

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Final guidance consultation

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

1 Recommendations

- 1.1 Ritlecitinib is recommended, within its marketing authorisation, as an option for treating severe alopecia areata in people 12 years and over.

Why the committee made these recommendations

There is no standard treatment for severe alopecia areata, and access to treatment varies widely. Hair loss can cause severe psychological distress.

Evidence from clinical trials shows that ritlecitinib is more effective than placebo at improving hair regrowth for up to 24 weeks.

The most likely cost-effectiveness estimates are within the range that NICE considers an acceptable use of NHS resources. So, ritlecitinib is recommended.

2 Information about ritlecitinib

Marketing authorisation indication

- 2.1 Ritlecitinib (Litfulo, Pfizer) is indicated for 'the treatment of severe alopecia areata in adults and adolescents 12 years of age and older'.

Dosage in the marketing authorisation

- 2.2 The dosage schedule is available in the [summary of product characteristics for ritlecitinib](#).

Price

- 2.3 Ritlecitinib costs £949.41 per pack of 30 capsules (dictionary of medicines and devices, accessed February 2024).
- 2.4 The company has a commercial arrangement (simple discount patient access scheme). This makes ritlecitinib available to the NHS with a discount. The size of the discount is commercial in confidence. It is the company's responsibility to let relevant NHS organisations know details of the discount.

3 Committee discussion

The [evaluation committee](#) considered evidence submitted by Pfizer, a review of this submission by the external assessment group (EAG), and responses from stakeholders. See the [committee papers](#) for full details of the evidence.

The condition

Effects on quality of life

- 3.1 The patient experts explained that living with severe alopecia areata has a profound impact on psychosocial health. They described the devastating impact of severe alopecia areata which can lead to depression, anxiety, social isolation and suicidal thoughts. The patient experts also explained that the condition can put immense stress on intimate relationships. They said that it can lead to social exclusion and can limit career progression or education because of an inability to fully participate in society. They further explained that this impact is also felt by their families who may provide care and emotional support. They emphasised that alopecia areata is much more than a cosmetic issue. They said that as well as the severe psychosocial impact, the lack of hair on parts of the body other than the scalp affects physiological health. This includes a lack of:
- eyelashes and eyebrows, which can lead to problems with sweat and grit getting into the eyes

- nasal hair to prevent mucus leaving the nose
- hair on skin, which impacts temperature regulation.

Consultation comments stated that young adults with severe alopecia areata have left education or work because of severe mental health challenges. Some people have suicidal ideation, and some people have died by suicide as a result of having severe alopecia areata. It was emphasised there was a need to understand that quality of life is improved with hair regrowth, with people describing 'getting their life back'. The committee concluded that severe alopecia areata has wide-ranging effects and can have a profound impact on quality of life.

Clinical management

Treatment options

- 3.2 There are no licensed treatments available on the NHS for severe alopecia areata. The clinical experts explained that there are some pharmacological treatment options available in secondary and tertiary care. These include topical corticosteroids and contact immunotherapy, and for those with more severe hair loss, systemic corticosteroids or immunosuppressants. But they said that none of these options are satisfactory. They explained that contact immunotherapy is only offered in some centres in England and Wales, that it requires weekly clinic attendance, and only targets scalp hair regrowth. The clinical experts further explained that systemic treatments can have side effects and need additional monitoring. The patient experts said that many people with the condition do not have any treatment, with referrals in some areas not being accepted for alopecia areata. The clinical experts explained that the inconsistent availability of treatments across England and Wales is in part because they are not licensed for alopecia areata, so not all clinics are willing to prescribe them. Non-pharmacological management of alopecia areata includes using wigs. The patient experts explained that the availability of wigs varies regionally and that those offered by the NHS are

often unsuitable. Because of this, people with alopecia areata often spend their own money on wigs and other appearance-altering treatments such as microblading. The patient and clinical experts agreed that there is no standard treatment pathway for alopecia areata and that the treatment options are very limited. They explained there is a high unmet need for a targeted treatment for severe alopecia areata. The committee concluded that there is no standard care for severe alopecia areata, that available treatments are not available equitably across England and Wales, and that there is an unmet need for new treatments.

Ritlecitinib

3.3 Ritlecitinib is a JAK inhibitor which downregulates the immune response at the hair follicles. Another JAK inhibitor, baricitinib, is licensed for severe alopecia areata in Great Britain but is not available for severe alopecia areata on the NHS. So, if recommended, ritlecitinib would be the first treatment available on the NHS with this mechanism of action that was licensed for severe alopecia areata. The patient experts explained that people want a licensed treatment that is specifically targeted at alopecia areata to be available on the NHS. The committee concluded that ritlecitinib is an innovative medicine and that JAK inhibitors provide a new mechanism of action for treating severe alopecia areata.

