NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Appraisal consultation document

Betibeglogene autotemcel for treating transfusion-dependent beta-thalassaemia

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

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

The appraisal 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 race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

Issue date: February 2021

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

For further details, see NICE's guide to the processes of technology appraisal.

The key dates for this appraisal are:

Closing date for comments: 4 March 2021

Second appraisal committee meeting: TBC

Details of membership of the appraisal committee are given in section 5

1 Recommendations

1.1 Betibeglogene autotemcel is not recommended, within its marketing authorisation, for treating transfusion-dependent beta-thalassaemia (TDT) in people 12 years and older who do not have a beta^0/beta^0 genotype, when haematopoietic stem cell transplantation (HSCT) is appropriate but a human leukocyte antigen-matched related HSC donor is not available.

1.2 This recommendation is not intended to affect treatment with betibeglogene autotemcel 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. When the person having treatment is under 18 years, this decision should be made jointly by the clinician and the child or young person, and their parents or carers.

Why the committee made these recommendations

Current treatment for TDT includes regular blood transfusions, and iron chelation therapy for excess iron build up in the body from the transfusions. HSCT is a curative therapy that is available for a few people with the condition. But they need a matched donor and to be within a specific age range. Betibeglogene autotemcel is approved for people who do not have a beta^0/beta^0 genotype, and when HSCT would be appropriate but there is no suitable donor.

Clinical trials for betibeglogene autotemcel are small and people have not been followed up for very long. However, the trials suggest that, after betibeglogene autotemcel, some people either eventually do not need blood transfusions any more or need them less often.

There are uncertainties about the cost effectiveness and the preferred estimate for betibeglogene autotemcel is considerably higher than what NICE normally considers

an acceptable use of NHS resources. So, betibeglogene autotemcel is not recommended.

2 Information about betibeglogene autotemcel

Marketing authorisation indication

2.1 Betibeglogene autotemcel (Zynteglo, bluebird bio) is indicated for 'the treatment of patients 12 years and older with transfusion-dependent beta-thalassaemia (TDT) who do not have a $beta^0/beta^0$ genotype, for whom haematopoietic stem cell (HSC) transplantation is appropriate but a human leukocyte antigen (HLA)-matched related HSC donor is not available'.

Dosage in the marketing authorisation

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

Price

2.3 The cost of betibeglogene autotemcel 1.2 to $20x10^6$ cells/ml dispersion for a one-time infusion is £1,450,000 per patient (list price, excluding VAT). The company has a commercial arrangement, which would have applied if the technology had been recommended.

3 Committee discussion

The appraisal committee considered evidence submitted by bluebird bio, a review of this submission by the evidence review group (ERG), NICE's technical report, and responses from stakeholders. See the committee papers for full details of the evidence.

The appraisal committee was aware that several issues were resolved during the technical engagement stage, and agreed that:

- Limited data were presented to support the degree of utility impact associated with infertility and it was agreed to remove it from the model. The utility decrement from subcutaneous chelation therapy during the iron normalisation period should be taken into account. The whole population utility dataset from Ara and Brazier (2011) should be used for age-adjusting utilities, rather than a subset excluding people with existing health conditions (issue 2, see technical report pages 16 to 20).

- The baseline population value for the proportion of people in UK with transfusion-dependent beta-thalassaemia (TDT) who have hypogonadism should be 20%. Category-specific body weights should be used for dosing, instead of the same weight for all modelled patients in the analyses (issue 3, see technical report pages 20 to 23).

- The model should reflect updated practices for iron chelation in the NHS, where some people with TDT now have a combination of oral treatments (issue 5, see technical report pages 24 to 26).

- The ERG's preferred standardised mortality ratio (2) should be used. This is because the data used to generate the ratio in the company's original analyses are based on out-of-date non-UK data (issue 6, see technical report pages 27 to 29).

- Prices from the drugs and pharmaceutical electronic market information tool should be used for the 3 generic drugs in the model. This is to better match the prices paid by the NHS (issue 7, see technical report pages 30 to 31).

- It is appropriate for the model to include 1 consultant-led appointment annually for 15 years for every modelled patient in the betibeglogene autotemcel arm. This is to better reflect follow-up costs (issue 8, see technical report pages 31 to 32).

