NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

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

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using ritlecitinib in the NHS in England. The evaluation committee has considered the evidence submitted by the company and the views of non-company stakeholders, clinical experts and patient experts.

This document has been prepared for consultation with the stakeholders. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from the stakeholders for this evaluation and the public. This document should be read along with the evidence (see the committee papers).

The evaluation committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex or sexual orientation?

Issue date: November 2023

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

- The evaluation committee will meet again to consider the evidence, this evaluation consultation document and comments from the stakeholders.
- At that meeting, the committee will also consider comments made by people who are not stakeholders.
- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using ritlecitinib in the NHS in England.

For further details, see NICE's manual on health technology evaluation.

The key dates for this evaluation are:

- Closing date for comments: 15 December 2023
- Second evaluation committee meeting: 16 January 2024
- Details of the evaluation committee are given in section 4

1 Recommendations

1.1 Ritlecitinib is not recommended, within its anticipated marketing authorisation, for treating severe alopecia areata in people 12 years and over.

1.2 This recommendation is not intended to affect treatment with ritlecitinib that was started in the NHS before this guidance was published. People having treatment outside this recommendation may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS clinician consider it appropriate to stop. For young people, this decision should be made jointly by them, their clinician, and their parents or carers.

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. Ritlecitinib may improve quality of life, but it is not clear by how much.

The most likely cost-effectiveness estimates for ritlecitinib are higher than what NICE normally considers an acceptable use of NHS resources. So, ritlecitinib is not 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 The list price is commercial in confidence and cannot be reported here.

2.4 The company has a commercial arrangement, which would have applied if ritlecitinib had been recommended.

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 not 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 eyes

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

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, 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, and that it requires weekly attendance in clinic 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 treatments. 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 equitable 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 on the NHS. So, if recommended, ritlecitinib would be the first treatment available on the NHS with this mechanism of action and 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 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 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 their experience that 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 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 achieve 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 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 The population in ALLEGRO 2b/3 included young people aged 12 to 17 years (14.5%) 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 was representative of 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. They also noted that the proportion of people with alopecia totalis or alopecia universalis was overrepresented and was closer to 10% in practice. The clinical experts explained that it is more difficult to achieve a response 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 and the proportion of young people overall was not.

Clinical effectiveness

3.7 The ALLEGRO 2b/3 trial showed that after 24 weeks, the response rate (the percentage of people achieving 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 compared 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 achieving 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 achieving a response is more difficult for these types of alopecia areata. They noted that it was impressive that ritlecitinib has 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 a total of 2 years of follow-up data from ALLEGRO 2b/3 and ALLEGRO-LT for people taking ritlecitinib (see section 3.7). The company noted that there is no evidence available for the effectiveness of ritlecitinib beyond this. But it suggested 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 start to 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 SALT less than or equal to 10 to SALT more than or equal to 50. Everyone entered the model with a SALT score of 50 or more. Stopping ritlecitinib treatment was assumed if SALT score worsened after 24 weeks on treatment, or if SALT score was more than 20 at 48 weeks or at any point after. 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 more than or equal to 50 best supportive care' health state by moving to a worse health state every cycle. The committee concluded that although using SALT score to define health states does not capture all aspects of alopecia areata (see section 3.4), the model was acceptable for decision making.

Best supportive care

3.11 The EAG explained that the on-treatment and best supportive care health states both included 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 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). The committee concluded that given the inconsistent use of pharmacological treatments for severe alopecia areata and the uncertainty around whether people would use pharmacological treatment after stopping ritlecitinib, it was acceptable to include only non-pharmacological treatment options in the best supportive care health states.

Utilities

The company's vignettes

3.12 The company did a vignette and time-trade-off study to estimate the utility values for people with alopecia areata and their carers. The company developed vignettes which described the impact of having alopecia areata with a specified SALT score, aligned with the health states in the model. The SALT scores were: 10 or less; 11 to 20; 21 to 49 or 50 or more. They also developed a vignette describing the impact of caring for a young person with severe alopecia areata. A time-trade-off approach using the vignettes was used 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 overestimation of the negative impact of the condition. The EAG was also concerned about the face validity (clinical plausibility) of the vignettes compared with the results from the AAPPO results in ALLEGRO 2b/3. The company explained that the differences seen 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 severely affected 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 was 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. The committee noted that the utility values represented the quality of life for the average person within the population in each health state. It considered that although the utility values could be representative for some people with severe alopecia areata, 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, although concerns remained around the validity of the results.

