT 50-100 applies to the other states in declining order of severity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
C) Anchoring the magnitude of caregiver disutility in each health state proportionally to the decrement in utility for each health state compared to the utility of patients with a SALT score of 0-10.***	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

** for age ≥ 12 years using the company's approach, not the EAG's weighted mean across age groups*

*** this was reported as [redacted] in the text of the company's TE response, but as [redacted] in the company's table, however, the EAG can only reproduce the former figure, so has reported this figure*

**** the EAG notes that the figures for scenario C change when selecting different age groups and different utility sources for the main health states but are reported here for the company's base case scenario, i.e. age ≥12 years and vignette study utilities*

The EAG has validated the application of these three scenarios within the model. However, it notes that each of these scenarios are based on assumptions regarding the proportion of the carer disutility experienced by for caregivers of patients with a SALT score ≥50 that will apply if a reduction in SALT score is achieved.

In addition, the EAG still considers that the estimates of carer disutility are uncertain because the company has not directly measured disutility in caregivers before and after a response to treatment. Instead it has estimated the disutility by comparing the absolute utility value from the vignette for a carer of a patient with a SALT score ≥50 with the utility for an age-matched member of the general population. The company has therefore not directly measured whether the disutility associated with caring for an adolescent with severe AA is likely to resolve if the patient responds to treatment. The EAG stands by its previous statement that the impact of severe AA on caregivers may remain even in the event of a treatment response, particularly if there were parental concerns regarding the long-term safety of remaining on treatment and the likelihood of relapse if the treatment is stopped.

4. Summary on the changes of the updated economic analysis presented by the company

Table 2 summarises the company's original base case model in the CS, the EAG's preferred analysis in the EAG report, and the company's updated base case model as presented in the company's TE response. It also indicates whether there is now agreement between the company's TE model and the EAG's preferences or whether the EAG considers a particular issue to remain unresolved.

In response to key issues 4, 5, 6, 8 and 9, the company has updated its base case analysis. These changes have been briefly described in Table 1, with further information provided in Section 3 on issues 6 and 8. In addition to these changes, the company also accepted the EAG's preferences with regards to: including mortality in the first 48 weeks of the model; restricting carer disutility to carers of adolescents; assuming the same psychological support for ritlecitinib as for BSC; assuming TEAEs are managed in primary care; and the correction of implementation errors identified by the EAG. In addition, the company has included the age-adjustment for utilities as requested (issue 9) and has incorporated the EAG's correction for handling missing data when implementing the stopping rules, which the EAG previously included in a scenario analysis. The company's updated base case does not implement the EAG's preferred utility values (issue 7). It also maintains the company's preference for assuming no waning of treatment effect in contrast to the EAG's approach which used average transition matrices from the second year (issue 3). It also incorporates an updated survival curve for estimating treatment discontinuation (issue 6). The company has also provided a subgroup analysis for the population with AT/AU (issue 1). It has also provided scenario analyses exploring the impact of including carer disutility for states other than SALT ≥50 in response to issue 8. The company has not accepted the EAG's preference for using a weighted average of outcome across the two age subgroups (12-18 years and ≥18 years), as discussed in issue 2. Table 6

Table 5 Summary of company's original base case (CS), EAG-preferred analysis (EAG report) and company's updated base case (TE response)

Aspect of model/ issue identified in the EAG report Section 5.3.4	Company's original base case	EAG-preferred analysis	Company's updated base case	Agreement between EAG-preferred and updated company's base case
EAG's corrected company base case: correcting implementation errors in the company's economic model [included in all subsequent rows]	No	Yes	Yes	✓
EA1: Using pooled counts from the second year to estimate the 3-month transition matrix applied from 2 years onwards	No	Yes	No, company assumes steady state	✗
EA2: Using only patients who were on the 50 mg dose to inform the long-term matrices	No	Yes	Yes	✓
EA3: Assuming no spontaneous remission in the ritlecitinib arm	No	Yes	Yes, with spontaneous remission also removed for BSC	EAG prefers the company's updated approach
EA4: Assuming double the hazard of discontinuation applied in the company's base case	No	Yes	New discontinuation survival analysis	EAG still considers the discontinuation risk to be uncertain
EA5: Allowing mortality in the first 48 weeks of the model	No	Yes	Yes	✓
EA6: Using the utility values reported by Bewley *et al.*[10]	No	Yes	No, company uses vignettes	✗
EA7: Carer disutility applies only during adolescent years	No	Yes	Yes	✓
EA8: Assuming the same psychological support for ritlecitinib as for BSC	No	Yes	Yes	✓
EA9: Assuming TEAEs are managed in primary care	No	Yes	Yes	✓
Additional issues highlighted but not included in EAG base case				
Age-adjustment for utilities	No	Not implemented, but requested	Yes	EAG prefers the company's updated approach