- The only myeloablative conditioning regimen in the model should be the busulfan combination in line with its marketing authorisation (issue 9, see technical report pages 32 to 33).

The committee recognised that there were remaining areas of uncertainty associated with the analyses presented (see technical report, table 2, pages 38 to 39), and took these into account in its decision making. It discussed issues 1, 2 (partial), 3 (partial), 4 to 6 and 10, which were outstanding after the technical engagement stage.

The condition and clinical management

TDT is a lifelong condition with difficult management regimens

3.1 Many people with TDT face a lifetime of regular blood transfusions and iron chelation treatment. The clinical experts explained that blood transfusions are given every 2 to 4 weeks (depending on the severity of anaemia), with 2 or 3 units of blood given each time. Iron chelation is needed to remove excess iron that builds up because of these transfusions. The aim is to try and avoid secondary conditions such as cardiac and liver disease, diabetes, hypogonadotrophic hypogonadism and osteoporosis. Iron chelation treatment is itself associated with adverse effects such as kidney toxicity, growth delay and problems with hearing and vision. Subcutaneous chelation is particularly painful, and causes large painful lumps at injection sites. People quickly run out of possible injection sites to use. Oral chelation, while seen as more convenient, can negatively affect the immune system and overall health. The patient experts explained that people with TDT experience 'brain fog', extreme fatigue, bone pain and low mood when due for a transfusion. They also said that it can take many people up to 5 days to recover after a transfusion. The United Kingdom Thalassaemia Society guidelines state that the 'expectation is that well monitored and chelated patients will have a near normal life expectancy'. However, 1 clinical expert explained that this might assume perfect compliance. They added that, in clinical practice, people who had less than optimal prior treatment still die young because of disease complications, despite chelation practices improving over the last 20 years. The committee concluded that TDT had an effect on life expectancy in the past. However, it concluded that it is unclear

what life expectancy with the condition is now because of changes in treatment practices.

TDT has a significant effect on the lives of people with the condition

3.2 The patient experts explained that people with TDT need a lot of time off education and work to manage their condition. They added that it is painful, and limits opportunities for travel, even within the same country. This is because of the need to be close to their hospital for transfusions and in case of complications. There is a risk of anaphylaxis with each transfusion, and sepsis or infections are a risk because of the multiple different ports and injection sites used. There is also stigma associated with scarring and needle marks. They explained that management options have improved compared with 20 years ago. However, people with TDT still struggle with regimen compliance (for example, oral chelation is seen as easier than subcutaneous administration, but still needs daily medication, sometimes several times per day). One patient expert explained that keeping up with treatment 'is a full-time job', and it is 'not possible to have a normal life'. They added that people with the condition try to remain positive despite their lives being unpredictable and, historically, limited in terms of life expectancy. The ERG pointed out that, in the company's UK patient preference report, only 37% of respondents said they would immediately accept a referral to a transplant specialist and betibeglogene autotemcel if offered it. This may indicate a preference for best supportive care over the new intervention, on balance. The committee understood that managing TDT is still difficult despite therapeutic improvements in recent decades. It concluded that TDT has a significant effect on the lives of people with the condition.

Curative treatment via haematopoietic stem cell transplantation (HSCT) is only available to a few people with TDT

3.3 Submissions from clinical experts and the company outlined improved outcomes in the last 20 years for people with TDT because of better iron monitoring and chelation methods. The clinical experts explained that, in

NHS clinical practice, curative treatment via allogeneic HSCT is typically only attempted in children younger than 9 years (although it is commissioned in young people younger than 19 years). They added that, even in children, HSCT is associated with a 5% to 15% mortality risk, and the mortality and morbidity risk is higher in people who are older. This is because of an increasing risk of complications such as graft versus host disease. Also, allogeneic HSCT needs a matched donor, which is not possible for everyone with the condition, and is inaccessible to many people with TDT because of their age.

