Atidarsagene autotemcel for treating metachromatic leukodystrophy

Highly specialised technologies guidance

Published: 28 March 2022

www.nice.org.uk/guidance/hst18

Your responsibility

The recommendations in this guidance represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, health professionals are expected to take this guidance fully into account, alongside the individual needs, preferences and values of their patients. The application of the recommendations in this guidance is at the discretion of health professionals and their individual patients and do not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

All problems (adverse events) related to a medicine or medical device used for treatment or in a procedure should be reported to the Medicines and Healthcare products Regulatory Agency using the Yellow Card Scheme.

Commissioners and/or providers have a responsibility to provide the funding required to enable the guidance to be applied when individual health professionals and their patients wish to use it, in accordance with the NHS Constitution. They should do so in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities.

Commissioners and providers have a responsibility to promote an environmentally sustainable health and care system and should assess and reduce the environmental impact of implementing NICE recommendations wherever possible.

Contents

1 Recommendations

1.1 Atidarsagene autotemcel is recommended, within its marketing authorisation, as an option for treating metachromatic leukodystrophy with mutations in the arylsulphatase A (ARSA) gene:

- for children who have late infantile or early juvenile types, with no clinical signs or symptoms

- for children who have the early juvenile type, with early clinical signs or symptoms, and who can still walk independently and have no cognitive decline.

  It is recommended only if the company provides atidarsagene autotemcel according to the commercial arrangement.

1.2 Atidarsagene autotemcel should be delivered in a highly specialised service by a specialist multidisciplinary team.

Why the committee made these recommendations

Metachromatic leukodystrophy is a genetic condition that affects the central nervous system. It has a significant effect on the quality of life of children with the condition, and their families and carers. It progresses rapidly, with loss of mobility and cognitive function, and causes early death. Treatment options are limited to managing symptoms and supportive care.

Clinical evidence suggests that the gene therapy atidarsagene autotemcel improves mobility and cognitive function and could correct the enzyme deficiency caused by the condition. But how well atidarsagene autotemcel works in the long term is uncertain.

The cost-effectiveness estimates show that atidarsagene autotemcel provides substantial extra health and quality-of-life benefits. But how much is uncertain, and it varies for the different types of the condition. Taking into account the long-term uncertainty, for children with late infantile and early juvenile forms of the condition, the cost-effectiveness estimates are within what NICE normally considers an acceptable use of NHS resources for highly specialised technologies. So, it is recommended for these children.

2 The condition

2.1 Metachromatic leukodystrophy (MLD) is an autosomal recessive genetic disorder, caused by a deficiency in the enzyme arylsulphatase A (ARSA). This deficiency causes sulphatides to accumulate, producing microglial damage, progressive demyelination and neurodegeneration, leading to neurological problems. MLD is a progressive and chronically disabling condition, which substantially reduces quality of life and life expectancy. MLD can broadly be divided into a presymptomatic stage with normal motor and cognitive development, followed by a developmental plateau and early onset of first symptoms. There are 3 main types based on genotype and age of symptom onset:

- The late infantile (LI) type is characterised by 2 null alleles (0/0 genotype). It is the most common (40% to 60% of children affected) and most aggressive form and usually starts before 30 months. Symptoms include peripheral neuropathy, muscle weakness, sight and hearing loss, difficulty walking, loss of speech, cognitive decline, and seizures. The condition progresses rapidly so that children lose awareness of their surroundings over a few years. Death normally occurs within 5 to 8 years.

- The juvenile type is characterised by either 1 null allele and 1 residual allele (0/R genotype) or, less frequently, 2 residual alleles (R/R genotype). About 20% to 35% of children affected have this type. Symptoms include impaired fine motor skills and concentration, behavioural problems, difficulties with movement, slurred speech, incontinence and seizures. Initial disease progression is slower than with the LI type but symptoms can progress rapidly. Death normally occurs within 10 to 20 years. It can be subdivided into:

  — early juvenile (EJ) disease, starting between 30 months and 6 years

  — late juvenile disease, starting between 7 and 16 years.

- The adult type (15% to 25% of people affected) is the rarest form and usually starts after 16 years. Symptoms include a decline in school or work performance, cognitive decline, personality changes and memory lapses. The decline can be slow and almost imperceptible. Death normally occurs within 25 years.

2.2   The prevalence of MLD is estimated as 1 in 147,000 live births in England and Wales, equating to about 4 to 5 children born with MLD per year.

2.3   Timely diagnosis of MLD may be challenging, particularly in families without a previous history of the condition. Generally, the LI type is identified because children are unable to meet a major motor development milestone, whereas children with EJ MLD may initially have some cognitive or coordination changes. Adults with MLD mainly present with dementia-like symptoms. Tests include brain MRI, and blood and urine tests to detect sulphatides. When MLD is confirmed, genetic testing is done to identify the specific mutation. Diagnosis of the later-onset forms may take longer than for the LI type because of their non-specific signs and symptoms.

2.4   MLD is managed in the NHS by neurodisability and metabolic consultants at regional centres, who advise local hospital and community-based teams. There are 3 paediatric lysosomal storage disorders specialist centres in England providing multidisciplinary treatment led by a paediatric metabolic consultant at:

- Birmingham Children's Hospital

- Great Ormond Street Hospital, London

- St Mary's Hospital, Manchester.

    It is expected that atidarsagene autotemcel would be delivered within the current specialised services framework by a multidisciplinary team.

3 The technology

3.1 Atidarsagene autotemcel (Libmeldy, Orchard Therapeutics), referred to as OTL-200 in the company submission, is a gene therapy medicinal product that expresses the human arylsulphatase A (ARSA) gene. Ex vivo autologous CD34+ haematopoietic stem and progenitor cells are collected from the person's bone marrow or peripheral blood. These are then transduced with a lentiviral vector, which inserts copies of human ARSA complementary DNA into the cell genome, and transplanted into the person. When successfully engrafted, the genetically modified cells secrete functional ARSA enzyme, which is absorbed by surrounding cells and used to break down or prevent build-up of harmful sulphatides. The effects of OTL-200 are potentially lifelong.