Aspect of model/ issue identified in the EAG report Section 5.3.4	Company's original base case	EAG-preferred analysis	Company's updated base case	Agreement between EAG-preferred and updated company's base case
Handling of missing patients	No	No, but explored in EAG scenario 8	Yes	EAG prefers the company's updated approach
Subgroup analysis for AT/AU	No	Included as scenario only with caveats	Included as scenario only	EAG prefers the company's updated approach for this subgroup

Abbreviations: AT - alopecia totalis; AU - alopecia universalis; BSC - best supportive care; EA - exploratory analysis; SALT - Severity of Alopecia Tool; TE, technical engagement; TEAE, treatment-emergent adverse events.

5. Methods of the EAG's TE exploratory analyses

Company changes adopted by the EAG

The EAG base case has incorporated the age-adjustment for utilities that the EAG expressed a preference for but was unable to incorporate at the time of the EAG report. It also incorporated the correction for missing data that the EAG included in EAG scenario 8, which the company has adopted in the company's TE base case.

Exploratory analyses 1 to 3

The EAG's TE base case differs from the company's TE base case in three ways explored individually in TE-EA1 to TE-EA3 using the company's TE base case as the starting point (see Table 6). These three changes are as follows:

- The EAG has maintained its preference for using the utility data from Bewley *et al.*[10] instead of the utility data from the vignette study (TE-EA1)
- The EAG has maintained its preference for using the average transition matrix from the second year to estimate the 3-month transition matrix applied from 2 years onwards instead of assuming no further change in SALT scores from 2 years (TE-EA2)
- Although the EAG's base case uses the updated survival analysis for treatment discontinuation, the EAG has selected the exponential rather than the Weibull distribution preferred by the company (TE-EA3).

EAG TE base case

The three changes explored in TE-EA1 to TE-EA3 are combined to produce the EAG's TE base case scenario. Results are then presented, in Table 6, for this base case for the whole cohort (≥12 years), when using the company's approach, which uses average baseline characteristics and pooled efficacy data across both age subgroups (12-18 years and ≥18 years). Results are then presented, in Table 6, for each age subgroup. In Table 7 the EAG's base case result for the whole cohort is presented using the EAG's preferred approach, which uses a weighted mean of the outcomes for the individual age subgroups.

EAG TE scenario analyses

EAG TE scenario analyses are then provided, in Table 7, using the EAG preferred base case with weighted average outcomes across the age subgroups, as the starting point. These scenarios explore the impact of using alternative extrapolation curves for treatment discontinuation (Gompertz and lognormal), removing the caregiver disutility and using the last transition matrix (months 21 to 24) to inform all post-24 months transitions for ritlecitinib. EAG TE scenario analyses for the adolescent subgroup are provided in

Table 8, exploring the alternative disutility assumptions for carers of patients with SALT scores <50, again using the EAG's preferred base case assumptions as the starting point. Finally, results are presented in

Table 9 for the AT/AU and non-AT/AU subgroups, using the EAG's preferred base case assumptions. This includes a scenario analysis exploring the potential impact of the apparent data discrepancy between the number of patients in the SALT 50-100 and SALT 21-49 when applying either the final only or interim and final stopping rules. In this analysis the number of patients with AT/AU at week 48 in the ritlecitinib arm is corrected from [redacted] and [redacted], for SALT 50-100 and SALT 21-49 respectively, to [redacted] and [redacted]. The analyses for the AT/AU and non-AT/AU subgroups do not use a weighted average approach across age cohorts as the model is only populated for ages ≥12 years for these subgroups. As the company has pointed out that the AT/AU subgroup is over-represented in the clinical trials, the EAG has presented results for the whole cohort using a weighted average across the AT/AU and non-AT/AU subgroups using both the proportion who are AT/AU in ALLEGRO 2b/3 and the company's assessment of what this proportion would be in clinical practice.

