NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Baricitinib for treating severe alopecia areata

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using baricitinib 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?

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- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using baricitinib in the NHS in England.

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

The key dates for this appraisal are:

Closing date for comments: 20 March 2023

Second appraisal committee meeting: 4 April 2023

Details of membership of the appraisal committee are given in section 4.

1 Recommendations

1.1 Baricitinib is not recommended, within its marketing authorisation, for treating severe alopecia areata in adults.

1.2 This recommendation is not intended to affect treatment with baricitinib 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.

Why the committee made these recommendations

Treatments available on the NHS for severe alopecia areata include topical corticosteroids, which are usually prescribed in primary care. If they do not work, people may be referred to a dermatologist and offered a range of medicines many of which are not licensed for this condition, or a wig.

Evidence from clinical trials suggests that baricitinib improves hair regrowth after 36 weeks of treatment compared with placebo. But treatment needs to be continued to prevent hair loss. Hair loss can cause severe psychological distress, but baricitinib did not show a meaningful improvement in many of the health-related quality of life assessments undertaken in the trials compared with placebo.

The cost-effectiveness estimates for baricitinib are uncertain and are higher than what NICE normally considers an acceptable use of NHS resources. So, baricitinib is not recommended.

2 Information about baricitinib

Marketing authorisation indication

2.1 Baricitinib (Olumiant, Eli Lilly and Company) is 'indicated for the treatment of severe alopecia areata in adult patients'.

Dosage in the marketing authorisation

2.2 The dosage schedule is available in the summary of product characteristics for baricitinib.

Price

2.3 The list price of baricitinib is £805.56 for a 28-tablet pack of either 2 mg or 4 mg tablets (excluding VAT, BNF online accessed February 2023).

2.4 The company has a commercial arrangement. This makes baricitinib available to the NHS with a discount and it would have also applied to this indication if the technology had been recommended. 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 Eli Lilly and Company, a review of this submission by the evidence 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 having severe alopecia areata affects their daily activities and it can have a profound psychosocial impact on all aspects of a person's life. This includes being able to work or study, socialise, take part in leisure activities and have intimate relationships. It can have a significant financial impact because people pay for non-NHS or private healthcare services and treatments. They highlighted the unpredictable nature of the condition and described feelings of shock, trauma, loss of control, disrupted identity, isolation, hopelessness, difficulty coping and sometimes suicidal thoughts. They emphasised that it is an autoimmune disease and not just a cosmetic issue, with no known

cure or effective treatment options. As well as the scalp, severe alopecia areata may also affect other hair such as beards, eyebrows and eyelashes, and can affect temperature regulation and nasal secretions. The patient experts highlighted that some people experience high levels of anxiety and depression. This can be exacerbated by stigma and a lack of understanding from others about the impact the condition can have on their emotional wellbeing and quality of life. The committee concluded that severe alopecia areata can have a profound psychosocial impact on a person's quality of life and that people with the condition would welcome new effective treatment options.

Clinical management

Treatment options

3.2 For severe alopecia areata with at least 50% of the scalp affected, topical corticosteroids may be prescribed in primary care. Clinical experts explained that if the condition does not respond to treatment, people may be referred to a dermatologist and offered different options, many of which are not licensed for alopecia areata. These can include oral or locally injected corticosteroids, dithranol, contact immunotherapy, minoxidil and immunosuppressive medicines such as oral azathioprine, ciclosporin, methotrexate and sulfasalazine. They emphasised that there is no standard care for severe alopecia areata and treatment options vary widely depending on geographic location, healthcare setting, availability and the person's preference. Both clinical experts said that they offer contact immunotherapy in their tertiary clinics but highlighted that it is only for treating scalp hair loss. They also offer immunosuppressant medicines and wigs on prescription. They explained that the criteria for providing wigs varies by region, with some regions requiring annual review by a dermatologist while in others, wigs can be provided without follow up. They explained that the standard wigs provided are the acrylic type, whereas human hair wigs may be offered in some regions once certain criteria are met. One patient expert explained that about 75% of people

with severe alopecia areata wear a wig most of the time. The other patient expert reiterated that there is no clear treatment pathway for severe alopecia areata. They added that there is wide variation in what is offered in secondary care with many clinics in England not offering contact immunotherapy or immunosuppressants. One of the clinical experts explained that dermatologists in secondary care may be less willing to prescribe these treatments because they are off-label, and there is limited evidence of their effectiveness. The committee agreed that there is wide variation in practice in treating severe alopecia areata both in terms of pharmacological options and wig provision. It considered that there is no standard care in the NHS for treating severe alopecia areata. It concluded that the company's and EAG's comparison with no active treatment in their base cases is an acceptable comparator for decision making. It concluded that there is an unmet need for safe and effective treatments for severe alopecia areata.