3.2 The marketing authorisation indication for atidarsagene autotemcel is for 'metachromatic leukodystrophy (MLD) characterized by biallelic mutations in the arylsulphatase A (ARSA) gene leading to a reduction of the ARSA enzymatic activity:

- in children with late infantile or early juvenile forms, without clinical manifestations of the disease

- in children with the early juvenile form, with early clinical manifestations of the disease, who still have the ability to walk independently and before the onset of cognitive decline'.

3.3 Atidarsagene autotemcel is administered as a single-dose intravenous infusion. The minimum recommended dose is 3×10^6 CD34$^+$ cells per kg. The product consists of 1 or more infusion bags containing 2 to 10×10^6 cells per ml suspended in a cryopreservative solution. A myeloablative conditioning regimen is needed before infusing atidarsagene autotemcel, to promote engraftment of the genetically modified cells. Before starting myeloablative conditioning, the treating clinician should confirm that atidarsagene autotemcel is clinically appropriate for the patient. Atidarsagene autotemcel must be administered in a qualified treatment centre with experience in haematopoietic stem cell transplantation as detailed in the summary of product characteristics for atidarsagene autotemcel.

3.4   Adverse reactions because of myeloablative conditioning or atidarsagene autotemcel include febrile neutropenia, stomatitis, mucosal inflammation, veno-occlusive disease and anti-ARSA antibodies. For full details of adverse reactions and contraindications, see the summary of product characteristics for atidarsagene autotemcel.

3.5   As a condition of the marketing authorisation for atidarsagene autotemcel, the company has to undertake a long-term study to assess both the efficacy and safety of atidarsagene autotemcel for children in its clinical development program and in the post-authorisation setting, for up to 15 years after treatment. The long-term study will also collect data from an additional 10 children with early symptomatic EJ MLD.

3.6   The list price for atidarsagene autotemcel is £2,875,000 (excluding VAT; company submission). The company has a commercial arrangement. This makes atidarsagene autotemcel available to the NHS with a discount. The size of the discount is commercial in confidence. It is the company's responsibility to let relevant NHS organisations know details of the discount.

4   Consideration of the evidence

The evaluation committee considered evidence submitted by Orchard Therapeutics, the views of people with the condition and those who represent them, clinical experts, NHS England and a review by the evidence review group (ERG). See the committee papers for full details of the evidence. In forming the recommendations, the committee took into account the full range of factors that might affect its decision, including in particular the nature of the condition, the clinical effectiveness, value for money and the impact beyond direct health benefits.

After the first meeting, the committee considered that it had not been presented with all the necessary analyses for decision making. The company provided an additional 2 years of follow-up data for 17 patients and updated analyses. The ERG also provided updated analyses and a critique of the company's additional information. The additional data and analyses were considered at the second committee meeting.

Nature of the condition

Effect of MLD on patients and their families and carers

4.1 The patient and clinical experts explained that metachromatic leukodystrophy (MLD) is a life-limiting, relentless, disabling and isolating condition, affecting all aspects of patients' and carers' lives. The patient experts told how MLD affects people with the condition, including progressive loss of their ability to sit, stand, walk, talk, see, hear and swallow. The ability to walk or talk can be lost overnight. They explained that living with MLD can be an unrelenting cycle of shock, fear, anxiety, desperation, grief and bereavement, with each further loss of function bringing new distress. In the later stages of the condition, people can develop painful spasticity, epilepsy, dementia, breathing problems, double incontinence and complex gastrointestinal dysfunction. Suctioning and multiple medications, which often need adjusting, are needed to help manage rapid disease progression. The clinical experts explained that spasticity, gastrointestinal dysfunction and intolerance to different feeding methods can present challenges for care. The patient experts explained that the suffering of people with MLD and the burden on families, including unaffected siblings, are immeasurable. They explained that people with MLD can become completely dependent, needing 24-hour care provided by 1 or 2 adults. They highlighted how the strain on carers negatively affects their quality of life and can be:

- physical (lifting and handling, chronic exhaustion)

- psychological (grief, worry, insomnia, chronic depression)

- financial (not being able to work)

- social (relationship breakdown).

  The committee concluded that MLD is a rare, serious and life-limiting condition that significantly affects the lives of people with the condition, and their families and carers.

Unmet need

4.2   The clinical experts explained that best supportive care is the main treatment for managing MLD symptoms. This can include:

- managing muscle spasms, infections, seizures or secretions

- pain relief or sedative drugs

- feeding support (including gastrostomy)

- psychological and social support (including specialist schooling)

- genetic advice and planning

- end of life care.

    The patient experts emphasised that there is an unmet need for effective disease-modifying treatments for MLD. They highlighted that atidarsagene autotemcel (referred to as OTL-200 in the company submission) can be life transforming, especially when offered early before symptoms appear. It could offer substantial benefits to people with MLD and their families. The committee recognised that treatment options are limited, and that there is a significant unmet need for disease-modifying therapies for MLD. It concluded that people with the condition and their families would welcome OTL-200 as an option for treating MLD.

Diagnosis

4.3   The classification system used to diagnose MLD is based on genotype and the age when symptoms appear. MLD type is a predictor of disease progression (see section 2.1). At the first committee meeting, the NHS England representative confirmed that routine MLD screening for newborn babies is not available in England and is unlikely to be introduced in the next 5 years. They explained that when a child has been diagnosed with MLD, other siblings can have genetic testing. The patient experts highlighted how difficult it is to make the initial diagnosis if there is no sibling with MLD. It may take on average 11 months for the late infantile (LI) type and 11 to 13 months for the early juvenile (EJ) type to be diagnosed because of inaccurate diagnoses and inappropriate referrals. The delays and uncertainties can cause anxiety for families. A clinical expert explained that people referred to a lysosomal storage disorders centre are usually seen by a specialist within a week. The clinical experts emphasised that an early diagnosis before the onset of symptoms is important and that the Inherited White Matter Disorders Service should help to speed up diagnosis. The patient and clinical experts highlighted that patient organisations are campaigning to have newborn screening introduced across all inherited metabolic disorders, including MLD. The committee recognised the difficulties with diagnosis in rare conditions such as MLD, particularly if there is no sibling with the condition.