Severity of Alopecia Tool

3.4 The company rated the severity of alopecia areata according to the Severity of Alopecia Tool (SALT). The SALT assesses the proportion of scalp surface area affected by hair loss. Using this tool, 0% scalp hair loss is represented by a SALT score of 0, and 100% scalp hair loss is represented by a SALT score of 100. The company defined severe alopecia areata as a SALT score of 50 or more. The patient experts explained that the SALT only measures hair loss and regrowth on the scalp, and that hair on other areas of the body is also important to consider (see [section 3.1](#)). The clinical experts explained that in their

experience, people who had ritlecitinib and had an improved SALT score also had improved hair growth on other areas of the body. The company said that clinical trial results showed that no one had eyebrow or eyelash regrowth without also having a SALT score improvement. The SALT score can be used as an absolute measure or a relative measure of treatment effect. The company used the absolute measure of a SALT score of 20 or below as a primary outcome in its pivotal clinical trial (see [section 3.7](#)). The clinical experts explained that it is difficult to reach a SALT score of 20 or below in severe alopecia areata. They considered that the relative measure may be more useful for determining treatment effect for people with this condition. But they also explained that SALT score is not routinely used to determine treatment effect in practice and that perception and acceptability of hair regrowth is more important. The patient experts also highlighted that a high relative reduction in SALT score may not be a meaningful outcome if this resulted in patchy hair regrowth. The committee concluded that both absolute and relative SALT scores can be measured in practice. It concluded that although it does not capture all aspects of the severity of alopecia areata, absolute SALT score reduction is an acceptable measure to demonstrate the clinical effectiveness of ritlecitinib and for use in the economic model.

Clinical evidence

Data sources

- 3.5 The main evidence for ritlecitinib was from the ALLEGRO phase 2b/3 trial (ALLEGRO 2b/3) and the ALLEGRO long-term follow-up trial (ALLEGRO-LT). ALLEGRO 2b/3 was a multi-arm mixed methods trial including 2 phases, in people 12 years and over with severe alopecia areata (defined by a SALT score of 50 or more). The first phase was a 24-week randomised controlled trial comparing ritlecitinib with placebo. In the second phase, people in the placebo arms were switched to ritlecitinib and people who had ritlecitinib in the first phase continued treatment, both

for a further 24 weeks. ALLEGRO-LT is an ongoing 36-month open-label follow-up trial that includes:

- people who took part in ALLEGRO 2b/3
- people who took part in the ALLEGRO phase 2a proof-of-concept study (ALLEGRO 2a)
- a de novo population who were newly enrolled.

The committee concluded that the ALLEGRO 2b/3 and the ALLEGRO-LT trials were appropriate to show the treatment effect of ritlecitinib.

Generalisability

3.6 ALLEGRO 2b/3 included young people (12 to 17 years; 14.6%) and adults (85.4%). They either had alopecia totalis or alopecia universalis at baseline (complete scalp hair loss [SALT score 100]; 46.0%), or they did not (some scalp hair [SALT score less than 100]; 54.0%). Similar proportions of each population were included in the ALLEGRO-LT trial de novo population (16.9% young people; 34.5% alopecia totalis or alopecia universalis at baseline). The clinical experts explained that in general, the population included in the ALLEGRO trials represented the people they see in clinical practice. But they stated that the proportion of young people included in the ALLEGRO trials underrepresented the proportion seen in clinical practice. After consultation, during which the company provided an estimate of 4.91% for young people with alopecia areata, the committee heard from the experts that the proportion of young people seen in clinic was more in line with the 14.6% seen in the trial. The clinical experts also noted that the proportion of people with alopecia totalis or alopecia universalis was overrepresented and was closer to 10% in practice. They explained that a response is less likely if a person has alopecia totalis or alopecia universalis than if they do not. The committee concluded that overall, the population in the ALLEGRO trials was mostly generalisable to clinical practice, but that the proportion of young people and adults with alopecia totalis or alopecia universalis was not.