Severity of Alopecia Tool

3.3 The company submission classified disease severity using the Severity of Alopecia Tool (SALT), which assesses the proportion of the scalp surface area affected by hair loss. A score of 0% represents no hair loss, while a maximum score of 100% represents total hair loss. The company submission defined severe disease as a SALT score of 50 to 94 and very severe disease as a SALT score of 95 to 100. The clinical experts explained that SALT is commonly used in research studies but is not widely used in clinical practice. Other measures such as the Physician Global Assessment may be used clinically to assess disease severity. There were differing views on a clinically meaningful SALT outcome. The company and EAG used an absolute measure of a SALT score of 20 or less which represents no more than 20% of scalp surface area involved in their base case. One clinical expert preferred the use of the absolute SALT score and highlighted that a SALT score of 20 or less after treatment had been validated in research studies as a clinically meaningful change in people with alopecia areata. However, the other

clinical expert highlighted concerns about using an absolute score of SALT 20 or less, which would represent a large change for people with very severe disease. They considered a relative measure of SALT50, that is, a 50% reduction from baseline SALT score or SALT75 to be more clinically meaningful. The committee recognised the lack of consensus on the use of absolute or relative SALT scores and thresholds for a clinically meaningful outcome. It noted clinical experts to the EAG considered SALT75 to be nearly equivalent to a SALT score of 20 or less in severe disease. It concluded that the company's and EAG's use of a SALT score of 20 or less for treatment response in their base case is appropriate for decision making.

Treatment pathway

Positioning of baricitinib

3.4 The marketing authorisation for baricitinib includes adults with severe alopecia areata, including people who have had previous treatment and those who have not. The clinical experts explained that they would use baricitinib at the same position as contact immunotherapy and immunosuppressants, in a secondary care setting rather than tertiary care. They explained that, if recommended, baricitinib will be the first licensed option for treating severe alopecia areata. However, its use will depend on person circumstances and preferences. For example, some people may prefer to have a local treatment such as contact immunotherapy rather than a systemic medicine like baricitinib. The clinical experts considered that distinguishing between people who have had treatment previously (treatment-experienced) and those who have not (treatment-naive) is not helpful in deciding who should have baricitinib. They noted that given the wide geographical variation in care, it is likely that most people will not have had pharmacological treatment for their condition. One clinical expert explained that baricitinib would not be offered to people with patchy alopecia areata of less than 6 months duration because they are more likely to regrow hair spontaneously.

However, hair regrowth is unpredictable in this condition. In addition, they explained that the chance of hair regrowth decreases the longer the duration of alopecia areata. The committee concluded that baricitinib is likely to be used to treat both newly diagnosed and long-term severe and very severe alopecia areata.

Clinical evidence

Data sources and generalisability

3.5 The main evidence for baricitinib is from 2 multi-national, multicentre (no UK or European centres), randomised, double-blind, parallel-group trials comparing 2 mg or 4 mg baricitinib with placebo for 36 weeks during the induction period. This was followed by a maintenance period in which people whose condition responded to treatment continued to have baricitinib or placebo. Adults (aged 60 years or under for males and 70 years or under for females) with severe alopecia areata were included in the trials. Severe alopecia areata was defined as a current episode of more than 6 months but less than 8 years, a SALT score of 50 or more at baseline, and no spontaneous improvement in the past 6 months (a reduction in SALT score of 10 or less). The primary end point was a SALT score of 20 or less at week 36, and the phase 3 data using baricitinib 4 mg was used to inform the economic model. The clinical experts considered the BRAVE trial populations to be broadly generalisable to those likely to have baricitinib in the NHS. They noted that many of the people were recruited from North America and had similar demographics. People tended to have had multiple treatments, although some treatments, such as cryotherapy would not be offered in the NHS. The EAG noted that people in the trials had more severe and difficult to treat alopecia areata, more similar to people likely to be seen in the NHS with long-term severe alopecia areata than newly diagnosed people with severe alopecia areata seen in the NHS. As such, they considered that the findings from BRAVE may underestimate the potential treatment effect in people with newly diagnosed severe alopecia areata. The clinical expert

also noted that the peak incidence of alopecia areata is in people who are about 25 years old. The committee concluded that the BRAVE trial populations were likely to be broadly generalisable to people likely to have baricitinib in the NHS, but acknowledged that the treatment effects may be underestimated in a newly diagnosed treatment-naive population.