Current treatment

4.4 The clinical experts explained that there are no effective disease-modifying treatments available in the NHS for MLD. Historically, haematopoietic stem cell transplantation (HSCT) was used, usually for people who were presymptomatic and who have late juvenile MLD. Over the past 10 years, clinicians have instead enrolled patients in OTL-200 trials. The clinical experts emphasised that even if there were no OTL-200 trials, HSCT is unlikely to be used because of poor outcomes and its potential to accelerate the condition. Best supportive care, the main treatment for MLD (see section 4.2), involves multidisciplinary care in partnership with local services. The clinical experts noted that local services are generally underfunded. The patient experts emphasised that because MLD progresses rapidly, delays between assessment and providing equipment may mean that the equipment is no longer appropriate. The clinical experts disagreed about the degree to which the lysosomal storage disorders specialist centres could ensure timely support. One clinical expert highlighted that people are often not referred to specialist centres because local clinicians do not think that treatment options are available. The committee acknowledged that HSCT is unlikely to be used for people with MLD. It recognised that effective treatment options are limited and that a dedicated service may help provide timely care and support to people with MLD and their families. It agreed that best supportive care is the relevant comparator in this evaluation.

Impact of the new technology

The population

4.5   The company submitted evidence for both groups covered by OTL-200's marketing authorisation:

- children who have LI or EJ types, with no clinical signs or symptoms

- children who have the EJ type, with early clinical signs or symptoms, and who can still walk independently and have no cognitive decline.

The company defined the first group as children with presymptomatic (PS)-LI and PS-EJ types. The company defined the second group as children with the early symptomatic (ES)-EJ type, who:

- can walk independently as shown by a Gross Motor Function Classification in MLD (GMFC-MLD) score of 0 (walking without support with normal performance for age) or 1 (walking without support but with reduced performance, that is, instability when standing or walking)

- have no cognitive decline, as shown by an intelligence quotient (IQ) of 85 or more.

The company explained that OTL-200 is most effective before disease progression. So over time, it had updated the definition of early symptomatic to identify children who are likely to benefit from OTL-200. The clinical experts stated that they have been identifying and assessing children for eligibility for OTL-200 studies for the past 10 years. They explained that in the absence of newborn screening or known family history, most children with MLD are diagnosed because they have symptoms (see section 2.3 and section 4.3). So, most patients would not be eligible for OTL-200 unless an older sibling was diagnosed before their symptoms started. The clinical experts considered that most children with the ES-EJ type would likely need a sibling diagnosed with the condition to be eligible for OTL-200, because the time to diagnosis is slow, and the condition progresses rapidly. However, they noted that the time to diagnosis could decrease if an effective treatment were available. They explained that, theoretically, OTL-200 could be offered immediately if the condition is detected in a newborn. The committee was concerned about the practicality of applying the eligibility criteria in the marketing authorisation for children in the ES-EJ subgroup. It noted the difficulty in diagnosing MLD before disease progression, and the need for an older sibling to be diagnosed first unless newborn screening for MLD becomes available.

Clinical evidence

4.6 The company submitted evidence for fresh and cryopreserved formulations of OTL-200. Study 201222 (fresh formulation, main registration trial; n=20) was a non-randomised, open-label, prospective, single-centre trial evaluating the efficacy and safety of OTL-200 in children with LI or EJ MLD. It measured:

- motor function using the Gross Motor Function Measure (GMFM) and GMFC-MLD

- biological markers of ARSA enzyme activity in both the peripheral blood and cerebrospinal fluid

- change in neurocognitive function using developmental quotient (DQ)

- change in neurological function using brain MRI

- stability of nerve conduction using nerve conduction velocity

- overall survival.

  The study's co-primary end points were:

- an improvement of at least 10% in total GMFM score compared with an untreated historical control MLD population (best supportive care)

- a statistically significant increase in residual ARSA enzyme activity by at least 2 standard deviations compared with pre-treatment values, measured in peripheral blood mononuclear cells at year 2 after treatment.

4.7 Expanded access programmes (fresh formulation; n=9) consisted of 1 hospital exemption (HE 205029) and 2 compassionate use programmes (CUP 207394 and CUP 206258). These studies were done at the same site and by the same staff as study 201222, and when appropriate, followed its design.

4.8 For the second committee meeting, the company provided an additional 2 years of data for 17 patients from the main registration trial and CUP 207394 (up to December 2019).

4.9   Study 205756 (cryopreserved formulation; n=10) was an open-label, single-arm study in children with presymptomatic early onset MLD (LI, EJ, or an intermediate variant between LI and EJ). In response to clarification, the company provided additional data up to November 2019 for 4 patients who had cryopreserved OTL-200.