Clinical effectiveness

Evidence on the clinical effectiveness of betibeglogene autotemcel is from a small number of people, and there are limited follow-up data

3.4 The company's main clinical evidence for betibeglogene autotemcel came from phase 1 and 2 studies (HGB-204, HGB-205), and 1 study in people with non-beta^0/beta^0 genotypes (HGB-207). All were multicentre single-arm trials, with transfusion independence (TI) as the primary outcome measure, and transfusion reduction as a secondary outcome measure. Across these trials, data from a total of 24 people were evaluable for TI at the time of the company submission, of which 83% (20 people) reached TI. No events for loss of TI had been recorded by the time of the company's submission. Of the 4 people who had not reached TI in the transplant population, 2 had 'substantial' transfusion reductions (a 60% or more reduction in frequency) and 2 were transfusion dependent. The committee noted that the clinical evidence base was limited, but that more people in the trials would reach the point of being 'transfusion evaluable'. It also noted that there was an ongoing long-term follow-up study. The committee was also aware of some abstracts with a later data cut (and more people evaluable for TI) entering the public domain in December 2020, but these data were not used in the analyses. The committee understood that the available clinical-effectiveness data came from a

small number of people and that follow-up data were limited. It concluded that further follow-up data would be welcome.

Discounting in the economic model

The life expectancy for people with TDT means the criteria for using a 1.5% discount rate for costs and benefits are not met

3.5 NICE's guide to the methods of technology appraisal uses a 3.5% discount rate for both costs and benefits in its reference case. Analyses that use a non-reference-case discount rate for costs and outcomes may be considered when:

- treatment restores people who would otherwise die or have a very severely impaired life to full or near full health
- this is sustained over a very long period (normally at least 30 years).

This is because, in these circumstances, cost-effectiveness analyses are very sensitive to the discount rate used. So, a discount rate of 1.5% for costs and benefits may be considered by a committee if it is highly likely that, on the basis of the evidence presented, long-term health benefits are likely to be reached. Also, the committee would need to be satisfied that the introduction of the technology does not commit the NHS to significant irrecoverable costs. One clinical expert referred to Hospital Episode Statistics (HES) data from the NHS. They stated that this showed that people with TDT are still dying at a younger age than people without TDT, despite improvements in its clinical management. The company referred to a 10-year forward-looking cohort analysis for the period 2009 to 2018 that used HES data (Jobanputra et al. 2020). The analysis incorporated 85% of people with TDT in the UK. It found that the median age of death for people with TDT in that period was 45 years (interquartile range 29 to 52). This specifically represented the age of patients in hospital from the cohort who died during the follow-up period of the HES data. It also included people with $beta^0/beta^0$ genotypes and comorbidities that would

have excluded them from having betibeglogene autotemcel. The ERG explained that betibeglogene autotemcel is licensed specifically for people who are fit enough to tolerate myeloablative treatment with busulfan before having betibeglogene autotemcel. These people would likely have better outcomes than the general TDT population, even without betibeglogene autotemcel (that is, would likely have a life expectancy of more than 45 years). As well as this, the committee understood that:

- the HES analysis included people who had decades of treatment before improved chelation techniques had become available
- the average life expectancy for people with better-managed TDT using newer chelation methods from the last 20 years would likely be higher.

The committee concluded that the life expectancy for people with TDT means that the criteria for using a 1.5% discount rate for costs and benefits were not met.

The argument used for considering a 1.5% discount rate for Strimvelis is not relevant to this appraisal

3.6 The company explained that, because betibeglogene autotemcel is a gene therapy, there is a rationale for benefits being sustained over a long period. It also said that it believed the growing evidence base showed that lasting positive outcomes are highly likely to be achieved. The company gave an example of NICE's highly specialised technology appraisal guidance on Strimvelis for treating adenosine deaminase deficiency–severe combined immunodeficiency (a different gene therapy and condition) to support the use of a 1.5% discount rate in the betibeglogene autotemcel appraisal. However, in the Strimvelis appraisal, the committee concluded that it was uncertain whether Strimvelis fully met the criteria to use a discounting rate of 1.5%. It also recognised that there were uncertainties in whether the long-term benefits of treatment would be achieved because of limited evidence. So, it was decided that both 3.5% and 1.5% should be considered in the Strimvelis committee's decision