Health-related quality of life measures

3.6 The health-related quality of life measures that were assessed in the BRAVE trials included the EQ-5D, Hospital Anxiety and Depression Scale, Short-Form 36 questionnaire and the Skindex-16 Alopecia Areata. At baseline, almost half the people with severe or very severe alopecia areata in the trials had EQ-5D scores of full health. The clinical experts explained that these measures are mainly used in research. Additionally, people who were more psychologically impacted by their condition may not have been eligible to take part in the trials. The patient expert suggested that people who enrol into a trial may have lower rates of anxiety than would be expected in the NHS, because people in trials have hope of being treated. However, they explained that if hair regrowth occurs, people may also become anxious about losing their new hair when they stop taking the treatment. They noted that alopecia areata can limit how people live their lives, and that over time, people can develop self-coping strategies such that the full impact on quality of life may not be readily observed. The clinical experts noted that high levels of anxiety and depression are common, occurring in about 1 in 3 people with severe alopecia areata. The committee acknowledged that severe alopecia areata can have a profound impact on quality of life that is not shown in the overall baseline EQ-5D scores for people taking part in the BRAVE trials. It considered whether this is because the EQ-5D may not be picking up important aspects of the condition or because the people in the trials are not representative of those with severe alopecia areata in terms of anxiety and depression.

Treatment response and health-related quality of life

3.7 Compared with placebo, people taking baricitinib 4 mg were more likely to have a treatment response, measured as a SALT score of 20 or less (about 34% for baricitinib compared with 4% for placebo) at 36 weeks. At 52 and 76 weeks, compared to placebo, people whose condition had responded to treatment and had continued to take baricitinib were less likely to have hair loss. Statistically significant improvements in the emotions and functioning domains of the Skindex-16 Alopecia Areata measure were observed for people having baricitinib compared with placebo at 36 weeks (results are academic in confidence and cannot be reported here). But the changes seen in many of the health-related quality of life measures that were assessed in the BRAVE trials, including the EQ-5D, Hospital Anxiety and Depression Scale and Short-Form 36 questionnaire, were either not statistically significant or clinically meaningful. The EAG acknowledged that there is likely to be a group of people for whom severe alopecia areata can have a large negative impact on quality of life and for whom treatment with baricitinib may result in large improvements in quality of life. However, large overall improvements in EQ-5D scores in the BRAVE trial populations were not observed. This may be because:

- almost half the people with severe alopecia areata had EQ-5D scores of full health at baseline (see section 3.6); and
- only about 1 in 3 people having baricitinib had a treatment response (measured as a SALT score of 20 or less) at 36 weeks.

The committee concluded that baricitinib is clinically effective at improving hair regrowth compared with placebo at 36 weeks. It noted that treatment with baricitinib in the maintenance period must be continued to prevent hair loss. It acknowledged that hair regrowth can have a profound impact on improving a person's quality of life, but based on the data from the BRAVE trials, the extent of this improvement in quality of life is uncertain.

Adverse events

3.8 The company reported adverse event data which showed that the short-term safety profile of baricitinib compared with placebo is favourable (findings are academic in confidence and cannot be reported here). It considered that these adverse events were mild, had little significant detriment to health-related quality of life and did not increase healthcare costs, so did not include adverse events in its economic model. The committee recognised the short-term safety profile of baricitinib was favourable but noted that the long-term safety is uncertain.

Economic model

Company's model structure

3.9 In its submission, the company made the case that baricitinib improves health-related quality of life, but does not prolong life. It developed a Markov model with 4 health states over a lifetime time horizon and 4 weekly cycles. In its base case, it assessed the cost-effectiveness of baricitinib 4 mg compared with no active treatment for treating severe alopecia areata in adults. The health states are induction, maintenance, best supportive care and death. Everyone enters the model by the induction state and either start treatment on baricitinib 4 mg or no active treatment for 36 weeks. During this phase, people can move to the best supportive care state based on all-cause treatment stopping from the BRAVE trials (except for stopping because of lack of efficacy). At the end of the induction period, people are categorised based on treatment response from the BRAVE trial data. At technical engagement, the company made various changes to its original base case, including its definition of treatment response. A SALT score of 20 or less was used in the company's revised base case. People whose condition responds to treatment move to the maintenance state where they stay until they lose treatment response or stop treatment because of any other reason, after which they move to the best supportive care state. People whose condition does not respond to treatment move to the best supportive care

state where they stay until the end of the model time horizon or death. At any point in the model time horizon, people can move to death from all health states, but no one can move from having a treatment non-response to a treatment response after the end of the 36 week induction period. The EAG considered the model structure to be appropriate and similar to other dermatological conditions such as atopic dermatitis. The committee concluded that the company's economic model for its revised base case was appropriate for decision making.