Response to treatment

4.10   The company did a naive comparison with a natural history cohort of 31 patients with untreated MLD enrolled since 2004 and, when possible, a comparison with a matched sibling. The ERG had concerns about the evidence; specifically the lack of baseline data from the natural history comparator cohorts and the gaps in OTL-200 baseline data. The ERG could not do any statistical analyses or verify the comparisons with the natural history cohorts. However, the committee considered that the evidence from the natural history cohorts showed that patients had very poor outcomes. Most had complete loss of movement and head and limb control (GMFC-MLD 6) and no cognitive function within a few years of diagnosis. However, for patients who had OTL-200, almost all had much better clinical outcomes. In its submission, the company considered that less than half of the patients showed a long-term treatment effect with normal motor and cognitive function (full response; see section 4.11). Other patients also showed a long-term effect on motor function without reaching the lowest GMFC-MLD classification states. The company considered this could be either long-term stabilisation or slower progression through the GMFC-MLD states than in the natural history cohort. The clinical expert explained that in the natural history cohort, DQ scores go down to 0, a state of no cognitive function or abilities. They explained that DQ scores should be about 100 in normal development. The committee noted that although there were fluctuations in DQ scores in the subgroups having OTL-200, scores were generally high and did not fall to cognitive impairment levels. In the company's latest data cut (December 2019), the ERG noted that some patients showed a decline in cognitive function. The ERG highlighted that the large fluctuations in individual patient profiles made interpretation difficult. The committee noted that the company used motor and cognitive function as the main outcomes to measure the clinical benefit of OTL-200. But there were other outcomes important to patients that had not been assessed in the studies, such as spasticity and quality of life. The committee noted that all patients in the 'full response' category had the presymptomatic LI or EJ types. It noted the differences in response trajectory between the presymptomatic and ES-EJ types. Comments from consultation also noted expected differences in kinetics and extent of treatment response for ES-EJ types. The committee concluded that when OTL-200 was effective, it had a substantial clinical benefit on both motor and cognitive function compared with the natural history cohort. It agreed that children who had OTL-200 could retain cognitive function, even if motor function declines.

Interpretation of treatment response

4.11   The company considered that OTL-200 could be effective for a person's lifetime because the progeny of the infused cells maintain the gene correction. However, successful engraftment and migration of cells into the central nervous system could take up to 2 years, so the condition could progress before there is a treatment effect. Therefore, the company proposed a classification system to identify the initial response and disease course for each person, including long-term stabilisation of disease symptoms. At the first committee meeting, the company suggested that GMFC-MLD score was the most appropriate outcome on which to base this classification and considered 3 categories:

- Full response: people had treatment before symptom onset and symptoms remained stable with motor and cognitive function fully intact. The company assumed that they remained in GMFC-MLD 0 for the full time horizon and led normal healthy lives in line with the general population.

- Stable partial response: people either had treatment after symptom onset (GMFC-MLD more than 0) and then stabilised, or had some progression after treatment but then stabilised in GMFC-MLD 1 or 2 (based on trial data and clinical expert opinion).

- Unstable partial response: treatment failed to stabilise the condition. People progressed through GMFC-MLD states but at a slower rate than patients having best supportive care (calculated compared with the natural history cohort and expert elicitation).

    The ERG considered that there was a biological rationale for full and partial response and for late stabilisation, but that the GMFC-MLD did not capture all the clinical signs and symptoms of the condition. The ERG was concerned that the difference between stable and unstable partial response was not clear. Also, some people could potentially stabilise in states with lower function than GMFC-MLD 2. The ERG considered that the classification criteria should be agreed in advance to prevent bias in interpretation. The ERG considered that patients in full response should remain in GMFC-MLD 0 for at least 12 months of follow up. It also considered that patients in stable partial response should show a decline in GMFC-MLD only in the first 12 months of treatment.

4.12   At the second committee meeting, the company responded by updating its
        classification system to include other clinical outcomes including GMFM, DQ, MRI
        and nerve conduction velocity to provide a more holistic interpretation of response:

- Full response: motor and cognitive function remained stable throughout the
  follow-up period, that is, no disease progression was observed throughout the
  follow-up period.

- Stable partial response: motor and cognitive function appeared to have
  stabilised after an initial period of worsening. To determine the GMFC-MLD
  level the person stabilised at, the following were considered:

  - DQ, MRI and nerve conduction velocity should have stabilised or continued
    to improve for 12 months.

  - The GMFM total score or relevant subdomain should be stabilising.

- Unstable partial response: a consistent trend of worsening in motor (GMFM
  and GMFC-MLD) and/or cognitive function, albeit at a slower rate than in the
  natural history cohort.

  Disease progression was defined as a worsening in motor impairment and/or
  cognitive function:

- Progression of motor impairment: worsening of GMFC-MLD and GMFM total
  score.

4.13   The committee was unclear about the definition of stability for all the clinical outcomes and how they individually contributed to the categories. The clinical experts explained that good engraftment can be shown by measurable ARSA enzyme activity in peripheral blood mononuclear cells and cerebrospinal fluid. But it is difficult to determine how long is needed to confirm stabilisation and exactly how much data is needed to be confident of a person's disease course. They noted that it is difficult to define what stabilisation means at this stage of child development for GMFC-MLD scores because scoring even in healthy children may vary for reasons unrelated to the condition. However, they considered that a 'flat line' in GMFC-MLD compared with the natural history cohort could be considered as stable. The clinical experts noted that sometimes there is a decline in a gross motor score after years of stabilisation. This may not necessarily be central nervous system deterioration because of a change in ARSA enzyme levels. It may be because of damage that occurred before treatment could take effect. For example, changes from a pre-existing abnormality of tone or power that become more obvious as the child grows, or progression over time of spasticity, or both. One clinical expert considered that people have stable engraftment at different levels, typically progress in the first 2 to 3 years after treatment and then stabilise at a specific GMFC-MLD level. However, some children who showed some decline later in the condition would not likely have lost engraftment, rather this would be related to long-term deterioration associated with secondary complications of the condition.

4.14   The committee considered that the company's and ERG's revised response classifications, taking account of other outcomes besides GMFC-MLD, provided a better framework for assessing response. However, it noted substantial uncertainty with predicting trajectories of disease progression for any of the response categories. It also noted that the economic model was based on these categories (see section 4.18) and was very sensitive to changes in classification. The committee considered that the ERG's response categories more easily allowed separation of treatment response from durability of treatment, which provided a more robust and transparent framework to explore assumptions around stability of treatment effects. It also considered that including cognitive outcomes in the categorisation of response was more appropriate for estimating outcomes than the company assumptions. It concluded that the ERG model structure was more appropriate. The committee noted that the ERG's classification of patients had the same unavoidable fundamental flaw as the company's classification of applying response criteria in a post hoc manner to observed data. It considered that classification of patients into response categories is still highly subjective with both sets of criteria, and subject to substantial uncertainty. However, it noted that the differences between the company and ERG cost-effectiveness results with each set of response criteria and classification were minimal when taking into account other committee assumptions.