The EAG has not explored the impact of applying more or less restrictive stopping rules, as the updated discontinuation analysis was only provided when assuming a SALT score of ≤20 at 48 weeks is used to determine treatment response. This is a limitation as the ICER was sensitive to the choice of stopping rule for the previous EAG base case.

EAG TE base case using probabilistic sensitivity analysis (PSA)

The EAG has run the PSA for the EAG's TE base case. This has been done firstly for the whole cohort, using the company's preferred approach of using average baseline characteristics and pooled efficacy data. The EAG has then run the PSA for each age subgroup and has used this to calculate the ICER for the whole cohort when using the weighted average approach preferred by the EAG. The results based on average outcomes across 10,000 PSA samples are provided in

Table 10. The cost-effectiveness acceptability curves (CEACs) and scatterplots on the cost-effectiveness plane for the whole cohort when using the company's preferred approach are provided in Figures 5 and 6. The CEACs and scatterplots for each age subgroup are also provided in Figure 7 to Figure 10. CEACs and scatterplots for the EAG's preferred approach of using a weighted average across age-subgroups are not possible as the parameter samples are not consistent when running the model multiple times and therefore the individual PSA samples for the two age cohorts cannot be combined.

6. Results of the EAG's TE exploratory analyses

The results in Table 6 show that the key driver of the difference in the ICER between the EAG's TE base case and the company's TE base case is the choice of utilities, as implementing this alone increases the ICER from £14,290 to £41,199 per QALY. The long-term extrapolation of treatment efficacy has a smaller, but still important impact, increasing the ICER from £14,290 to £16,980 per QALY. Conversely, the cost-effectiveness does not seem particularly sensitive to the choice of survival curve used to model time to treatment discontinuation as shown in the scenario analysis presented in Table 7, in which varying the survival curve from a Gompertz to a lognormal had a minimal impact on the ICER, when using the EAG's TE base case as the starting point.

The EAG notes that the ICERs for the EAG's TE base case are higher for the adult subgroup (£50,203 per QALY) compared with the adolescent subgroup (£44,269 per QALY). This is related to both the different efficacy data applied and the lack of carer disutility for adult patients. The difference in ICERs between the EAG's preferred approach of using the weighted mean outcomes across the two age subgroups (£48,987 per QALY) and the company's preferred approach (£47,812 per QALY), is less marked for the EAG's TE base case than it was for the previous EAG base case. The cost-effectiveness estimates appear to be very sensitive to whether the patient starts treatment with AT/AU (£59,616 per QALY) or without AT/AU (£42,557 per QALY). The EAG's scenario analysis to explore the apparent data discrepancy in the model for the AT/AU subgroup shows that this is not associated with significant decision uncertainty.

The model appears to remain relatively sensitive to the choice of long-term extrapolation as repeated application of the final transition matrix resulted in an increased in the ICER from £48,987 to £53,593 per QALY. Restricting the long-term transition matrix to data from patients who received the licensed dose throughout also increased the ICER to £55,711 per QALY. The cost-effectiveness in the adolescent subgroup also appears to be relatively sensitive to the choice of caregiver disutility estimates, with ICERs ranging from the base case approach of £43,269 per QALY to £49,790 per QALY when excluding caregiver disutility. The latter brings the ICER for adolescents close to that reported for adults in whom a caregiver disutility was not applied. The company's various scenarios including caregiver

disutility but extending it to health states with a SALT score <50 fall within this range, suggesting that the EAG's base case, which adopts none of these scenarios may overestimate the QALY gains due to carer disutility.

The PSA results for each age cohort are within 3% of the estimates from the deterministic ICERs, suggesting that the deterministic model provides a reasonable approximation to the probabilistic ICER. For this reason the various scenario analyses were not repeated using the PSA. The EAG believes that the previous discrepancy between the deterministic and probabilistic ICERs for the EAG base case at the time of the EAG report was being driven by the inclusion of spontaneous remission in the model for the BSC arm but not the ritlecitinib arm, and the company's decision to remove this from both arms of the model has resolved this issue. The CEACs (Figures 6, 8 and 10) show that the probability of the ICER falling below £30,000 per QALY in the EAG's TE base case scenario is close to zero, both for the cohort as a whole and for either age subgroup.