Modelling of best supportive care

Composition of best supportive care

3.10 In its base case, the company included a basket of treatments in best supportive care that contained different treatments costed from a primary care perspective (ciclosporin, methotrexate, azathioprine, intralesional steroids, contact immunotherapy, prednisolone, topical corticosteroids, topical minoxidil foam, oral minoxidil, mycophenolate mofetil, anthralin cream) and modacrylic wigs. The choice of treatments was based on data collected from 117 people with severe or very severe alopecia areata in the UK, referred to as the Adelphi study which was a company-sponsored online survey. The EAG noted that most people in the Adelphi study were treatment-experienced and had already tried many previous treatments. So the EAG considered it unlikely that people would be willing to try more pharmacological treatments that have limited effectiveness over a lifetime horizon, after all other options had been exhausted. As such, in its base case, it assumed that no one had best supportive care, and only included wigs and orthotics (see section 3.11). The committee recalled that there is no standard care for treating severe alopecia areata in the NHS and that there is great geographical variation in access to different pharmacological treatments and wigs. The committee noted this may suggest that people in the NHS may be more likely to be treatment-naive to pharmacological options (see section 3.2). It noted that the Adelphi study included people who were recruited by their dermatologists which

suggests they may be more likely to be engaged in their care. Also, some treatments such as contact immunotherapy and immunosuppressants are less likely to be prescribed in secondary care than in tertiary care (see section 3.2). The committee considered that the composition of best supportive care, particularly over a lifetime horizon, is uncertain. Based on the clinical and patient experts' feedback, it concluded that there is wide variation in access to treatments, and that it is likely people would have limited pharmacological options and are more likely to use wigs and orthotics.

Best supportive care use after non-response

3.11 In its original base case, the company assumed that the same proportion of people on baricitinib or no active treatment whose condition does not respond would have best supportive care for the remaining lifetime time horizon of the model. The proportion was based on the Adelphi study (the exact figure is academic in confidence and cannot be reported here). After technical engagement, the company assumed that people who have baricitinib are less likely to have best supportive care when their condition no longer responds to treatment (about half) compared with no active treatment. The EAG noted that this differential best supportive care use is a major driver of costs in the model because of the relatively high costs of best supportive care compared with baricitinib. It did not consider that there was evidence to suggest that people on baricitinib whose condition has not responded are less likely to have further treatment than people who have not had baricitinib. So, it assumed that people whose condition had not responded to treatment in both arms would only have wigs and orthotics. The clinical experts considered that it is unlikely that people who had been offered all possible treatment options and whose condition did not respond to treatment would be followed up indefinitely. They considered that people would likely have wigs and orthotics, but could not provide specific proportions of people. They also could not confirm that these proportions are likely to be different depending on whether people had previously had baricitinib or no active treatment. The EAG provided

scenario analyses which assume the same proportion of people have best supportive care after non-response in both arms, but adjusted the proportions, that is, 0% (base case), 25% and 50%. The EAG's base case and scenarios included the use of wigs and orthotics. The committee considered that this was an area of high uncertainty. Given the lack of evidence, it made a conservative conclusion that the same proportion in both arms should have best supportive care after non-response and considered the impact of the range of proportions provided by the EAG.

Modelling utility values

Source of utility values

3.12 In its base case, the company preferred to use utility values derived from EQ-5D data from its Adelphi study. This is because it considered that the utility values were more plausible because only about 1 in 5 respondents had scores of full health at baseline, compared with almost half the people in the BRAVE trials (see section 3.6). The EAG preferred to use utility values from the BRAVE trials because it could be used to estimate within-person changes in EQ-5D after treatment response (SALT score of 20 or less). Also, the BRAVE trials are in line with the model structure, and are of high quality with a sample size that is more than 4 times that of the Adelphi study. The EAG highlighted that the baseline EQ-5D scores from BRAVE are similar to those reported in other studies for people with severe alopecia areata. The committee recalled the feedback from the patient and clinical experts about the issues of capturing health-related quality of life data in people with severe alopecia areata (see section 3.6). It considered that the quality-adjusted life year (QALY) gains with treatment in the BRAVE trials may be underestimated. But, the committee agreed that the EQ-5D data over a longer period (36 weeks up to 76 weeks) from BRAVE is more robust compared with data from the smaller, cross-sectional Adelphi survey that captured quality of life data at one time point. It concluded that the utility values derived from the EQ-5D data from the BRAVE trials are preferred for decision making.