Clinical effectiveness

3.7 The ALLEGRO 2b/3 trial showed that after 24 weeks, the response rate (the percentage of people with a SALT score of 20 or less) was statistically significantly greater for people having a 50 mg dose of ritlecitinib compared with people having placebo (ritlecitinib 50 mg response rate: 23.0%; difference in response rate between ritlecitinib 50 mg and placebo: 21.4%, 95% confidence interval [CI] 13.4 to 29.5). After 48 weeks, the response rate for people having 50 mg ritlecitinib improved further (ritlecitinib 50 mg response rate: 43.2%). The ALLEGRO-LT trial showed that response rates continued to improve for people taking ritlecitinib for up to 2 years. ALLEGRO 2b/3 also showed that more people had eyebrow and eyelash regrowth with ritlecitinib than with placebo after 24 weeks. The clinical experts said that existing treatments for alopecia areata target scalp hair regrowth and that the benefits seen with ritlecitinib in eyebrow and eyelash regrowth were promising. The patient experts highlighted that eyebrow and eyelash regrowth are also important outcomes and that this was a benefit of ritlecitinib over other available treatments. The committee concluded that ritlecitinib is more effective than placebo for clinically meaningful hair growth on both the scalp and other areas of the body.

Subgroups

3.8 ALLEGRO 2b/3 reported a statistically significant difference in response rate between people having a 50 mg dose of ritlecitinib and placebo for the subgroups of:

- young people aged 12 to 17 years
- adults
- people with alopecia totalis or alopecia universalis
- people without alopecia universalis or alopecia totalis.

The clinical experts said that the results for the young people and adult subgroups reflected what was expected in clinical practice. This is

because there is no reason to expect a difference in treatment effect based on age. The EAG noted that the results suggested that there was a lower response rate for people with than without alopecia totalis or alopecia universalis. The clinical experts highlighted that this was because a response is less likely for these types of alopecia areata. They noted that it was impressive that ritlecitinib had been shown to be statistically significantly more effective than placebo in the subgroup of people with alopecia totalis or alopecia universalis. The committee concluded that ritlecitinib is more effective than placebo in the subgroups presented by age and alopecia areata severity.

Long-term treatment effects

3.9 The company presented 2 years of follow-up data from ALLEGRO 2b/3 and ALLEGRO-LT for people taking ritlecitinib (see [section 3.7](#)). It noted that there is no evidence available for the effectiveness of ritlecitinib beyond this, but that 36-month follow-up data from ALLEGRO-LT may be available in the future. The clinical experts explained that alopecia areata is a chronic disease and that people may want to use ritlecitinib long term. But they noted that, based on their experience with systemic treatments, people may discuss stopping treatment once they feel they have satisfactory hair regrowth. With systemic treatments, this is often after 2 or more years of successful response and the clinical experts expected that this could be similar with ritlecitinib. They explained that there are various reasons for someone wanting to stop taking ritlecitinib, such as family planning or side effects. They noted that side effects associated with ritlecitinib included acne in young people, and respiratory tract infections. The company explained that serious adverse event rates were similar in the ritlecitinib and placebo groups in ALLEGRO 2b/3. They also explained that there is no evidence available from ALLEGRO 2b/3 or ALLEGRO-LT to show what happens to hair growth when ritlecitinib is stopped. The EAG said that data on this may be available from the ALLEGRO 2a proof-of-concept study. The clinical experts suggested that long-term data from

registries may also be able to answer this in the future but that it is difficult to predict based on the evidence available. The committee concluded that people taking ritlecitinib would likely stop treatment rather than taking it indefinitely. It also concluded that it was uncertain what the effect of stopping treatment would be, but that any evidence available to inform this would be useful for decision making.

Economic model

Company's model structure

3.10 The company's model had 9 states: 4 for on-treatment, 4 for best supportive care, and death. The 4 on-treatment and best supportive care health states were defined by SALT score, ranging from 10 or below, to 50 or more. All people entered the model with a SALT score of 50 or more. Stopping ritlecitinib treatment was assumed if the SALT score worsened after 24 weeks on treatment, or if it was more than 20 at 48 weeks or at any point after this. After stopping, a transition to the equivalent best supportive care health state was assumed. This was followed by (if applicable) a transition to the SALT score 50 or more best supportive care health state by moving to a worse health state every cycle, informed by the transitions observed from the trial data. The committee concluded that although using SALT score to define health states did not capture all aspects of alopecia areata (see [section 3.4](#)), the model was acceptable for decision making.