EQ-5D utilities from the trials

3.13 ALLEGRO 2b/3 collected health-related quality-of-life data using a number of 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 (HADS).

Section 4.3 of NICE's health technology evaluations: the manual (2022) states that 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 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 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 EQ-5D is not appropriate then the following sources of utility values can be used, in order of preference:

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

The company argued that 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. Therefore, 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 in people with alopecia areata. It said that this could 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, high baseline scores and the short follow-up period of the trial. But the EAG highlighted that the company had not presented longer-term evidence on EQ-5D from ALLEGRO-LT and that this might help inform the suitability of EQ-5D for estimating health-related quality of life in this population. It also noted that the company had not presented a scenario analysis using the EQ-5D results from the ALLEGRO trials. The committee accepted the limitations of the EQ-5D results from the ALLEGRO 2b/3 trial. But it concluded that it would like to see the longer-term EQ-5D data, up to 36 months, from ALLEGRO-LT as well as the impact of using EQ-5D from the ALLEGRO trials presented in a scenario analysis.

Other utility sources

3.14 The company further argued that EQ-5D from the literature is not appropriate, for many of the same reasons that the EQ-5D data from ALLEGRO 2b/3 is inappropriate, such as content validity (see section 3.13). The clinical and patient experts also said that EQ-5D 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 people with severe alopecia areata such as the psychosocial impact of the condition (see section 3.1). The

EAG disagreed that 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 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 EQ-5D is sensitive to varying alopecia areata severity in a European population. The EAG explained that Bewley et al. was a conference poster but that an equivalent study in a Japanese cohort had been published in a peer-reviewed journal article. The EAG assumed that the same methods had been used to generate the utility values in the Bewley et al. data, which were similar. Based on the methods outlined in the NICE health technology evaluation manual (see section 3.13), the EAG preferred to use this data in the model 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 ALLEGRO 2b/3 to estimate utility values in the model (see section 3.13). It instead chose to estimate utility values from its vignette study (see section 3.12). The committee acknowledged the limitations of the Bewley et al. utility values, noting that the publication was a conference poster and that the mild, moderate and severe disease populations do not directly match to the health states in the model. It also noted that the utilities in Bewley et al. were based on longer-term evidence than the EQ-5D reported in ALLEGRO 2b/3 so may be more sensitive to changes in health-related quality of life. The committee considered the methods in the NICE health technology evaluation manual and noted that if EQ-5D is not considered to be an appropriate measure for estimating utilities, then evidence that it performs poorly on tests of validity and responsiveness, sourced from synthesis of peer-reviewed literature should be presented to

demonstrate this. But it was not convinced that this had been adequately done by the company. It also noted that other generic and condition-specific preference-based measures were in the hierarchy of evidence sources for utility values (see section 3.13). It noted that data using these from ALLEGRO 2b/3 and ALLEGRO-LT may be useful to show the impact of ritlecitinib on health-related quality of life in people with severe alopecia areata. The committee concluded that there was not sufficient evidence that EQ-5D was an inappropriate measure for evaluating changes in disease severity. It also concluded that further evidence was needed to show that other generic (such as SF-6D) and condition-specific preference-based measures could not be used to estimate utilities before it could consider the use of vignettes, which were lower in the hierarchy of evidence sources. The committee recalled that it would like to see analysis using EQ-5D data from the ALLEGRO clinical trials (see section 3.13). But, based on the evidence presented, it considered that the utility values estimated from the Bewley et al. study were the most appropriate to include in the model.

Carer utilities

3.15 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 3.12). 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 carer disutility to carers of young people in the model. 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 have an effect on family members of adolescents. 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

3.16 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 who present 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 which weighted the incremental cost-effectiveness estimate (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 this was more generalisable to the population seen in clinical practice.

Weighting by age

3.17 The EAG highlighted that the company's ICERs did not use the weighted average of outcomes for young people and adults in the model but used average baseline characteristics across the full ALLEGRO 2b/3 population. The company said that age doesn't modify treatment effect and 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. 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 (see section 3.6). The committee considered that there may be real-world evidence available to help inform the estimate of the proportion of people with severe alopecia areata who are young people. The committee concluded that if this were available, it would be appropriate to weight the ICER according to the proportion of young people expected in clinical practice. It further concluded that it was appropriate to weight the ICER according to average outcomes for young people and adults separately given that the carer disutility was only applied for young people.