Table 6: Results of the EAG's exploratory analyses for the whole cohort (≥ 12 years) and the age subgroups (12-18 years and ≥ 18 years) [a]

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
Company's original base case					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,179
EAG's original preferred analysis					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£60,735
Company's TE base case					
BSC	[redacted]	[redacted]	-	-	
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£14,290
TE-EA 1: Company's TE base case with utilities from Bewley *et al.*[10]					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£41,199
TE-EA 2: Company's TE base case using pooled counts from the second year to estimate the 3-month transition matrix applied from 2 years onwards					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£16,980
TE-EA 3: Company's TE base case with exponential discontinuation curve					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£14,358
TE-EA base case (but using average baseline characteristics and pooled efficacy data across whole cohort) [*]: TE-EA 1 to 3 combined					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£47,812
EAG base case for subgroups by age					
TE-EA base case for age 12-18 years subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£43,269
TE-EA base case for age ≥ 18 years subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£50,203

[a] Deterministic and for the whole population covered by the decision problem (aged ≥ 12 years) unless otherwise stated
[*]This is for the whole cohort using the company's preferred approach of using average baseline characteristics and efficacy data pooled across the whole cohort.- the EAG prefers the weighted average approach shown in Table 7.

Table 7: EAG scenario analysis for whole cohort using the weighted average across both age subgroups [a]

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
TE-EA base case: Whole cohort (age ≥ 12 years) using weighted average					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£48,987
TE-EA scenario 1: TE-EA base case with Gompertz time to discontinuation curve					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£49,551
TE-EA scenario 2: TE-EA base case with lognormal time to discontinuation curve					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£48,412
TE-EA scenario 3: TE-EA base case with no carer disutility					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£50,138
TE-EA scenario 4: TE-EA base case with the transition matrix from month 21-24 applied to transitions beyond 2 years for patients on ritlecitinib					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£53,593
TE-EA scenario 5: TE-EA base case with transition matrix restricted to those who received licensed 50mg dose throughout					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£55,711

[a] *Deterministic and using the average weighted outcomes across adult and adolescent subgroups*

Table 8: EAG scenario analysis for the adolescent subgroup (age 12-18 years) exploring alternative assumptions for caregiver disutility

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
TE-EA base case: age 12-18 years subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£43,269
TE-EA scenario 6: TE-EA base case for age 12-18 but with company scenario approach A for caregiver disutility					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£43,701
TE-EA scenario 7: TE-EA base case for age 12-18 but with company scenario approach B for caregiver disutility					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£45,404
TE-EA scenario 8: TE-EA base case for age 12-18 but with company scenario approach C for caregiver disutility					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£45,289
TE-EA scenario 9: TE-EA base case for age 12-18 but with no caregiver disutility					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£49,790

[a] *Deterministic unless otherwise stated*

Table 9: EAG base case and scenario results for subgroups with and without AT/AU [a]

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
TE-EA base case for AT/AU subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£59,616
TE-EA scenario 10: TE-EA base case for AT/AU subgroup with data for AT/AU matching data for final stopping rule					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£59,356
TE-EA base case for non-AT/AU subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£42,557
TE-EA base case for whole cohort using weighted average approach across the AT/AU and non-AT/AU subgroups and the proportion with AT/AU from the ALLEGRO 2b/3 trial					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£47,876
TE-EA base case for whole cohort using weighted average approach across the AT/AU and non-AT/AU subgroups and the proportion with AT/AU estimated by the company (9.52%)					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£43,461

[a] *Deterministic unless otherwise stated; does not use weighted average across age subgroups approach as this is not available for the AT/AU and non-AT/AU subgroups*

Table 10 EAG TE base case results when using probabilistic sensitivity analysis [a]

Option	QALYs	Costs	Incremental		ICER
			QALYs	Costs	
TE-EA base case: Whole cohort (age ≥ 12 years) using average baseline characteristics and pooled efficacy data across whole cohort					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£48,715
TE-EA base case: Age 12-18 years subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£44,073
TE-EA base case: Age ≥ 18 years subgroup					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£51,415
TE-EA base case: Whole cohort (age ≥ 12 years) using weighted average					
BSC	[redacted]	[redacted]			
Ritlecitinib	[redacted]	[redacted]	[redacted]	[redacted]	£50,123