Best supportive care

3.11 The company's original submission stated that best supportive care should only include non-pharmacological treatments. The EAG explained that the on-treatment and best supportive care health states would both include the use of wigs, psychological support and dermatology and GP visits. But these were used at varying rates across health states and arms. It noted that there was no pharmacological treatment included in the model for the best supportive care health states, whether that was as the

comparator arm from the first cycle or after stopping ritlecitinib. The clinical experts highlighted that there are some pharmacological treatments available for alopecia areata which might make up best supportive care, but these are used inconsistently across England and Wales (see [section 3.2](#)). The committee also heard that these treatments were not supported by a recent Cochrane review. The EAG and the company also explained that there is limited evidence available to estimate the effectiveness of these unlicensed treatments, and that what evidence is available is contradictory and low quality. The clinical experts explained that the decision to offer another pharmacological treatment after ritlecitinib would be made on a case-by-case basis. This would be based on treatment history and discussion with the person with severe alopecia areata. The patient experts highlighted that there is no treatment pathway for alopecia areata (see [section 3.2](#)).

- 3.12 In response to consultation, the company noted that pharmacological treatments were accepted as part of best supportive care in the [NICE technology appraisal guidance for baricitinib for treating severe alopecia areata](#) (TA926) and that not including it in this appraisal could be considered a conservative assumption. Other stakeholders also suggested that pharmacological treatments should be included. The company did not change its assumption of only non-pharmacological treatments in its base case, but it presented scenarios including ‘baskets’ of pharmacological treatments. These scenarios assumed that 88% of people (based on an Adelphi Disease Specific Programme) or 87% of people (based on UK key opinion leader data) who received best supportive care would also receive pharmacological treatment. Some of the scenarios assumed the same use of pharmacological treatments in both arms (ritlecitinib and placebo), and some assumed less use in the ritlecitinib arm. The EAG explained that if a high proportion of people have pharmacological treatments, then ALLEGRO placebo data is not appropriate for estimating expected costs and benefits in best supportive care. In addition, the composition of the baskets was very uncertain. For

Final draft guidance consultation – Ritlecitinib for treating severe alopecia areata in people 12 years and over

example, data from the ADAAGIO study suggested a very different composition to the company's baskets. The EAG also thought that the assumption of a 10-year treatment duration for best supportive care was too long. The clinical and patient experts said that many of the treatments can only be given for short periods, and because of the inconsistent availability of treatments across England and Wales, many people did not receive any (see [section 3.2](#)). The EAG highlighted that in the TA926, only the same use of pharmacological treatments in both arms had been accepted. The committee noted that the scenarios presented only included the cost of the pharmacological treatments. None of the administration costs, clinical-effectiveness data, or adverse events associated with these pharmacological treatments had been included. The committee agreed that in ALLEGRO, in which people had severe alopecia areata for a median of 7 years, people would have been beyond pharmacological treatment options at this point. So it did not have the evidence to consider pharmacological treatments as a comparator. The committee concluded it was acceptable to include only non-pharmacological treatment options in the best supportive care health states.

Utilities

The company's vignettes

- 3.13 The company did a vignette and time-trade-off study to estimate utility values for people with alopecia areata and their carers. The company developed vignettes that described the impact of having alopecia areata with a specified SALT score, aligned with the health states in the model. The SALT score ranges were: 10 or less; 11 to 20; 21 to 49; and 50 or more. It also developed a vignette describing the impact of caring for a young person with severe alopecia areata. It used a time-trade-off approach using the vignettes to estimate the utility values for each health state in the model, as well as the carer disutility associated with caring for someone with severe alopecia areata. The EAG explained that the

company had followed best practice methods to develop the vignettes, but it had 2 concerns. Firstly, the vignettes only described the negative nature of the health state and did not include information on the aspects of life that were unaffected, for example mobility. It suggested that this may have biased the time-trade-off exercise, leading to the negative impact of the condition being overestimated. The EAG was also concerned about the face validity (clinical plausibility) of the vignettes compared with the results from the Alopecia Areata Patient Priority Outcomes (AAPPO) results in ALLEGRO 2b/3. The company explained that the differences between the AAPPO results and the vignettes was because the vignettes were developed based on a variety of sources, making them less subject to bias than the AAPPO results alone. The company also noted that ALLEGRO 2b/3 excluded people with suicidal thoughts or depression and so the data may underrepresent the impact of severe alopecia areata. The patient experts said that the utility values that were generated from the vignette study for the most severe health state were clinically plausible. They emphasised the severe psychosocial impact that alopecia areata has on people. Drawing on personal experience, 1 patient expert said that the effect of severe alopecia areata on their quality of life had been greater than recovery from a brain haemorrhage. The patient and clinical experts suggested that for some people with suicidal thoughts, their utility value could be as low as that estimated for the most severely affected health state from the vignette study. The clinical experts said that for the average person with severe alopecia areata the true utility values might be higher than suggested by the vignette study, although it was highly uncertain and difficult to estimate.