Cost-effectiveness estimates

Acceptable ICER

3.13 NICE's guide to the methods of technology appraisal notes that above a most plausible incremental cost-effectiveness ratio (ICER) of £20,000 per QALY gained, judgements about the acceptability of a technology as an effective use of NHS resources will take into account the degree of certainty around the ICER. The committee will be more cautious about recommending a technology if it is less certain about the ICERs presented.

The committee noted the high level of uncertainty, specifically around the following issues:

- no clear consensus on standard care (see section 3.2)
- no clear consensus on a clinically important SALT outcome (see section 3.3)
- the evidence of baricitinib's effectiveness in the treatment-naive population is uncertain but likely to be underestimated based on BRAVE outcomes (see sections 3.5, 3.6 and 3.7)
- the effect of baricitinib on health-related quality of life is uncertain (see sections 3.6 and 3.7)
- the long-term safety of baricitinib is unknown (see section 3.8)
- no clear consensus on composition of best supportive care (see section 3.10)
- no clear consensus on the proportion of people likely to have best supportive care after all other options have been exhausted, over a lifetime time horizon (see section 3.11)
- no evidence on the differential use of best supportive care between the baricitinib and no active treatment arms after all other options have been exhausted over a lifetime time horizon (see section 3.11)
- the QALY gains with treatment may be underestimated in the BRAVE trials (see section 3.12).

Because this uncertainty could mean the true ICER is above what NICE normally considers cost-effective, the committee agreed that an acceptable ICER would be towards the lower end of the range normally considered a cost-effective use of NHS resources (£20,000 to £30,000 per QALY gained).

Preferred assumptions

3.14 The committee's preferred assumptions were for the model to:

- use no active treatment as the comparator (see section 3.2)
- use a SALT score of 20 or less as a clinically meaningful outcome (see section 3.3)
- include adverse events (see section 3.8)
- include only wigs and orthotics in best supportive care (see section 3.10)
- apply the same proportion of people having best supportive care after all other options have been exhausted for both the baricitinib and no active treatment arms, but consider a range of proportions in best supportive care use (see section 3.11)
- use utility values derived from EQ-5D data from the BRAVE trials (see section 3.12).

The committee considered that the EAG analyses reflected their preferred assumptions.

Other factors

Equality issues

3.15 The committee acknowledged that beard hair loss may have a greater religious implication for people of some faiths. Also, alopecia areata may be more common in people of Asian family background, lower socioeconomic status and in people living in urban areas. The clinical experts explained that referral to secondary care in these groups are likely to be lower compared with other people with severe alopecia areata. The

committee noted that all these patient groups are included in this appraisal. It recalled the issues of measuring health-related quality of life in this condition, and acknowledged that some people may be affected by stigma (see section 3.1 and section 3.6). It noted that stigma may affect people's behaviour in a way that changes the effectiveness of an intervention and that routine quality of life assessments may not capture the benefits of treatment. However, it considered that these factors did not alter its conclusions.

Innovation

3.16 The clinical experts considered baricitinib to be a step-change in managing severe alopecia areata for which there are limited licensed treatment options. The committee acknowledged that there may be benefits with baricitinib that were not captured in the modelling, but evidence for this was limited. The committee concluded that baricitinib is innovative.

Conclusion

Recommendation

3.17 All the ICERs in the EAG analyses (base case and scenario analyses using 25% or 50% best supportive care use in both arms) were higher than the range considered to be a cost-effective use of NHS resources (£300,710 to £425,560 per QALY gained). So, baricitinib could not be recommended for routine commissioning in the NHS for treating severe alopecia areata in adults.

4 Evaluation committee members and NICE project team

Appraisal 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

Radha Todd

Chair, technology appraisal committee A

NICE project team

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

Sharlene Ting

Technical lead

Eleanor Donegan

Technical adviser

Thomas Feist

Project manager

ISBN: [to be added at publication]