Figure 5

Figure 6

Figure 7

Figure 8 [redacted]

Figure 9 [redacted]

Figure 10 [redacted]

7. Discussion

The EAG considers that there remains significant uncertainty regarding the long-term extrapolation of treatment efficacy and the likely rate of discontinuation in the long-term. Whilst the company has provided data from a later cut of ALLEGRO-LT, which it says is relatively complete with only one patient still to reach the 24-month follow-up out of the cohort still enrolled in the study, there was a high proportion of missing data. Similarly, the updated analysis of discontinuation events was limited by a high degree of censoring before 1.4 years but few details were provided by the company on the reasons for and timing of the censoring events. The TE Response from the BAD[19] suggests that suggest real-world registry/practice data may also be useful to examine treatment waning in the future. The BAD notes that a national AA safety and effectiveness register is currently being set up in the UK and the EAG considers that this registry would provide additional information on the rate of discontinuation of treatment and whether patients discontinue due to loss of response in the long-term.

The EAG's attempt to understand the company's evidence has been hampered by a lack of detail in the company's TE response, particularly with regard to the reasons for missing data in the new data cut from ALLEGRO-LT and the timing of censoring events in the company's discontinuation analysis, which appears to be based on the previously presented data cut. The company is arguing in response to issue 3 that treatment response is maintained in the long-term, and it is arguing in response to issue 6 that there is a low rate of treatment discontinuation in patients who have maintained their treatment response. The EAG does not understand how both these can be true when a high proportion of patients are reported as having missing data in the long-term follow-up of ALLEGRO-LT and there is a high degree of censoring in the discontinuation analysis. If patients withdrew from ALLEGRO-LT due to lack of response, then this contradicts the company's position on maintenance of treatment effect. If they withdrew from the study for other reasons, despite maintaining a good treatment response, then this should be reflected in a high level of treatment discontinuation in the survival analysis. It may be that there is another explanation that allows both the company's assertions to be correct, but the EAG's ability to provide critique on both these issues has been hampered by the company failing to provide sufficient information in both cases.

The EAG notes the company's point that the average cost-effectiveness across the whole trial population may be biased in favour of BSC due to the over representative of patients with AT/AU within the trial population. However, the company has not provided a similar commentary on whether adolescents were over or under sampled in the ALLEGRO 2b/3 trial. The EAG notes its previous conclusion that the average cost-effectiveness across the whole population covered by the licensed indication is dependent on whether the age distribution in clinical practice is similar to that observed in

the ALLEGRO 2b/3 trial. This is likely to remain true even if the efficacy data applied to both groups was to be taken from the whole cohort because caregiver disutility is only applied to adolescents.

The ICER based on the EAG's preferred data and assumptions following the company's response to TE is £48,987 per QALY. However, the average ICER across the whole cohort covered by the licensed indication may be between £43,000 and £50,000 per QALY depending on the uptake of treatment in different age groups. In addition, there is the potential for the ICER to be higher due to uncertainty regarding the best approach to extrapolate long-term effectiveness. The EAG was also unable to explore applying more restrictive stopping rules using the updated model and this was found to have a significant impact on the ICER previously.

The choice of utility values remains a major difference between the company's TE base case and the EAG's TE base case. However, the EAG does not consider that that the company has provided any substantive new evidence that would justify a change to the EAG's previous position that the published estimates of EQ-5D by physician-rated severity from Bewley *et al.* are preferable to the TTO utilities obtained from the company's vignette study.