- 3.14 In response to consultation, the company presented an extension vignette study of alopecia areata and a proxy review study of atopic dermatitis, the results of which supported the original vignette study. The EAG explained that the same methodology was used in the extension study and the original study, so the concerns were the same. In addition, the company used an unvalidated video conference method, adding to the uncertainty

Final draft guidance consultation – Ritlecitinib for treating severe alopecia areata in people 12 years and over

about the study's validity. The EAG did not critique the proxy study, because it felt that suitable EQ-5D data was available in the literature. The committee noted that the utility values represented the quality of life for the average person in each health state. It felt that the utility values could represent some people with severe alopecia areata, but that for the average person, the utility values estimated from the vignette study were likely to be too low. The committee concluded that the company had mostly followed best practice when doing the vignette study, but that concerns remained about the validity of the results.

EQ-5D utilities from the trials

3.15 ALLEGRO 2b/3 collected health-related quality-of-life data using multiple measures, including:

- EQ-5D-5L
- EQ-5D-Y (for young people)
- EQ Visual Analogue Scale (VAS)
- the AAPPO tool
- short form-36 (SF-36)
- Hospital Anxiety and Depression Scale.

[Section 4.3 of NICE's health technology evaluations: the manual \(2022\)](#)

states that the EQ-5D should be used to generate utility values, and if these are not available from the trials they can be sourced from literature. It also states that to make the case that the EQ-5D is inappropriate, qualitative empirical evidence on lack of content validity (whether a test actually measures all the areas it should measure) should be presented. Alongside this there should be evidence that the EQ-5D performs poorly on tests of validity and responsiveness, sourced from a synthesis of peer-reviewed literature. If, based on this evidence, the committee is satisfied that the EQ-5D is not appropriate, then the following sources of utility values can be used, in order of preference:

- another generic preference-based measure
- a condition-specific preference-based measure
- vignettes
- direct valuation of own health.

The company argued that the EQ-5D lacks content validity for people with severe alopecia areata. This is because it contains no domains on social functioning, relationships, emotional impact, physical appearance or financial impact. So it said that it was inappropriate to use the EQ-5D data collected in ALLEGRO 2b/3 to estimate utility values for people with severe alopecia areata. The company also noted that the EQ-5D data collected in ALLEGRO 2b/3 had further issues, including a ceiling effect caused by high baseline scores and a relatively short 24-week placebo-controlled follow-up period. It noted that these both made any improvement in health-related quality of life difficult to measure. The company explained that the average time since diagnosis in ALLEGRO 2b/3 was 10 years and that this may have led to high levels of adaptation. It said that this may have been reflected in the high baseline scores. The company also explained that people with major psychiatric conditions were excluded from ALLEGRO 2b/3 and these were the people who were most likely to have the biggest improvement in health-related quality of life from ritlecitinib. The patient experts explained that people with alopecia totalis or alopecia universalis try to convince themselves and others that they are well because they are often mistaken for people having chemotherapy. This may have led to the high baseline scores seen in the ALLEGRO 2b/3 EQ-5D results. The patient and clinical experts agreed that EQ-5D data from the ALLEGRO 2b/3 trial is unlikely to capture the severity of the condition. The EAG agreed that using the EQ-5D results from ALLEGRO 2b/3 was unlikely to be appropriate, because of selection bias, many years of exposure to severe alopecia resulting in adjustment, high baseline scores and the short follow-up period of the trial.

3.16 In response to consultation, the company presented longer-term EQ-5D data from the ALLEGRO-LT trial as requested. It did not present a scenario analysis using this data because it did not consider this to be appropriate, because the ALLEGRO-LT data had similar issues to the ALLEGRO 2b/3 data. The EAG agreed that using the EQ-5D results from ALLEGRO trials was unlikely to be appropriate, because of the exclusion of people with psychiatric comorbidities, the long average duration since diagnosis, high baseline scores and the short follow-up period of the trial. But for completeness the EAG provided the requested scenario analyses using the ALLEGRO-LT EQ-5D data. The committee agreed with the EAG and concluded that given the data limitations, the resulting incremental cost-effectiveness ratio (ICER) from the scenario analysis using the ALLEGRO-LT EQ-5D data needed to be interpreted with caution. But it agreed that it was helpful for understanding the uncertainty around the EQ-5D data used in the model.