Long-term treatment effect

3.18 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. It 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. 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

3.19 The length of time people used ritlecitinib in the model was estimated using extrapolated data from ALLEGRO 2b/3. The company chose to use a Weibull model to extrapolate time on treatment. This was based on it being an 'accelerated failure time' model with good AIC/BIC (Akaike information criterion and Bayesian information criterion) ranking and a good 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 and so preferred to use an exponential model to extrapolate time

on treatment, which had better AIC/BIC ranking. It highlighted that there was very little difference in any of the extrapolation curves presented in terms of the AIC/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 and so it was highly uncertain how long on average ritlecitinib would be used for. The committee 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 which reflected that people may request to stop treatment.

Cost effectiveness

Acceptable ICER

3.20 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 top of the range usually considered a cost-effective use of NHS resources.

Committee's preferred assumptions

3.21 The committee's preferred assumptions mostly aligned with the EAG's, which included:

- including non-pharmacological treatments only 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 Bewley et al. (see section 3.14)
- including a disutility for carers of young people with severe alopecia areata (see section 3.15)
- using the average transitions in health states over the final year for which data was available to estimate long-term treatment effect (see section 3.17)
- using the exponential model to extrapolate time to treatment stopping (see section 3.18)
- weighting the average outcomes for young people and adults in the model separately (see section 3.15).

This resulted in a probabilistic EAG base case ICER of £50,123 per QALY gained.

The committee also preferred to weight the proportion of people with alopecia totalis or alopecia universalis (see section 3.16) and the proportion of young people (see section 3.17) in the model according to the proportions expected to be seen in clinical practice. The committee noted that the proportion of people with alopecia totalis or alopecia universalis had been estimated by the company as 9.52%, which it accepted as reasonable based on clinical expert opinion (see section 3.16). It noted that the scenario analysis on the EAG base case which weighted the ICER according to the proportion expected to have alopecia totalis or alopecia universalis in practice reduced the EAG's base case ICER to £43,461 per QALY gained. The committee was less clear what the proportion of people with severe alopecia areata who are young

people is in clinical practice, and no weighted analysis to adjust for this was presented. The committee agreed that of the analyses presented, the EAG's scenario which weighted the ICER according to the proportion expected to have alopecia totalis or alopecia universalis in practice most closely reflected its preferred assumptions. It also noted that it had not been presented with a scenario including the EQ-5D data from the ALLEGRO clinical trials (see section 3.13) and that this was important for understanding the level of uncertainty around the cost-effectiveness estimates. It concluded that it would like to see cost-effectiveness estimates which included all its preferred assumptions as well as a scenario using the EQ-5D data from ALLEGRO-LT. This would include a weighting for the proportion of young people included in the model using an estimate of the proportion of people with severe alopecia areata who are young people in clinical practice.

Other factors

Equality

3.22 The committee considered that some people with severe alopecia areata may be more affected by the psychological impact of hair loss because of the religious significance of hair. Clinical and patient experts also explained that severe alopecia areata can have a particularly high impact on psychosocial health and quality of life for young people. Religion and age are protected characteristics under the Equality Act 2010. However, given that the cost-effectiveness estimates preferred by the committee were not within the range usually considered a cost-effective use of NHS resources, including those for the subgroup of young people aged 12 to 17 years, the committee was unable to make recommendations for these groups.

Conclusion

Recommendation

3.23 None of the cost-effectiveness estimates presented reflected all of the committee's preferred assumptions. The committee requested to see the following evidence and analyses to inform decision making and to reflect all of its preferred assumptions:

- the longest available EQ-5D data from the ALLEGRO trials, according to the SALT score health states in the model (see section 3.13)
- a scenario analysis using the longest available EQ-5D data from the ALLEGRO trials (see section 3.13)
- further evidence that EQ-5D performs poorly on tests of content validity and responsiveness from peer-reviewed literature (see section 3.13)
- further evidence to demonstrate that generic (such as SF-6D) and condition-specific preference-based measures of health-related quality of life are not suitable for estimating utility values for use in the model (see section 3.13)
- an estimate of the proportion of people with severe alopecia areata in clinical practice who are young people (see section 3.15)
- a scenario using all the committee's preferred assumptions, including a weighting based on age using the estimated proportion of people with severe alopecia areata in clinical practice who are young people (see section 3.20).

The committee did not have a cost-effectiveness estimate that reflected all of its preferred assumptions. But the analysis which most closely reflected these resulted in an ICER above the range that is normally considered a cost-effective use of NHS resources. So, ritlecitinib is not recommended for treating severe alopecia areata in people 12 years and over.

4 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
Technical lead

Joanna Richardson
Technical adviser

Thomas Feist
Project manager

ISBN: [to be added at publication]