Generalisability

4.15   The committee noted that although 29 people were recruited to the OTL-200 studies, only 25 people were included in the company's efficacy analysis. The company explained that 4 people were excluded from the post hoc analysis because they did not meet the eligibility criteria in the marketing authorisation. One person was in the PS-LI subgroup but had symptoms at treatment. Three people were in the ES-EJ subgroup and had treatment after they had entered a rapid disease progression phase. The ERG considered that:

- There was 1 person with ES-EJ MLD who met the marketing authorisation criteria (GMFC-MLD and IQ thresholds) at treatment and should have been included in the post hoc efficacy analysis. The company explained that this person's symptoms had progressed between assessment and treatment and so they would not have been eligible for treatment in line with the marketing authorisation.

- There was 1 person with ES-EJ MLD who could have been considered as having a borderline IQ threshold. The IQ test is not precise, and so they should have been included in the analysis. The company explained that the eligibility criteria had been updated over the past 10 years to identify people who are likely to benefit from OTL-200 (see section 4.5).

- Of the 5 people with ES-EJ MLD included, 2 did not represent the typical EJ natural disease course because they had treatment when they were over 7 years old (GMFC-MLD score of 0 or 1). At this age, most people in the natural history EJ cohort had progressed to the lowest GMFC-MLD state (GMFC-MLD 6). The ERG queried whether these 2 people had a disease course more similar to slow progressing late juvenile type.

Clinical outcomes

4.16   The committee considered that motor and cognitive function were appropriate
clinical outcomes to measure a patient's response and progression (see
section 4.10). However, the patient experts explained that overall quality of life was
not wholly captured by these measures. They considered that other outcomes
were important, such as preserving the ability to eat, continence and
communication. One patient expert noted that many children who had OTL-200
had older siblings who did not have treatment. Although many children who had
OTL-200 were alive and well, siblings who had not had treatment had died or were
very debilitated. The committee commended the patient organisations for the
submissions providing detailed feedback from a survey on the effect of OTL-200
on quality of life. The company did not collect health-related quality-of-life data in
its studies so some of these additional outcomes were not captured in the
analyses. The ERG considered that the analyses did not fully capture the
differences between clinical outcomes and health-related quality of life of children
with MLD. One patient expert noted that there was little correlation between nerve
conduction velocity scores and clinical outcomes. The clinical experts explained
that OTL-200's effect on the peripheral nervous system seems to be slower than
on the central nervous system, but the underlying mechanisms for these
differences are not understood. The committee acknowledged that biological
markers may not necessarily correlate with clinical outcomes. It considered that all
outcomes would be taken into account when evaluating OTL-200's response.

Cryopreserved formulation

4.17   The company highlighted that the European Medicines Agency considered the fresh and cryopreserved formulations to be comparable. The company emphasised that similar cerebrospinal fluid ARSA enzyme activity was seen at day 19 and at 1 year for both formulations. The ERG noted that comparability data from 4 people who had the cryopreserved formulation were limited (see section 4.6). The clinical expert agreed, but considered there is no reason that the cryopreserved formulation would be inferior to the fresh formulation. The committee noted that the European Medicines Agency accepted this, but the committee did not consider there was enough evidence to confirm that both formulations are equivalent. The company explained that all 10 patients have now been recruited to the cryopreserved formulation study, but no new data is available. However, feedback from clinicians suggests that peripheral engraftment in the first few months happens at the same rate with both formulations. The committee considered that there was some uncertainty about potential differences between the fresh and cryopreserved formulations. This was because of the lack of evidence for the cryopreserved formulation, which will be used commercially.

Cost to the NHS and value for money

The company's economic model

4.18 The company submitted a Markov model approximating a partition survival model to compare the cost effectiveness of OTL-200 with best supportive care (natural history cohort). This provided incremental cost-effectiveness ratios (ICERs) for individual subgroups (PS-LI, PS-EJ and ES-EJ) and for the whole population (pooled). The model consisted of 8 health states (7 GMFC-MLD health states [GMFC-MLD 0 to GMFC-MLD 6] and death), a monthly cycle length and a lifetime time horizon. Patients progressed through the model depending on whether they had best supportive care or had OTL-200 and were categorised as having a full response, a stable response or an unstable partial response in the company's base case (see section 4.11). Patients could only become progressively worse, that is, they were only allowed to move to higher GMFC-MLD health states. For the EJ subgroups only, the company included treatment-dependent cognitive impairment (DQ) substates. The starting ages were 18, 45 and 80 months for the PS-LI, PS-EJ and ES-EJ groups respectively. About half of the population were male. For the PS-EJ and ES-EJ subgroups, 20% of patients having best supportive care were considered to start with moderate cognitive impairment compared with no patients having OTL-200 (see section 4.19). The ERG considered that a lifetime time horizon is appropriate given that OTL-200 is a potential cure. However, it noted that input parameters for children were extrapolated to adults and that short-term effectiveness evidence was projected over a very long period, increasing uncertainty in the results. It also noted that the concept of stabilisation was difficult to validate because the model structure is based on categorising and extrapolating unique response patterns seen in very few people. Limited follow up also increased uncertainty in the results. The ERG also provided some revisions to the model structure in response to consultation that included an additional response group, and integration of cognitive impairment into the response categorisation (see section 4.12).

Assumptions after the first committee meeting

4.19   After the first committee meeting, the company incorporated some of the ERG's preferred assumptions or corrections in its model, as follows:

- The ERG highlighted that the company's assumption that more patients on best supportive care would start with moderate cognitive impairment was unjustifiable (see section 4.18). Also, baseline differences would introduce bias. The ERG amended the baseline characteristics to ensure consistency across arms.

- The ERG noted that the time spent in GMFC-MLD 0 in the company's model was inconsistent with the observed data. The ERG re-estimated the time spent in GMFC-MLD 0 using the company's reported starting ages and data from the natural history study.