making. This committee noted that the incremental cost-effectiveness ratios (ICERs) for Strimvelis were cost effective with both discount rates, so this was not material to the decision making. At technical engagement, the Cell and Gene Therapy Catapult, a stakeholder for this topic, explained that discounting disproportionately affects benefits of therapies with high upfront costs but longer-term benefits. It explained that long-term evidence on sustainability of effect would be very difficult to generate in time for launch (without delaying patient access significantly). It also explained that traditional assessment frameworks could penalise lack of such evidence at the time of appraisal. The committee noted that the long-term data for betibeglogene autotemcel were much more uncertain than the data for Strimvelis. Like betibeglogene autotemcel, Strimvelis has a high upfront cost and longer-term benefits. However, at the time of its appraisal, there were data from a maximum follow up of 13 years, compared with just over 5 years for betibeglogene autotemcel. The committee also noted that the condition Strimvelis has an indication for is fatal within infancy if it is untreated and, at the time of that appraisal, HSCT was the only treatment option outside trials. Strimvelis was considered to be transformative for people who, without treatment, would otherwise die. TDT, however, has other management options apart from HSCT. The committee concluded that the rationale used for the consideration of a 1.5% discount rate as part of decision making for Strimvelis was not relevant to this appraisal.

The criteria for using a 1.5% discount rate for costs and benefits are not met, so a 3.5% discount rate should be used

3.7 The committee noted that some people who have treatment with betibeglogene autotemcel do not see a reduction in transfusions, let alone become TI. It noted the lack of long-term evidence for betibeglogene autotemcel (see section 3.16) and the utility decrement associated with TDT (see section 3.9). It considered that the criteria of 'treatment restores people who would otherwise die or have a very severely impaired life to full or near full health, and when this is sustained over a very long period

(normally at least 30 years)' were not met. Also, it was not satisfied that the introduction of the technology would not commit the NHS to significant irrecoverable costs. Although no loss of efficacy has been seen, follow-up data are limited and there may be irrecoverable costs if betibeglogene autotemcel does not work as well in clinical practice. So, the committee concluded that, because the criteria for using a 1.5% discount rate were not met, the reference case value of 3.5% should be used for both costs and benefits.

Utility values used in the model

Utilities from the company's UK chart review may not be representative of the population who would be eligible for betibeglogene autotemcel

3.8 The company did not use the utility values collected from its clinical trials in the modelling. This was because it thought the baseline values collected were 'artificially high', possibly because of an 'adjustment bias'. It stated that this made it hard to detect a utility benefit after treatment with betibeglogene autotemcel. Instead, it used the mean utility (of 0.69) of all people with TDT aged 16 years or older from its UK chart review (gathered via a retrospective vignette study). This generated a utility decrement of 0.27 associated with being transfusion dependent. It applied this decrement to the general population utilities. The ERG explained that there were only 10 adults with utility data in the clinical trials. The demographics in the company's chart review were very different from those in the trials. The health-related quality of life (HRQoL) data were derived from an older population in the chart review, while the model adjusted utilities over time to account for ageing. So, the company's model may have double counted the effects of ageing. The ERG considered that using a mean utility score for people who were transfusion dependent based on the older chart review population might have been inappropriate. HRQoL may be lower for older people than younger people. This is because of changing chelation practices and more optimal management in recent years. The committee noted that the population in

the UK chart review may not be representative of people who would be eligible for betibeglogene autotemcel. It concluded that this meant the utilities sourced from the full chart review population may not be suitable for decision making.

Utilities from a subsection of the chart review population are more comparable to the clinical trials

3.9 The ERG did a reanalysis of the chart review data. This included people with TDT aged 12 to 35 years to match the trial population. It excluded people with existing comorbidities that were already separately considered in the modelling. The ERG pointed out that this resulted in a higher mean utility than that of the complete chart review population. It preferred to use this utility value for the transfusion-dependent state in the model. This was because it was more comparable to the baseline value for people in the trials, and to the population who might be eligible for treatment with betibeglogene autotemcel in clinical practice. The ERG also considered that utilities from the chart review were not the most appropriate representation of HRQoL associated with TDT at baseline. This was because there were 2 sources of EQ-5D data available that were collected directly from people with TDT. The ERG thought it was unreasonable for company to dismiss both these patient-derived values as unrealistic. They already represented a substantial utility decrement of about 0.1 compared with general population values for the age group. So, they may have already incorporated a reduced baseline utility score compared with the general population. The company countered that people with TDT and their families are fully aware of the natural history of the condition and its potential effect on future quality of life. It thought that this would likely be a motivating factor for seeking gene therapy, rather than indicating the current disease burden on their health and wellbeing. The committee considered that the utilities from the ERG's reanalysis of the chart review data better matched the clinical trials.