Other utility sources

3.17 At the first committee meeting, the company argued that EQ-5D data from the literature is not appropriate, for many of the same reasons it argued that the EQ-5D data from the ALLEGRO trials is inappropriate, such as content validity (see [section 3.15](#)). The clinical and patient experts also said that EQ-5D data from any source would be unlikely to detect a change in health-related quality of life in people with severe alopecia areata. This is because it does not adequately cover aspects important to these people, such as the psychosocial impact of the condition (see [section 3.1](#)). The EAG disagreed that the EQ-5D as a measure is inappropriate for showing changes in treatment effect for people with severe alopecia areata. It highlighted that some aspects of severe alopecia areata are captured by the EQ-5D, in the anxiety and depression and usual activities domains. It presented evidence from the Adelphi real-world evidence database (Bewley et al. 2022) that indicated that the EQ-5D is sensitive to varying alopecia areata severity in a European

population. Based on the methods outlined in the NICE health technology evaluation manual (see [section 3.15](#)), the EAG preferred to use this data to estimate utility values for each health state. It mapped the mild, moderate and severe disease described in Bewley et al. to the SALT score-based health states in the model. The company argued that the mild, moderate and severe disease states in Bewley et al. were graded based on clinician judgement, so were subject to bias. It said that it was not appropriate to use any of the other health-related quality-of-life measures used in the ALLEGRO trials to estimate utility values in the model (see [section 3.15](#)). In response to consultation, the company also presented a post hoc psychometric evaluation of ALLEGRO-2b/3 EQ-5D and SF-36 data which showed that the EQ-5D and SF-36 measures did not respond to a change in health-related quality of life in people with severe alopecia areata. It also explained that no condition-specific preference-based utility measures exist. So the company continued to use the utility values from its vignette study in its preferred base case, and provided scenario results using the vignette in alopecia areata and proxy atopic dermatitis utilities (see [section 3.14](#)). The EAG explained that given the data limitations in the ALLEGRO trials (see [section 3.15](#)), it was not surprising that the psychometric report did not show that the ALLEGRO data was sensitive to changes in health-related quality of life in severe alopecia areata. It noted that no SF-6D data was presented. The EAG reviewed the new evidence presented by the company but did not consider that any of it showed that the EQ-5D data was not suitable for use in alopecia areata. The EAG highlighted that the Adelphi data was sensitive to changes in alopecia areata-related health-related quality of life, and stated the usual activity domain was statistically significantly associated with physician-rated alopecia areata severity. Also, 29% of participants scored 11 or more on the Hospital Anxiety and Depression scale for anxiety and 27% for depression. Utility values from a full-text publication (Vañó-Galván et al.) describing the European cohort from the Adelphi database, previously described only in abstract form, are now

available. The EAG preferred to use the Vañó-Galván et al. utility values in its revised base case. The committee concluded that there was not sufficient evidence that the EQ-5D was an inappropriate measure for evaluating changes in disease severity in alopecia areata. The committee acknowledged the limitations of the Vañó-Galván et al. utility values, noting that the mild, moderate and severe disease populations did not directly map to the health states in the model. But, based on the evidence presented, it concluded that the utility values estimated from the Vañó-Galván et al. study were the most appropriate to include in the model.

Carer utilities

3.18 In its submission, the company included a disutility for carers of both adults and young people with severe alopecia areata. The company estimated utilities for a carer of someone with severe alopecia areata from its vignette study (see [section 3.13](#)). It subtracted the carer utility value from an age-matched general population utility value to estimate the carer disutility that was applied in the model. The EAG highlighted that the utility value for carers was estimated using a vignette that described the impact of caring for a young person with severe alopecia areata and not an adult. So, at technical engagement, the company agreed to only apply a carer disutility to carers of young people. The patient experts highlighted that family members of adults with alopecia areata may provide care and are also affected by the condition (see [section 3.1](#)). The committee accepted that it is plausible that the impact of severe alopecia areata is not limited to the person with the condition but may also affect family members of young people. It concluded that the company's approach was acceptable and made little difference to the cost-effectiveness estimates. So, it concluded that it was appropriate to include disutilities for carers of young people in the model.

Other assumptions

Weighting by alopecia severity

3.19 The company said that the proportion of people with alopecia totalis or alopecia universalis in ALLEGRO 2b/3 (46.0%) was greater than the proportion of people with severe alopecia areata with alopecia totalis or alopecia universalis in clinical practice, which it estimated as 9.52%. The clinical experts agreed with the company and estimated that around 10% of people with severe alopecia areata have alopecia totalis or alopecia universalis (see [section 3.6](#)). The company presented a scenario analysis that weighted the ICER according to the expected distribution of alopecia totalis or alopecia universalis seen in clinical practice. The committee agreed that it was appropriate to consider the weighted ICER in decision making because it was more generalisable to the population in clinical practice.