- The company assumed general population levels of all-cause mortality in all health states (GMFC-MLD 0 to 5) except GMFC-MLD 6. The company's assumption meant that there would be no mortality risk from MLD until GMFC-MLD 6. The ERG corrected implementation errors in the company's parametric survival analysis of the natural history cohort to estimate risk of death over time while in GMFC-MLD 6.

- The ERG considered that in stable and/or unstable partial response, mortality would be associated with lifelong neurodisability. The ERG included standardised mortality ratios for GMFC-MLD 1 to 5, informed by values applied in NICE's highly specialised technologies guidance on cerliponase alfa for treating neuronal ceroid lipofuscinosis type 2 (HST12).

- The ERG considered it appropriate to model a 1.25-increase in long-term mortality associated with having myeloablative conditioning. This was informed by NICE's appraisal of betibeglogene autotemcel for treating transfusion-dependent beta-thalassaemia.

- The company assumed that no carers were needed until GMFC-MLD 5, when 2 carers were needed. The ERG considered that carers would be needed from GMFC-MLD 1 (0.5 carers) to GMFC-MLD 6 (2 carers).

- The company adjusted utilities for patients' ages only in GMFC-MLD 0 with normal cognition or mild cognitive impairment. The ERG corrected the use of the predictive equation and applied it for all patients regardless of GMFC-MLD state.

Progression modifiers

4.20   The company modelled the unstable partial response group to progress at a rate that was a multiplier of best supportive care. For the LI subgroup, this multiplier was calculated using the OTL-200 evidence and natural history cohort. In the company's original submission, there was not enough evidence to calculate these multipliers for the EJ subgroups, so the company used values from clinical expert elicitation. At the second committee meeting, the company used the additional data from the patients who had progressed at the latest data cut to calculate the progression modifiers for the ES-EJ subgroup. Using data from the ES-EJ patients who had progressed between GMFC-MLD 2 and 3, the company calculated the average time to progression to be more than 5 times longer than in the natural history cohort. The ERG considered that this may be inappropriate because it included 2 patients whose condition may not represent the EJ disease course (see section 4.15). The ERG considered that in principle, the progression modifiers should be based on data from specific subgroups, but there was not enough evidence to populate this in the model. Therefore, it preferred to use the progression modifiers calculated for the LI population. The company considered that the progression modifiers used in the model for OTL-200 between GMFC-MLD 0 to 1 and GMFC-MLD 1 to 2 were implausible. The ERG highlighted that these progression modifiers were taken directly from the company's original model. In response to consultation, the ERG accepted changes that limited progression modifiers to be equal to the natural history cohort for these states. The committee acknowledged the lack of data on which to base progression modifiers across different GMFC-MLD health states and across various subgroups. It noted that this was an important source of uncertainty and there was probably not enough data to be certain of the true values. It concluded that the progression modifiers calculated from the LI population were likely to be the most appropriate. This was because of the limited evidence available for the EJ population and concerns about the generalisability of that evidence.

Stabilisation of treatment response

4.21   The company's interpretation of response relied on assumptions about long-term stabilisation of disease symptoms in the economic model (see section 4.11). After consultation, the company assumed in its revised base case that stabilisation occurred for an average of 50 years. This was based on OTL-200's supposed mechanism of action, which supports long-term stabilisation (see section 3.1). The company noted that HSCT has shown an ongoing lasting effect for metabolic disease beyond 30 years and has been used for over 50 years to successfully treat other conditions. The ERG considered that stabilisation assumptions should be based on OTL-200 evidence rather than inferred from technologies used for other conditions. It considered that the additional 2 years of data from the company's updated data cut showed that some patients' motor function declined even after periods of apparent stabilisation (2 to 3 years). The ERG also maintained that the updated data cut showed continued decline in cerebrospinal fluid ARSA enzyme activity to an average of the lower limit of the normal range. The ERG considered that the observed rate of decline would mean the company's scenario was likely to be overly optimistic. As part of the ERG's updated model structure, response categories allowed for separate consideration of stability assumptions (see section 4.12). The ERG noted the difficulty in accurately estimating an appropriate rate of progression with the current stabilisation evidence and the GMFC-MLD health state in which patients are likely to stabilise. At the third committee meeting, the clinical experts noted that there was no new evidence on OTL-200 about stabilisation of treatment effects. However, they reiterated evidence from HSCT for other similar indications and NICE's highly specialised technologies guidance on Strimvelis for treating adenosine deaminase deficiency–severe combined immunodeficiency (HST7), which suggest patients have stable grafts 30 and 20 years, respectively, after transplantation. They also highlighted the variable views about interpreting brain MRI data to confirm clinical stabilisation.

4.22   The committee considered the possibility that some people would stabilise in lower GMFC-MLD states, which would substantially reduce OTL-200's modelled treatment benefit. It recalled that even with stable engraftment, people's conditions could deteriorate because of secondary complications of the disease rather than failure of treatment (see section 4.13). This meant that further loss of stabilisation may be possible, even without loss of engraftment, and this was not accounted for in the model with an average of 50 years' stability. The committee noted a recent publication that showed a lack of long-term benefit of HSCT for people with Fabry disease, but also noted the differences in disease type and treatment mechanism. The committee noted that the cost-effectiveness estimates were highly sensitive to the average length of stabilisation applied in the various scenarios, although this effect was restricted to those with a stable response in the updated ERG model structure. It considered that it was unlikely that collecting further data in the short term would reduce this uncertainty. The committee considered that the most plausible period of stabilisation would be an average of 20 years, to account for HSCT evidence from other plausibly similar treatments, although this assumption was highly uncertain.