Utility values should be taken from a population that matches the clinical-effectiveness population

3.10 The company proposed 1 reason why using the baseline values in the trial may have underestimated the utility gain from betibeglogene autotemcel. This was that people with TDT have a degree of adaptation to, acceptance of or trying to keep a positive outlook with the condition. So, the committee considered a possible approach of looking at patient experience from before and after HSCT to gain insights into how much betibeglogene autotemcel might improve HRQoL. However, the clinical experts explained that they do not see people who have had successful HSCT in their day-to-day practice because they are effectively cured. This meant that the committee was unable to get any insight on this. A patient expert outlined the effect of graft versus host disease and specific skin concerns in the population who have HSCT, which would not occur with betibeglogene autotemcel. Therefore, this proxy approach would not give accurate utility values for people with TDT before and after betibeglogene autotemcel treatment. The committee considered that the company's preferred utility decrement was too large (see section 3.8). It concluded that the ERG's approach (see section 3.9), using utility data from people aged 12 to 35 years (as in the trial) who are fit enough to have betibeglogene autotemcel, should have been used in the modelling.

The clinical-effectiveness data and utility data in the company's model comes from populations with different age ranges

3.11 The marketing authorisation for betibeglogene autotemcel includes people 12 years and older. In the company submission, only people 35 years or younger were included in the clinical-effectiveness population (that is, people with evaluable data for TI from HGB-204, HGB-205 and HGB-207). In HGB-204 and HGB-205, the protocols meant people in the trials could be no older than 35 years. The company explained that the population of 12 to 34 year olds was chosen for their model base case to reflect a cohort of people wanting to have a gene therapy. However, the

company's UK chart review, which it used for utility data, included people younger than 12 years (which did not match the marketing authorisation). Also, a large proportion of people with TDT in the review were over 35 years (which did not match the company's base case or clinical-effectiveness population). The company justified this by explaining that age is not a variable that determines eligibility for treatment, and the marketing authorisation does not state an upper age limit for treatment. The committee noted the discrepancies between the population in the full UK chart review and the clinical-effectiveness evidence population.

The population characteristics for the utility data from the company's UK chart review should match those of the clinical-effectiveness population

3.12 The ERG considered that the full chart review included people who would not be eligible for betibeglogene autotemcel. It highlighted that the company resubmitted its analysis using a distribution of chelation therapies that had been age-matched to the clinical trials, which resulted in a lower ICER. However, the company also maintained its previous position on the estimation of utilities (which, if changed, would result in a higher ICER). The ERG explained that data on quality of life and chelation treatment in the trials were more likely to reflect the data from the younger population in the chart review. This was because everyone in the trials was younger than 35 years and because people 35 years or older are less likely to be fit enough to have betibeglogene autotemcel in clinical practice. The clinical experts thought that the therapy would be given in the UK before complications of TDT develop, which would most likely be before the age of 35 years. They explained that, if the marketing authorisation were updated, they would favour using the treatment in children younger than 12 years, rather than young people aged 12 to 17 years. This was because they considered a younger age the most clinically favourable time for the treatment to be used. The committee understood that the company's updated analysis used the lower age range for chelation treatments, but not for utilities. The committee considered that, for consistency, the same approach should be taken for

both. It concluded that the population characteristics for the utility data from the company's UK chart review should match those of the clinical-effectiveness population used in the modelling.

Modelled population

Adjustment to fix possible underrepresentation of the population with a severe non-beta^0/beta^0 genotype is inappropriate for decision making

3.13 Severe non-beta^0/beta^0 genetic mutations are included in the marketing authorisation for betibeglogene autotemcel. These mutations are associated with dramatically reduced beta-globin production, and behave like a beta^0 genotype despite being grouped with other non-beta^0/beta^0 genotypes. The ERG considered that the proportion the company used in its modelling (from HGB-212, which included patients with severe non-beta^0/beta^0 genotypes) may be too low. It found evidence that severe non-beta^0/beta^0 genotypes represent up to 28% of the population with TDT in the UK. The ERG did scenario analyses that varied the proportion of modelled patients with a severe non-beta^0/beta^0 genotype. It found that, when the proportion of severe non-beta^0/beta^0 genotypes was higher than the figure used by the company, the resulting probability of transplant success decreased. Because of data limitations, the ERG's scenarios were exploratory and not incorporated in its base case. The committee considered that there was not enough evidence available yet on whether or not the severe non-beta^0/beta^0 genotype was underrepresented in the trials compared with the UK population who might have betibeglogene autotemcel. It concluded that it would be inappropriate to include such an adjustment in its decision making.