Weighting by age

3.20 The EAG highlighted that the company's ICERs did not use the weighted average of outcomes for young people and adults but used average baseline characteristics across the full ALLEGRO 2b/3 population. The company said that age did not modify treatment effect, so there was no reason to use weighting for different age groups. The EAG noted that this had a limited impact on the ICER when disutilities for carers of adults with severe alopecia areata were not included in the model. At the first committee meeting, the clinical experts highlighted that the proportion of young people included in ALLEGRO was lower than expected in clinical practice. But they could not reliably estimate what proportion of people in clinical practice were young people. The committee requested real-world evidence to help inform the estimate of the proportion of people with severe alopecia areata who are under 18, if available. It would be appropriate to weight the ICER according to the proportion of young people expected in clinical practice and to weight the ICER according to average outcomes for young people and adults separately, given that the

Final draft guidance consultation – Ritlecitinib for treating severe alopecia areata in people 12 years and over

carer disutility was only applied for young people. At the second committee meeting, the clinical and patient experts stated that the proportion of young people in the trial better reflected the proportion seen in clinical practice than the company's proposed figure (see [section 3.6](#)). The EAG's analyses of the weighted ICER, which used adult-only data for adults and combined adult and young person data for young people, did not differ that much from the unweighted ICER, which used average efficacy across both age groups applied it to both age groups. The EAG highlighted that because of not having the efficacy data cut by both age and alopecia totalis or alopecia universalis status, it was not possible to weight both age and severity at the same time. The EAG's revised base case was weighted for severity and assumed the proportion of adolescents was 14.6%. As this scenario captured the potentially large differences in efficacy between having and not having alopecia totalis or alopecia universalis, and that the weighting by age had a minimal effect on the ICER, the committee concluded that not weighting by age in the base case was acceptable.

Long-term treatment effect

3.21 The company's model assumed that after 96 weeks of ritlecitinib treatment, a person's SALT score remained stable for the full time horizon unless ritlecitinib was stopped. The company said that this was supported by ALLEGRO-LT data for up to 2 years. The clinical experts said that it was unclear what the long-term effects of continued ritlecitinib treatment would be. The EAG preferred to use the average transitions in health states over the final year for which data was available to estimate long-term health-state transitions. In response to consultation, the company stated that the committee's preferred assumption for the time on treatment was conservative. The company reiterated its preference of staying in state until discontinuation. It provided a range of scenarios with time on treatment longer than that used in the committee's preferred approach, based on its stay-in-state effectiveness assumption and different

discontinuation rate curves. The EAG noted that the choice between staying in state or using average transitions, rather than the choice of discontinuation curve (see [section 3.22](#)), had a large effect on ICERs. The EAG noted that no new evidence had been submitted to support the staying-in-state assumption, and because of the limited follow up in the ALLEGRO-LT trial it was hard to verify any long-term effect. It was unclear how the missing data had been dealt with in the company's analysis, whereas the data at 2 years was relatively complete. It was also unclear how the non-responders were accounted for if they stopped treatment. The committee agreed that the company's approach was optimistic. So, it concluded that the EAG's approach to modelling long-term treatment effect was more appropriate.

Stopping treatment

3.22 The length of time people used ritlecitinib in the model was estimated using extrapolated data from ALLEGRO 2b/3. The company used a Weibull model to extrapolate time on treatment. This was based on it being an 'accelerated failure time' model with good Akaike information criterion (AIC) and Bayesian information criterion (BIC) ranking and a good visual fit to the Kaplan–Meier data from ALLEGRO 2b/3. The EAG explained that it was not necessary to use an accelerated failure time model, so it preferred to use an exponential model to extrapolate time on treatment, which had better AIC and BIC ranking. It highlighted that there was very little difference in any of the extrapolation curves presented in terms of the AIC and BIC ranking or the fit to the Kaplan–Meier data. So, it also explored other extrapolations, which showed that the choice of extrapolation curve had a minor impact on the ICER. The clinical experts explained that ritlecitinib is expected to be a long-term treatment for a chronic condition (see [section 3.9](#)). But the committee noted there are reasons that people would choose to stop treatment, so it was highly uncertain how long on average ritlecitinib would be used for. The committee noted that the choice of curve had a minimal effect on the

ICER. It concluded that it was likely that this would remain an uncertainty, but that the EAG's approach to extrapolating time on treatment was a more conservative approach that reflected that people may stop treatment.