Health valuation study

4.23   The company did not collect any EQ-5D data in its OTL-200 studies. But it commissioned an elicitation study to generate health state utilities using vignettes and time trade-off exercises with the general public. The ERG had several concerns about this study. The study design did not follow NICE's reference case because it directly modelled public preferences with no explicit consideration of the patients' quality of life. This was a problem when the public considered cognitive impairment outside the context of a condition affecting children such that many participants chose extreme values for cognitive impairment. It also considered that the results lacked face validity; more challenging health states were rated as better than less challenging health states. Also, the results lacked external validity compared with utility values used in other appraisals, for example utility values that were lower than the EQ-5D worst health state. The ERG considered that the content and construction of the vignette descriptions were inconsistent. The committee concluded that the elicitation study had serious methodological limitations. It would have preferred the company to follow the ERG's suggestion of using clinical experts as proxies for patients to derive utilities for each health state (as done in HST12). In response, the company supplied alternative utility value sets (see section 4.24).

Utility values

Page 34 of 45

4.24   At the second committee meeting, the company provided a utility set using a linear regression model. This rescaled the negative utility values so that no value was lower than the lowest possible EQ-5D utility value from the time trade-off exercise for the EJ subgroups. These were applied to the normal cognition and the moderate or severe cognitive impairment health substates. The company also did a second scenario in which it used a 'top-up' health-related quality-of-life increment for OTL-200 patients only, for retained cognitive function not captured by loss of motor skills (see section 4.10). The company suggested that patients who had OTL-200 in GMFC-MLD 3 to 6 had additional benefits beyond GMFC-MLD scores, for example, improved cognitive function, no swallowing or feeding problems, reduction in seizures and bowel and bladder problems, and improved vision. The company used this second 'top-up' set in its base case. However, at the second meeting, the company highlighted that it preferred to use the rescaled utility set. The ERG considered that the rescaled utility values were more appropriate and resolved some of the face validity issues (see section 4.23). However, the ERG continued to use its utility set because this maintained a negative utility value in patients without cognitive decline in the lowest GMFC-MLD health states. The ERG considered the 'top-up' utility values were inappropriate for decision making because there were no negative utility states, which did not reflect the evidence. The patient and clinical experts emphasised the poor quality of life that people have without OTL-200 as the condition progresses. They have severe spasticity, seizures, poor gut motility making feeding difficult, and difficulty passing urine. They can become doubly incontinent, have breathing problems, scoliosis and little communication. The clinical experts explained that palliative care is difficult and complex because so many body systems are affected. The patient experts did not consider that the 'top-up' utility values accurately represented the condition. In response to the evaluation consultation document, the company provided a new scenario that applied a utility top up of 0.1 for patients in GMFC 3 to 6. The ERG reiterated its views that a top up is not necessary given the significant issues with the utility set and substantial utility benefits already claimed. The committee acknowledged the limitations of the original utility study and noted that the rescaled values did not address the methodological weaknesses. However, it considered that the rescaled utility set had better external validity relative to other appraisals (for example, HST12) than the other utility sets (company's original utilities or 'top-up' utility sets), and was an acceptable compromise. It noted that the most negative utility value for the rescaled set was more than that of some other comparable appraisals. But given the patient and clinical experts' statements

Distribution of subgroups

4.25 At the first committee meeting, the company presented a single pooled ICER. This weighted the individual subgroups of the MLD population by the distribution expected in clinical practice. The ERG highlighted that the company's modelled distribution of subgroups did not reflect known MLD epidemiology (see section 2.1). It amended the distribution based on epidemiological evidence and elicited clinical evidence. At the second committee meeting, the company agreed with the ERG's subgroup distribution. The committee considered that these issues make the pooled ICER uncertain:

- There are substantial differences in the cost-effectiveness estimates by subgroup. The clinical evidence suggests that people with ES-EJ MLD have much worse outcomes than other subgroups, so the ICERs are higher.

- The distribution of MLD subgroups in clinical practice is unknown and any assumptions based on the data are likely to be inaccurate.

- Given the very low patient numbers, modelled treatment response categories could be affected by individual patients or clinical decisions about treatment eligibility (see section 4.15).

- The pooled ICER is very sensitive to the distribution used.

  The committee concluded that:

- Any assumptions about the distribution of subgroups are likely to be inaccurate.

- The pooled ICER depends on how diagnosis might change in clinical practice in the future.

- The evidence for each of the EJ subgroups was extremely uncertain because of the low patient numbers.

- The response trajectories of the presymptomatic and early symptomatic EJ subgroups were likely different (see section 4.10).

- Ideally OTL-200 should be cost effective for all subgroups, to minimise the risk to the NHS.

Discount rate

4.26   The company considered that a 1.5% discount rate was appropriate because many people's condition stabilises in states with high motor and cognitive function. In response to the evaluation consultation document, the company stated that it 'believes that for presymptomatic patients the 1.5% discount rate is appropriate for decision making, as these patients have the potential to live in full or near full health for over 30 years'. NICE's methods guide states that 'In cases when 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) ... a discount rate of 1.5% for costs and benefits may be considered by the appraisal committee if it is highly likely that, on the basis of the evidence presented, the long-term health benefits are likely to be achieved. The appraisal committee will need to be satisfied that the introduction of the technology does not commit the NHS to significant irrecoverable costs.' The ERG also noted that myeloablative conditioning has a significant adverse event burden that would likely impact all patients. In addition, the ERG questioned whether it was appropriate to apply a differential discount rate across subgroups as proposed by the company. The committee recalled that in the company's own base case, less than 50% of the overall population were likely to have a full response (see section 4.10). There was substantial uncertainty about how long benefits of OTL-200 last (see section 4.13). It noted that OTL-200's cost is a single cost that could commit the NHS to significant irrecoverable costs. And there are also potential ongoing irrecoverable costs for patients who have OTL-200 and stabilise in worse health states for longer periods. So, the committee considered that the non-reference discount rate of 1.5% was not appropriate for decision making.