Iron normalisation and effects of iron overload

One year is insufficient for all modelled patients to reach cardiac iron normalisation after betibeglogene autotemcel

3.14 At technical engagement, the company and the ERG agreed on a 5-year time to normalisation for liver iron. However, they disagreed on time to normalisation for cardiac iron. The company preferred a time period of 1 year for this, based on its clinical trial results at the time of its submission. However, the ERG thought it was possible that people in the trials were not typical of the eligible population. It thought that they may reach cardiac iron normalisation more quickly than a 'typical' patient because of their reduced levels at baseline. It was concerned that a very small number of people in the trials reached the company's described cut-off of a 'normal' cardiac T2* of 40 msec (the number cannot be shown here because it is considered confidential). Also, the company presented results for cardiac iron based on a cut-off for normal levels of 20 msec (note, for this measure, a lower T2* score indicates higher iron levels). The clinical experts explained that, in UK clinical practice, a cardiac T2* of 20 msec is the cut-off used for 'normal' cardiac iron. This is supported by a key review paper on cardiovascular function and treatment in beta-thalassemia major (Pennell et al. 2013). They also said that clinicians do not aim for a T2* of 40 msec. They explained that there will always be a risk of cardiac complications like arrhythmia for people with TDT if they have ever had high cardiac iron (even if asymptomatic at the time). This is because of residual iron which cannot be removed. They also outlined the process for iron loading and removal; excess iron organ deposits are removed from the liver before the heart. The committee noted that all the people in the trials had 'normal' cardiac iron according to UK clinical practice (a T2* of 20 msec). The committee considered that the modelled time to liver iron normalisation should be a minimum of 5 years. This was because there were still some people in the trials whose liver iron had not yet reached the normalisation threshold by 48 months. The committee

took into account that it should take longer for people in clinical practice to reach cardiac iron normalisation because liver iron is removed first, as explained by the clinical experts. Therefore, the time to cardiac iron normalisation should not be shorter than the time to liver iron normalisation. So, the committee concluded that 1 year was insufficient for time to cardiac iron normalisation and that 5 years should have been used in the model.

The approach to modelling complications from iron overload is uncertain

3.15 In its initial model, the company assumed that people who have normalised iron levels are no longer at risk of developing complications from iron overload. The ERG's clinical adviser suggested that there may be pre-existing irreversible damage caused by iron overload in many people who would be eligible for betibeglogene autotemcel (but not severe enough to rule out treatment). They explained that these people could potentially develop complications from pre-existing iron overload damage in the long term (in line with view of other clinical experts on risk of cardiac complications after iron overload; see section 3.14). The ERG uniformly applied an annual cardiac complication rate for people with TDT with 'low' cardiac iron to all people with TI who had had their iron normalised. The company thought this approach would overstate the risk associated with potential irreversible cardiac iron damage. Instead, it proposed to lower the annual cardiac complication rate for people with TI with 'normalised iron' (using the complication rate for low cardiac iron). The ERG thought it was possible that people in the trials were not typical of the eligible population (that is, they represented a more favourable population). The people in the trial might have had a lower risk of cardiac complications, and so mortality, compared with the population that would be seen in clinical practice. The ERG thought the company's revised rate of complications had insufficient evidence to support it. The committee considered the baseline annual mortality risk used by the company, and the populations that it had been derived from. It considered that the

baseline annual mortality risk may have been too high and not applicable to the population eligible for the technology. The committee concluded that the results of the ERG and company approaches were exploratory estimates. Because there was no supporting evidence from the literature, it considered both approaches in its decision making.

Long-term outcomes

Zero mortality from a myeloablative regimen plus betibeglogene autotemcel is clinically implausible

3.16 The company assumed:

- that the engraftment procedure was successful for everyone in the model because there had been no engraftment failures in the betibeglogene autotemcel trials
- no graft loss leading to a return to transfusions for people who did not need transfusions after betibeglogene autotemcel.