Cost effectiveness

Acceptable ICER

3.23 The committee discussed there being no licensed treatments for severe alopecia areata available on the NHS. It also noted that there is a large unmet need for a new treatment that specifically targets the condition (see [section 3.2](#)). It noted that ritlecitinib is innovative in that it has a different mechanism of action to other treatments used in the NHS. Also, unlike other treatments, it targets hair regrowth in areas of the body other than the scalp, which is an important outcome for people with the condition (see [section 3.7](#)). The committee accepted that there were likely to be uncaptured benefits in any measure of health-related quality of life for severe alopecia areata. So, to account for uncaptured benefits and the innovative nature of ritlecitinib, the committee agreed that an acceptable ICER for ritlecitinib for treating severe alopecia areata in people 12 years and over would be towards the higher end of the range usually considered a cost-effective use of NHS resources.

Cost-effectiveness estimates

3.24 The committee's preferred assumptions aligned with the EAG's revised base case and were:

- including only non-pharmacological treatments as part of best supportive care (see [section 3.11](#))
- using utility values for each health state mapped from the mild, moderate and severe disease utility values from Vañó-Galván et al. (see [section 3.17](#))

- including a disutility for carers of young people with severe alopecia areata (see [section 3.18](#))
- weighting the proportion of people with alopecia totalis or alopecia universalis in the model (see [section 3.1](#))
- using the average transitions in health states over the final year for which data was available to estimate long-term treatment effect (see [section 3.21](#))
- using the exponential model to extrapolate time to treatment stopping (see [section 3.22](#)).

This resulted in a deterministic ICER of £25,406 per quality-adjusted life year (QALY) gained. The committee noted the uncertainty around using EQ-5D data for the utilities in the model. Although there was uncertainty around the cost-effectiveness estimates, the committee noted that the extrapolation of time to treatment stopping in the model was potentially conservative. And given that this medicine was likely to have uncaptured benefits and was innovative (see [section 3.23](#)), it concluded that the most likely cost-effectiveness estimate for ritlecitinib compared with best supportive care was within the range that NICE considers a cost-effective use of NHS resources.

Other factors

Equality

3.25 The committee acknowledged that some people with severe alopecia areata may be more affected by the psychological impact of hair loss because of the religious or cultural significance of hair. The clinical and patient experts also explained that severe alopecia areata can have a particularly large impact on psychosocial health and quality of life for young people. Responses to consultation stated that severe alopecia areata is more common in Asian and African groups and that alopecia areata incidence is higher in people with low socioeconomic status, and people from non-White groups. Religion, race and age are protected

characteristics under the Equality Act 2010. They also noted that people with autoimmune skin conditions including alopecia are at higher risk of spontaneous abortions than people without these conditions, and that severe alopecia areata is associated with severe physical disfigurement. Severe physical disfigurement is a protected characteristic under the Equality Act 2010. The committee noted that alopecia areata is a condition of high unmet need and that treatments are not available equitably across England and Wales (see [section 3.2](#)), but that the higher prevalence of the condition in some groups cannot be addressed by a technology appraisal. The committee agreed that there were potential equality issues for this appraisal. But the recommendation applies to all groups covered by the marketing authorisation and will improve access to treatment for alopecia areata in the NHS. So, the committee did not need to amend its recommendation.

Conclusion

Recommendation

- 3.26 Having concluded that ritlecitinib is a cost-effective use of NHS resources (see [section 3.24](#)), the committee recommended it for routine use in the NHS, for treating severe alopecia areata in people 12 years and over.

4 Implementation

- 4.1 Section 7 of the [National Institute for Health and Care Excellence \(Constitution and Functions\) and the Health and Social Care Information Centre \(Functions\) Regulations 2013](#) requires integrated care boards, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this evaluation within 3 months of its date of publication.
- 4.2 The Welsh ministers have issued directions to the NHS in Wales on implementing NICE technology appraisal guidance. When a NICE technology appraisal guidance recommends the use of a drug or

treatment, or other technology, the NHS in Wales must usually provide funding and resources for it within 2 months of the first publication of the final draft guidance.

- 4.3 When NICE recommends a treatment ‘as an option’, the NHS must make sure it is available within the period set out in the paragraphs above. This means that, if a patient has severe alopecia areata and the doctor responsible for their care thinks that ritlecitinib is the right treatment, it should be available for use, in line with NICE’s recommendations.

5 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by [committee A](#).

Committee members are asked to declare any interests in the technology being evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The [minutes of each evaluation committee meeting](#), which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

James Fotheringham

Vice chair, technology appraisal committee A

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser and a project manager.

Albany Chandler and Marcela Haasova

Technical leads

Joanna Richardson

Technical adviser

Thomas Feist

Project manager

ISBN: [to be added at publication]