Applying QALY weighting

4.27   The interim process and methods of the highly specialised technologies programme (2017) specifies that a most plausible ICER of below £100,000 per quality-adjusted life year (QALY) gained for a highly specialised technology is normally considered to be an effective use of NHS resources. For a most plausible ICER above £100,000 per QALY gained, judgements about the acceptability of the highly specialised technology as an effective use of NHS resources must take account of the size of the incremental therapeutic improvement. This is revealed through the number of additional QALYs gained and by applying a 'QALY weight'. It is understood that a weight between 1 and 3 can be applied when the QALY gain is between 10 and 30 QALYs. The committee discussed the QALY gains with OTL-200, highlighting that they were highly uncertain and varied substantially between subgroups for the most plausible scenario (see section 4.28). The company considers the exact QALY gains to be commercial in confidence, so they cannot be reported here. Taking into account the incremental QALY gains with OTL-200, the committee concluded that it likely met the criteria for a QALY weight of between 1 and 3. But the exact weighting was uncertain and dependent on the MLD subgroup (see section 4.25).

The committee's preferred assumptions

4.28   In addition to the assumptions incorporated by the company after the first committee meeting (see section 4.19), the committee preferred:

- using the ERG's revised model structure and classification for OTL-200 response (see section 4.11)

- including a benefit for cognitive function separate from gross motor function in patients having OTL-200 (see section 4.10)

- using the same progression modifiers as those used for LI MLD for people with EJ who had an unstable response to OTL-200 (see section 4.20)

- including that OTL-200's effects are likely to be stable over an average period of 20 years (see section 4.21)

- including the company's rescaled utility set (see section 4.24)

- using individual ICERs rather than the pooled ICER because of differences in response trajectories, particularly between presymptomatic and early symptomatic EJ groups (see section 4.10 and section 4.25)

- including a discount rate of 3.5% for costs and benefits (see section 4.26).

Cost-effectiveness estimate

4.29   The committee considered the cost effectiveness of OTL-200 compared with best supportive care. It recognised the limited amount of evidence available, especially for the EJ subgroups (see section 4.15), and the uncertainty about how long response to OTL-200 lasts (see section 4.13). The committee examined the individual ICERs for each of the subgroups separately using the committee's preferences (see section 4.28). It considered that these were within the range NICE normally considers to be a cost-effective use of NHS resources for highly specialised technologies, taking into account a QALY weighting for each subgroup (see section 4.27). The ICERs are subject to a commercial arrangement and are commercial in confidence, so they cannot be reported here. The committee concluded that it could recommend OTL-200 as an option for treating MLD in children who have late infantile or early juvenile types, with no clinical signs or symptoms, and in children who have the early juvenile type, with early clinical signs or symptoms but who still have the ability to walk independently and have no cognitive decline.

Delivery of OTL-200

4.30   The committee recalled its concerns about the practicality of applying the eligibility criteria in the marketing authorisation, particularly for children in the ES-EJ subgroup (see section 4.5). The committee considered that the cost-effectiveness estimates relied on accurately identifying people who could benefit from the technology, before rapid progression of the condition. Therefore, it considered that treating MLD within the scope of the marketing authorisation was essential. It concluded that eligibility in relation to the marketing authorisation would most effectively be assessed by a multidisciplinary team in highly specialised services.

Impact of the technology beyond direct health benefits

4.31   The committee discussed OTL-200's effect beyond its direct health benefits, and the patient experts' statements. It was aware of the large impact of MLD on families, including the emotional effect on carers, siblings with the condition and other family members. It also noted the substantial financial impact on families, with parents possibly having to give up work to provide full-time care and adapt their home. Parents explained that OTL-200 had completely changed their experience of having children with MLD. This was because some children who have treatment remain healthy, and are able to live a normal life and attend mainstream school and activities. The committee considered that some of these aspects were included in the economic analysis. However, it recognised that the full effect of benefits beyond the direct health benefits had not been quantified. The committee considered the uncaptured benefits qualitatively in its decision making.

Other factors

Equality issues

4.32   The committee noted the potential equality issue with identifying patients with early symptomatic disease that may discriminate against people with learning disabilities. It noted that OTL-200's marketing authorisation states that patients should have treatment 'before the onset of cognitive decline' (see section 4.5). The committee considered the practicality of applying the IQ threshold of 85 or less for cognitive decline. The clinical experts explained that the threshold is there to identify a decline in cognitive function because of MLD, rather than to establish a strict IQ-based treatment criterion. The committee considered that it would be important to ensure that anyone with pre-existing learning difficulties would not be disadvantaged in accessing the technology by using this criterion. The clinical experts also noted other equality issues about speed of diagnosis that could affect access to early treatment. These included family background, socioeconomic status and geographical access to services. The committee acknowledged that some of these could be equality issues but did not consider that the guidance could resolve them.

Innovation

4.33   The committee acknowledged that OTL-200 is an innovative technology and
       represents a step change in managing MLD. It recalled the patient and clinical
       experts' statements that the technology is life transforming (see section 4.2). It
       considered that all the health benefits of OTL-200 were not likely to be captured in
       the economic model (see section 4.31).

5 Implementation

5.1 Section 8(6) of the [National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013](#) requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this evaluation within 3 months of its date of publication.

5.2 The Welsh ministers have issued directions to the NHS in Wales on implementing NICE highly specialised technologies guidance. When a NICE highly specialised technologies guidance recommends the use of a drug or treatment, or other technology, the NHS in Wales must usually provide funding and resources for it within 2 months of the first publication of the final evaluation document.

5.3 When NICE recommends a treatment 'as an option', the NHS must make sure it is available within the period set out in the paragraph above. This means that, if a patient has metachromatic leukodystrophy with mutations in the arylsulphatase A (ARSA) gene and the doctor responsible for their care thinks that atidarsagene autotemcel is the right treatment, it should be available for use, in line with NICE's recommendations.

6 Evaluation committee members and NICE project team

Evaluation committee members

The highly specialised technologies evaluation committee is a standing advisory committee of NICE.

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

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.

NICE project team

Each highly specialised 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

Adam Brooke
Technical adviser

Joanne Ekeledo and Daniel Davies
Project managers

ISBN: 978-1-4731-4496-5

Accreditation

NICE accredited

www.nice.org.uk/accreditation

Page 45 of
45