The ERG pointed out that the need to collect back-up cells for rescue treatment as part of the betibeglogene autotemcel protocol indicates that risk of engraftment failure exists. The company's submission included 24 people from the trials who were transfusion evaluable. There was also a maximum follow up of 61 complete months of data. The ERG thought this was insufficient to determine whether long-term permanent engraftment occurs in everyone (and its clinical adviser thought it would be too optimistic to assume permanent engraftment for all). One clinical expert explained that graft versus host disease would not be a concern with betibeglogene autotemcel treatment because the infusion uses the recipients' own cells (unlike in HSCT). However, gram-negative sepsis, infections and other complications of myeloablative conditioning represent a non-zero mortality risk. So, there will be some mortality and morbidity (for example, infertility) associated with the betibeglogene autotemcel protocol in clinical practice. As a result, the committee concluded that a

zero mortality assumption for myeloablative conditioning followed by betibeglogene autotemcel was overly optimistic and clinically implausible.

Company's model and cost-effectiveness evidence

The model should be run with a higher number of profiles to get stable ICER values

3.17 The company selected a discretely-integrated condition event (DICE) simulation as the structure for its model. The model simulated a number of 'profiles', or hypothetical patients defined by age and sex. Each profile was weighted to reflect the distribution of patients in the eligible treatment population. The company presented cost-effectiveness results based on analyses with 600 profiles. It stated that increasing the number of profiles also increased the model runtime by an order of magnitude without providing significant additional information about cost effectiveness. The ERG and NICE technical team ran the company's base-case analysis in the model but varied the number of simulated patient profiles (up to 50,000). They found that the ICER stabilised at around 5,000 profiles. The ERG was concerned that the company did not provide an explanation as to why the latter's preferred analysis (600 simulated profiles) generated a lower ICER than analyses run with a higher number of iterations. The ERG explained that stochastic models, such as DICE, typically need a large number of iterations to generate stable results. This is because some variation will appear as a result of 'random noise', or first-order uncertainty. Long runtime for DICE models was known in the literature before this appraisal, and should have been factored into the company's decision to use this modelling approach. The ERG confirmed that increasing the number of profiles would not decrease accuracy, with the only negative impact being an increased model runtime. The committee considered that 600 profiles was not sufficient, and that a higher number of profiles should have been used to get stable ICERs. It concluded that analyses with a higher number of profiles (5,000 instead of 600) should be used in its decision making to get stable ICER values.

Costs of fertility preservation should be included

3.18 Myeloablative conditioning can cause infertility, so fertility preservation is now common in the early stages of the HSCT process. The ERG noted that myeloablative conditioning is also needed as part of the betibeglogene autotemcel process. However, the costs for fertility preservation were not included in the company's model. One patient expert explained that it is usually parents or carers who make the decision for their child with TDT to have HSCT. However, they said that many people with TDT would have made a different decision for themselves as an adult. They explained that, because treatment with betibeglogene autotemcel can be done at a later age than HSCT, being able to make a choice about their own fertility is important. The committee considered that the costs for fertility preservation ahead of treatment with myeloablative conditioning followed by betibeglogene autotemcel should be included in the model.

Cost effectiveness

The company's base case does not reflect the committee's preferred assumptions

3.19 The ICERs in this appraisal are considered to be commercially sensitive by the company, so cannot be reported here. The committee agreed that its preferred approach to modelling would:

- use the reference case discount rate for costs and benefits (3.5%)
- use the ERG's preferred approach to utilities
- limit the UK Chart Review population data to match the population in the clinical-effectiveness data
- set the time to normalisation for cardiac iron and liver iron to 5 years
- incorporate a non-zero mortality associated with myeloablative conditioning followed by betibeglogene autotemcel (there were no data available to run a scenario featuring this assumption, but it would increase the ICER)

- use 5,000 simulated profiles
- include costs of fertility preservation before myeloablative conditioning followed by betibeglogene autotemcel (there were no data available to run a scenario featuring this assumption, but it would increase the ICER).

The ICER using the committee's preferred assumptions considerably exceeded the top end of the range normally considered a cost-effective use of NHS resources. Even when the company's preferred non-reference case discount rate for costs and benefits was used alongside the committee's preferred assumptions, the ICER still exceeded the top end of the range usually considered a cost-effective use of NHS resources.