References

1. National Institute for Health & Care Excellence. Single Technology Appraisal - Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007]. Technical engagement response form. (Company response) - Pfizer. In; 2023.
2. Pfizer. Ritlecitinib for treating moderate to severe alopecia areata in people 12 years and over. Company evidence submission summary to NICE. *Document B* 2023.
3. Ara R, Brazier J. Populating an economic model with health state utility values: moving toward better practice. *Value in Health* 2010;**13**:509-18.
4. Mesinkovska N, King B, Mirmirani P, Ko J, Cassella J. Burden of Illness in Alopecia Areata: A Cross-Sectional Online Survey Study. *Journal of Investigative Dermatology Symposium Proceedings* 2020;**20**:S62-S8. https://doi.org/10.1016/j.jisp.2020.05.007
10.1016/j.jisp.2020.05.007.</p>
5. Pfizer data on f. Clinical study report. Study B7981048: alopecia areata benefit-risk trade-off study. 2022.
6. Pfizer data on f. Vignette study for utility estimation in Alopecia Areata. In; 2022.
7. Burge R, Anderson P, Austin J, et al. The patient-reported burden of alopecia areata by current severity: a real-world study in the United States [Poster 26158]. American Academy of Dermatology (AAD); 2021/03/19/, abstract no. 73.
8. Lloyd A, Aggio D, Law EH, Price T. *Does the EQ-5D measure the full impact of alopecia areata on patients' quality of life? Running title: EQ-5D measurement in patients with alopecia areata*: Pfizer; 2023.
9. Sellami R, Masmoudi J, Ouali U, Mnif L, Amouri M, Turki H*, et al.* The relationship between alopecia areata and alexithymia, anxiety and depression: a case-control study. *Indian J Dermatol* 2014;**59**:421. https://doi.org/10.4103/0019-5154.135525
10. Bewley A, Galvan SV, Johansson E, Durand F, Petto H. Measuring the Burden of Alopecia Areata with the European Quality of Life-5 Dimensions (EQ-5D): Results from a real-world survey in 5 European countries. *Value in Health*;**25**:S428-9.
11. 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. https://doi.org/10.1111/1346-8138.16360
12. Pfizer data on f. Qualitative research in alopecia. Targeted literature review and patient advocacy group representative interviews in the United Kingdom. 2022.
13. Rencz F, Mukuria C, Bató A, Poór AK, Finch AP. A qualitative investigation of the relevance of skin irritation and self-confidence bolt-ons and their conceptual overlap with the EQ-5D in patients with psoriasis. *Qual Life Res* 2022;**31**:3049-60. https://doi.org/10.1007/s11136-022-03141-y
14. Swinburn P, Lloyd A, Boye KS, Edson-Heredia E, Bowman L, Janssen B. Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 2013;**16**:1156-62. https://doi.org/10.1016/j.jval.2013.10.003
15. Hoogendoorn M, Oppe M, Boland MRS, Goossens LMA, Stolk EA, Rutten-van Mölken MPMH. Exploring the Impact of Adding a Respiratory Dimension to the EQ-5D-5L. *Med Decis Making* 2019;**39**:393-404. https://doi.org/10.1177/0272989X19847983
16. Yang Y, Rowen D, Brazier J, Tsuchiya A, Young T, Longworth L. An exploratory study to test the impact on three "bolt-on" items to the EQ-5D. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 2015;**18**:52-60. https://doi.org/10.1016/j.jval.2014.09.004
17. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *Eur J Health Econ* 2015;**16**:927-39. https://doi.org/10.1007/s10198-014-0638-9

18. NICE. 2021. Barcitinib for treating moderate to severe atopic dermatitis: technology appraisal guidance. Available at: https://www.nice.org.uk/guidance/ta681/resources/baricitinib-for-treating-moderate-to-severe-atopic-dermatitis-pdf-82609375014853. In.
19. National Institute for Health & Care Excellence. Single Technology Appraisal - Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007] - Technical engagement response form (Stakeholder Response) British Association of Dermatologists. In; 2023.
20. National Institute for Health and Care Excellence. Single Technology Appraisal - Ritlecitinib for treating severe alopecia areata in people 12 years and over [ID4007] - Technical engagement response form (Stakeholder Response) Alopecia UK. In; 2023.
21. Reid EE, Haley AC, Borovicka JH, Rademaker A, West DP, Colavincenzo M*, et al.* Clinical severity does not reliably predict quality of life in women with alopecia areata, telogen effluvium, or androgenic alopecia. *Journal of the American Academy of Dermatology* 2012;**66**:e97-102. https://doi.org/10.1016/j.jaad.2010.11.042
22. NICE 2023. Final draft guidance: Baricitinib for treating severe alopecia areata. 2023. Available at: https://www.nice.org.uk/guidance/gid-ta10941/documents/final-appraisal-determination-document. In.