Equalities

The incidence of a condition in different population groups is outside the remit of a technology appraisal

3.20 In the UK, TDT is mostly seen in ethnic minority populations, the largest groups being in people with Pakistani, Indian or Bangladeshi family background. Some people with TD, and parents or carers of people with TDT, noted the stigma still associated with the condition in particular communities. Some families said they were unable to tell anyone outside of their immediate family for fear that they (patients) or their children (parents) would be victimised or wrongly judged. One patient expert explained that people with TDT are 'seen as a burden to their parents, spouse and the NHS', and that there is a lack of education around the condition in some communities. One clinical expert involved in the appraisal explained that there is a cultural element that is important for this topic because the UK TDT population is increasingly of Asian family background. They also noted that the stigma burden for some patients and their families may not have been captured in the quality-adjusted life years (QALYs). The NHS England commissioning expert submission explained that, if recommended, betibeglogene autotemcel would be

available to a wider age range of people with TDT (that is, people older than 12 years rather than 19 years and younger in the current commissioned policy for stem cell transplantation). This would help address an inequality in the current pathway for people older than 19 years. They explained that the recently set up National Haemoglobinopathies Panel would oversee all referrals to identify any inequalities seen in the people referred or accepted for betibeglogene autotemcel. They pointed out that current care pathways in TDT are not genotype-specific and so are applicable to all people with the condition. Only people with the specific genotype will qualify for betibeglogene autotemcel. This could disadvantage people from ethnic groups who are not eligible for this treatment because they do not have the relevant genotype. The committee was aware of these viewpoints from people with TDT, parents and carers of people with TDT, clinical experts and commissioning experts, when making their decision. It understood that the wording of the licensed indication limits use to people with a non-beta^0/beta^0 genotype. It noted that differences in incidence of a condition in different ethnic groups is outside the remit of a technology appraisal.

Other factors

Betibeglogene autotemcel does not meet end-of-life criteria but is considered innovative

3.21 The committee concluded that NICE's advice about life-extending treatments for people with a short life expectancy did not apply. The technology is considered innovative because it is a potentially curative treatment for people with TDT who cannot have the existing potentially curative treatment, HSCT. It is also the first gene therapy for TDT approved by the European Medicines Agency. The committee was aware that the marketing authorisation presents an important treatment option for people with TDT. It concluded that betibeglogene autotemcel is innovative.

Conclusion

Betibeglogene autotemcel is not recommended

3.22 The committee could not recommend betibeglogene autotemcel, within its marketing authorisation, for treating TDT in people 12 years and older who do not have a beta0/beta0 genotype, for when HSCT is appropriate but a human leukocyte antigen-matched related HSC donor is not available. NICE's guide to the methods of technology appraisal notes that: 'above a most plausible ICER of £20,000 per QALY gained, judgements about the acceptability of a technology as an effective use of NHS resources will specifically take account of…the degree of certainty around the ICER. In particular, the committee will be more cautious about recommending a technology when they are less certain about the ICERs presented'. Betibeglogene autotemcel did not meet NICE's end-of-life criteria but was considered innovative, so an ICER closer to £30,000 per QALY gained was considered acceptable. Above a most plausible ICER of £30,000 per QALY gained, the guide to the methods of technology appraisal notes that an increasingly stronger case will need to be identified for supporting the technology as an effective use of NHS resources. The most plausible ICER for betibeglogene autotemcel was considerably higher than £30,000 per QALY gained and associated with a high level of uncertainty relating to longer-term clinical effectiveness. Therefore, the committee could not recommend betibeglogene autotemcel.

4 Proposed date for review of guidance

4.1 NICE proposes that the guidance on this technology is considered for review by the guidance executive 3 years after publication of the guidance. This allows for consideration of data from bluebird bio's long-term follow-up study, LTF-303. NICE welcomes comment on this proposed date. The guidance executive will decide whether the

technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Gary McVeigh
Chair, appraisal committee
January 2021

5 Appraisal 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 D.

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

The minutes of each appraisal committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

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.

Amy Crossley
Technical lead

Caron Jones
Technical adviser

Kate Moore
Project manager

ISBN: [to be added at publication]